CHEMICAL ENGINEERING TRANSACTIONS

VOL. 33, 2013

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at: www.aidic.it/cet

Guest Editor: Enrico Zio
Copyright © 2013, AIDIC Servizi S.r.l.,
ISBN 978-88-95608-24-2; **ISSN** 1974-9791

Framework for a Hybrid Prognostics

Kamal Medjaher, Noureddine Zerhouni

FEMTO-ST Institute, UMR CNRS 6174 - UFC / ENSMM / UTBM, Automatic Control and Micro-Mechatronic Systems Department. 24, rue Alain Savary, 25000 Besançon, France.
kamal.medjaher@ens2m.fr

Fault detection and isolation, or fault diagnostic, of physical systems has been subject of several interesting works. Detecting and isolating faults may be convenient for some applications where the fault does not have severe consequences on humans as well as on the environment. However, in some situations, detecting and diagnosing faults may not be sufficient. In these cases, it is more interesting to anticipate the time of the fault, what is the purpose of prognostics. This latter activity aims at estimating the remaining useful life of systems by using three main approaches: data-driven prognostics, model-based prognostics and hybrid prognostics.
Data-driven prognostics concerns the transformation of the raw monitoring data to relevant models or trends which are then used to continuously assess the health state of the system and predict its remaining useful life. This approach is easy to implement, but suffers from precision. In addition, the method is applied in most cases on single physical components (bearings, gears, belts, etc.) and thus the interaction between the components of the whole system is not addressed. Model-based prognostics (also called physics of failure prognostics) deals with analytical modelling of the system including its degradation. This approach gives more precise results, but it is difficult to apply on complex systems for which the construction of the behaviour and degradation models is not a trivial task. Finally, the hybrid approach combines both model-based and data-driven approaches by keeping their advantages.
This paper presents a framework allowing the development of hybrid prognostics. The framework relies on two main phases: an offline phase and an online phase. The first phase concerns the construction of the nominal and degradation models, and the definition of the faults and performance thresholds needed to calculate the remaining useful life of the system. The second phase deals with the utilisation of the models and thresholds obtained in the first phase to detect the fault initiation, assess the current health state of the system, predict its future health state and calculate its remaining useful life.

1. Introduction

Fault Detection and Isolation (FDI), and fault prognostics of industrial systems are two necessary functions as they allow avoiding non-desirable situations and catastrophes. FDI can be applied on both abrupt and incipient faults. Several research and industrial works have been conducted in the domain. Interesting reviews related to FDI methods can be found in (Venkatasubramanian, 2005, Jardine et al. 2006, Samantaray and Ould Bouamama, 2008). The reported methods can be classified in two main categories: qualitative methods and quantitative methods. FDI can be used to do reconfiguration and accommodation and is suitable for systems where the fault does not have severe consequences. For example, detecting and isolating a fault on a valve controlling inflammable liquids may not avoid possible explosions. In this case, the fault is diagnosed *a posteriori* and thus is undergone.
Contrary to FDI, which is done a posteriori after the appearance of the faults, prognostics aims at anticipating the time of a failure by predicting the Remaining Useful Life (RUL) of the system (AFNOR, 2005). Prognostic results can then be used to take appropriate decisions on the system (change of set points, reduce the production load, stop the system, etc.).
Fault prognostic methods can be grouped in three main approaches: data-driven prognostics, model-based prognostics and hybrid prognostics. Interested readers can get more details on these three

approaches through the reviews published by Jardine et al. (2006), Heng et al. (2009a), Sikorska et al. (2011) and Tobon-Mejia et al. (2012).

Data-driven prognostics is based on the utilization of monitoring data to build behaviour models including the degradation evolution, which are then used to predict the RUL. Thus, Medjaher et al. (2012) proposed a method to estimate the RUL of critical components by using mixture of Gaussians hidden Markov models. Heng et al. (2009b) suggested a data-driven prognostic method based on the utilization of a feed-forward neural network. The training targets of this latter model are asset survival probabilities estimated using a variation of the Kaplan–Meier estimator and a degradation-based failure probability density function (PDF) estimator. Finally, Dong and He (2007) published a method based on a segmental hidden semi-markov model (hsmm) to do diagnostics and prognostics.

Model-based prognostics, also called physics of failure prognostics, uses models generated from fundamental laws of physics to calculate the RUL as suggested by Luo et al. (2003) and Chelidze and Cusumano (2004).

Finally, hybrid prognostics combines both previous approaches and can be considered as the one which gives the trade-off in terms of precision, applicability and complexity. Furthermore, the hybrid approach can be applied on physical systems rather than single physical components. This approach is for example easy to implement on mechatronic systems for which the construction of the behaviour and degradation models is possible. For this class of systems, the hybrid approach allows the estimation of the RUL of the whole system, this information can then be used to take appropriate decisions (reconfiguration of the mission, fault accommodation, etc.).

This paper presents a framework for the development of hybrid prognostics, which can be applied on multi-physical systems, particularly the mechatronic systems. The proposed framework is based on two phases: an offline phase and an online phase. The first phase concerns the construction of the nominal and degradation models of the systems. This phase concerns also the definition of the faults' thresholds and the system's performance limits needed to calculate its remaining useful life. The second phase uses the models and the thresholds obtained during the first phase to detect the fault initiation, assess the current health state of the system, predict its future health state and calculate its remaining useful life.

The paper is organized as follows. After the introduction, section 2 presents a brief comparison between the model-based and the data-driven approaches to introduce the motivation behind the proposition of a framework for a hybrid approach. Section 3 deals with the presentation of the proposed framework and finally, section 4 concludes the paper.

2. Model-based prognostics vs data-driven prognostics

Fault prognostics can be done according to three main approaches: data-driven prognostics, model-based (also called physics of failure) prognostics and hybrid prognostics. The first approach uses the data provided by sensors (monitoring data) and which capture the degradation evolution of the system. The data are then pre-processed to extract features which are used to learn models for health assessment and RUL prediction, as proposed by Heng et al. (2009b) and Dong and He (2007). Examples of models are neural networks, regressions, hidden Markov models, support vector regression, etc. The second approach requires a deep understanding of the physical phenomena of the system, including the degradation evolution. This approach uses physical laws to build the global model of the system, which is then used for simulations and predictions to calculate the RUL. In this approach Luo et al. (2003) developed a model-based prognostic method using a mathematical model including the degradation of a car suspension. Similarly, Chelidze and Cusumano (2004) proposed a method based on a dynamical systems approach applied to the problem of damage evolution in a two-well magneto-mechanical oscillator. Note that in model-based prognostics the construction of the model supposes the availability of a degradation model. Examples of degradation models are those related to crack by fatigue, corrosion and wear. A summary of the advantages and drawbacks of each approach is given in Figure 1. As mentioned in this figure, the data-driven approach gives less precise results than the model-based approach. The other drawback of data-driven prognostics is the variability of the data used to learn the degradation models. Indeed, to implement this approach, one needs to do several experiments to acquire data representing the behaviour of the degradation. But, in practice the data acquired for example for bearings having the same reference and degraded by using same operating conditions will vary. Thus, the model learned from these data is in fact a "mean" model and the RUL predictions obtained by its utilization will suffer from precision (the precision of the results depends strongly on the variability of the data).

Data-driven prognostics

- **Advantages**
  - Simplicity of implementation
  - Low cost
- **Drawbacks**
  - Need of experimental data that represent the degradation phenomena
  - Variability of test results even for a same type of component under same operating conditions
  - Less precision
  - Difficult to take into account the variable operating conditions
  - Component-oriented approach rather than system-oriented approach
  - Difficult to define the failure thresholds

Model-based prognostics

- **Advantages**
  - High precision
  - Deterministic approach
  - System-oriented approach: propagation of the failure in the whole system
  - The dynamic of the states can be estimated and predicted at each time
  - The failure thresholds can be defined according to the system's performance (stability, precision, …)
  - Possibility to simulate several degradations (drifts on the parameters)
- **Drawbacks**
  - Need of degradation model
  - High cost of implementation
  - Difficult to apply on complex systems

Figure 1: Data-driven prognostics vs Model-based prognostics

However, in terms of applicability, cost and simplicity of implementation, the practitioners prefer the data-driven methods. Indeed, in practice, the model-based methods are not easy to generalize on industrial systems due particularly to the difficulty to build the physical model of the systems including their degradation phenomena. The model-based methods can be applied on systems for which the models are known or for a class of systems such as the mechatronic ones. But even for these systems, it is necessary to do some experiments in order to catch the behaviour of the degradation which takes place in the system. Thus the implementation of a hybrid prognostics becomes a reality and allows taking benefit from both model-based and data-driven approaches.

3. Framework for a hybrid prognostics

The framework proposed in this paper applies on systems for which the construction of physical (or mathematical) behaviour models is possible. Also, the framework supposes that the system under study can be monitored by appropriate sensors in order to track the evolution of the degradation of its components. Furthermore, the proposed framework is system-oriented rather than component-oriented. This means that the whole system is considered for RUL prediction. For this purpose, we suppose that the variations (or drifts) in the parameters of the system are propagated to the whole system and are taken into account in the global dynamic model for simulations, predictions and RUL calculation.

The hybrid prognostics proposed in this framework is done in two main phases, as shown in Figure 2: an offline phase to build the dynamic model of the system and learn its degradation models, and an online phase (or exploitation phase) where the learned models are used to detect the initiation of the degradation and predict the RUL of the system.

The first phase concerns the construction of the nominal behaviour model of the physical system and the degradation models of its components. This phase concerns also the definition of the faults' and system's performance thresholds. The performance of the system can be defined by one of the following aspect: degree of precision, stability margins, magnitude of the faults, etc.

The second phase deals with the exploitation of the models learned or constructed previously to continuously assess the health state of the system, predict its future state and calculate its RUL. Note that in this framework, we suppose that the calculation of the RUL is triggered by the detection of the fault initiation.

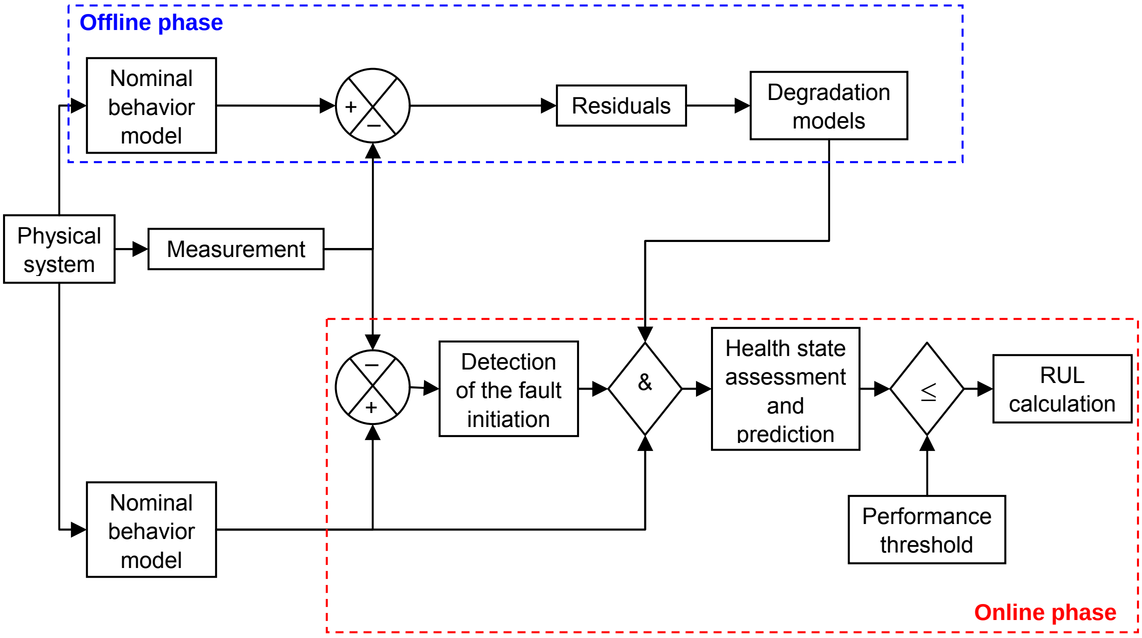

Figure 2: The detailed process of the proposed prognostic framework

Details on the steps of the two phases of the hybrid prognostic framework are given hereafter.
As stated previously, the first phase includes three main steps: the construction of the nominal behaviour model of the system, the generation of its degradation model and the definition of the thresholds (faults' thresholds and system's performance thresholds). The nominal model consists of a set of mathematical equations, which can be obtained by using the fundamental laws of physics or appropriate modelling formalisms and techniques such as bond graphs, which are well described in the following books (Karnopp et al., 2006) and (Samantaray and Ould Bouamama, 2008). The output of the nominal model is compared to the measurements acquired on the real system to generate residuals which are then used to build (or extract) the degradation models of the system's components.
A residual is a numerical evaluation of an Analytical Redundancy Relations (ARR) obtained from an over-determined system of equations (number of equations is greater than the number of variables) (Ould Bouamama et al., 2006). An ARR can represent a mass balance, energy balance, etc., and contains only known variables (inputs, outputs and parameters of the system). An ARR is given by the following expression:

$$ARR : \Phi(K) = 0 \tag{1}$$

Where K is a set of known variables or parameters of the system. A residual r(t) is a numerical evaluation of an ARR:

$$r(t) = \Phi(K) \tag{2}$$

Theoretically, the residuals should be equal to zero in the absence of faults and different from zero in the presence of faults. However, in practice and due to modelling and measurement noises, the value of each residual remains within a given threshold when there is no fault and exceeds the threshold otherwise.
In practice, several faults can occur simultaneously on the system. However, in this framework, only the single fault case is considered (several faults can't occur at a same time). This assumption is made to simplify the approach and proof its applicability. Other assumptions considered for the generation of the degradation models are given below.

1. Only incipient faults are considered (no abrupt faults). This assumption is due to the fact that this paper deals with fault prognostics.
2. A fault is a consequence of a change in a physical parameter of the system. Thus, any change in the physical parameter will affect the residuals in which this parameter appears.
3. The faults in the sensors are not considered. We suppose that the sensors are fault free and provide correct measurements.
4. The faults in the actuators are not taken into account.

Let α_1, α_2, α_3 … α_n be the set of physical parameters of the system which are involved in its dynamic model and in the corresponding residuals. The residual equation given in Eq.2 can then be re-written as follows:

$$r(t) = \Phi(\alpha_1,\ \alpha_2,\ \alpha_3,\ \dots\ \alpha_n) \tag{3}$$

Then, because only a single degradation can occur at a same time, its evolution can be determined by inverting Eq.3. For example, in the case where the degradation corresponds to the variation of the parameter α_1, its evolution can be calculated by the following equation:

$$\alpha_1 = \Phi^{-1}(r(t),\ \alpha_2,\ \alpha_3,\ \dots\ \alpha_n) \tag{4}$$

Note that during the offline phase, several experiments should be carried out to extract the degradation model (represented by the corresponding parameters) of each component of the physical system.

The second phase of the proposed approach concerns the exploitation of the models and knowledge obtained in the first phase to assess the health state of the system and calculate its RUL. During this phase, the output of the nominal behaviour of the system is continuously compared to the measurements provided by the sensors to detect whether the fault starts to occur or not. If a fault initiation is detected, the process of health assessment and RUL calculation is launched. The detection of a fault initiation is done by continuously evaluating the residuals and by analyzing the corresponding binary fault signature matrix formed by the residuals and given in table 1. In this matrix, each cell "i" contains "1" if the parameter α_i is present in the residual "i" and "0" otherwise.

Table 1: Example of a binary fault signature matrix

	α_1	α_2	α_3	…	α_n
Residual 1	1	1	0	…	0
Residual 2	0	0	1	…	0
…	0	1	0	…	0
…	0	0	0	…	0
Residual n	1	0	1	…	1

As depicted in Figure 2, the assessment of the system's health state is done by using its global model composed by the nominal model, the degradation model and the result of the fault detection step. Indeed, once the current fault initiation is detected, its corresponding degradation model, obtained in the offline phase, is replaced in the nominal model. Then, the updated whole model (the nominal model including the degradation model) is used to do predictions and calculate the RUL of the system. The RUL is calculated according to a defined performance, which can be the precision of the system, its stability, etc. (Figure 3).

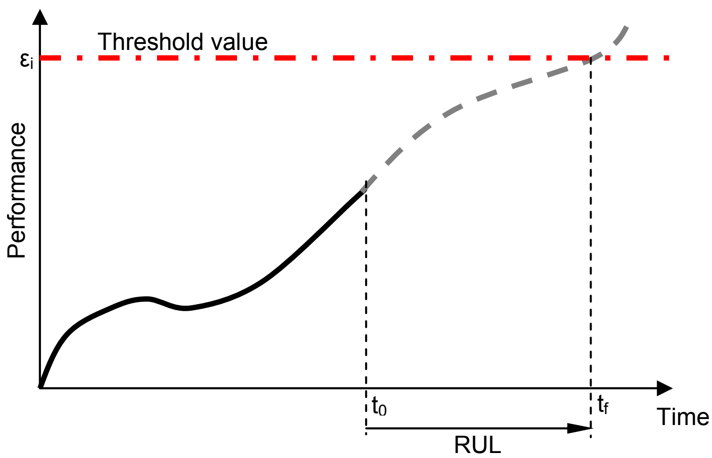

Figure 3: RUL calculation according to a given system's performance

4. Conclusion

A framework for a hybrid prognostics, with particular application to mechatronic systems is proposed in this paper. The framework is a system-oriented approach rather than a component-oriented. It has thus the interest of combining both model-based and data-driven approaches to take benefits from their advantages. The framework relies on two phases. The first phase concerns the generation of the dynamic and the degradation models and also the definition of the thresholds needed for the RUL calculation. The second phase deals with the exploitation of the knowledge gathered during the first phase to detect the initiation of the degradation, assess the health state of the system and predict its RUL. The predicted RUL can then be used to take appropriate decisions on the system. These decisions can be a reconfiguration of the mission to delay the fault, a preparation of the maintenance resources to repair the system and extend its utilization, the accommodation of the fault, etc.

The presented framework can be easily applied on mechatronic systems for which the construction of dynamic behavior models is possible. Furthermore, for this category of systems, the generation of the degradation models is also feasible by using the concept of residuals and by respecting the assumptions made in this contribution. Finally, the implementation of the framework on a real mechatronic system would proof its applicability and its effectiveness.

References

AFNOR, 2005, Condition monitoring and diagnostics of machines – prognostics - part 1: General guidelines, NF ISO 13381-1.

D. Chelidze, J. Cusumano, 2004, A dynamical systems approach to failure prognosis, Journal of Vibration and Acoustics, 126, 2 – 8.

M. Dong, D. He, A segmental hidden semi-markov model (hsmm)-based diagnostics and prognostics framework and methodology, 2007, Mechanical Systems and Signal Processing, 21, 2248–2266.

A. Heng, S. Zhang, A. C. Tan, J. Mathew, 2009a, Rotating machinery prognostics: State of the art, challenges and opportunities, Mechanical Systems and Signal Processing, 23 (3), 724 – 739.

A. Heng, A. C. Tan, J. Mathew, N. Montgomery, D. Banjevic, A. K. Jardine, 2009b, Intelligent condition-based prediction of machinery reliability, Mechanical Systems and Signal Processing, 23 (5), 1600 – 1614.

R. Isermann, Supervision: fault-detection and fault-diagnosis methods, 1997, An introduction, Control Engineering Practice, 5, 639–652.

A. K. Jardine, D. Lin, D. Banjevic, 2006, A review on machinery diagnostics and prognostics implementing condition-based maintenance, Mechanical Systems and Signal Processing, 20 (7), 1483 – 1510.

D. Karnopp, D. Margolis, R. Rosenberg, 2006, Systems Dynamics: Modeling and Simulation of Mechatronic Systems, 2nd Edition, John Wiley, New York.

J. Luo, K. R. Pattipati, L. Qiao, S. Chigusa, 2003, Model-based prognostic techniques applied to a suspension system, Transactions on Systems, Man, and Cybernetics, 38, 1156–1168.

K. Medjaher, D. A. Tobon-Mejia, N. Zerhouni, 2012, Remaining useful life estimation of critical components with application to bearings, IEEE Transactions on Reliability, 61 (2), 292 – 302.

B. Ould Bouamama, K. Medjaher, A. Samantaray, M. Staroswiecki, 2006, Supervision of an industrial steam generator. part I: Bond graph modelling, Control Engineering Practice, 14 (1), 71–83.

A. Samantaray, B. Ould. Bouamama, 2008, Model-Based Process Supervision: A Bond Graph Approach, Springer.

J. Sikorska, M. Hodkiewicz, L. Ma, 2011, Prognostic modelling options for remaining useful life estimation by industry, Mechanical Systems and Signal Processing, 25, 1803 – 1836.

D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, 2012, CNC machine tool's wear diagnostic and prognostic by using dynamic bayesian networks, Mechanical Systems and Signal Processing, 28, 167 – 182.

V. Venkatasubramanian, 2005, Prognostic and diagnostic monitoring of complex systems for product lifecycle management: Challenges and opportunities, Computers & Chemical Engineering, 29 (6), 1253 –1263.