CHEMICAL ENGINEERING TRANSACTIONS
VOL. 33, 2013
Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791
A publication of
AIDIC
The Italian Association
of Chemical Engineering
Online at: www.aidic.it/cet
DOI: 10.3303/CET1333052

Risk Sensitive Particle Filtering in Support of Predictive Maintenance

Michele Compare[a], Enrico Zio*[,a,b]

[a]Energy Department, Politecnico di Milano, Milan, Italy
[b]Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, France
enrico.zio@polimi.it;

Predictive Maintenance (PrM) exploits the estimation of the equipment Residual Useful Life (RUL) to identify the optimal time for carrying out the next maintenance action. Particle Filtering (PF) is widely used as prognostic tool in support of PrM, by reason of its capability of robustly estimating the equipment RUL without requiring strict modeling hypotheses. However, a precise estimate of the RUL requires tracing a large number of particles, and thus large computational times, often incompatible with the need of rapidly processing information for making decisions in due time. To circumvent this problem, the Risk Sensitive Particle Filtering (RSPF) technique is exemplified in this work by way of a case study concerning a mechanical component affected by fatigue degradation.

1. Introduction

In recent years, the relative affordability of on-line monitoring technologies has led to a growing interest in new maintenance paradigms such as PrM. This is founded on the possibility of monitoring the equipment to obtain information on its conditions, which is then used to identify problems at an early stage and predict their evolutions in the future for estimating the equipment RUL. An accurate estimation of the RUL is of great interest, as it would provide lead time to opportunely plan, prepare and execute the repair or the replacement of the equipment, e.g., by delaying the maintenance to the next planned plant outage, by provisioning with spare parts only at time of necessity, by optimizing staff utilization, while remaining acceptably confident that the system will not fail before maintenance and the equipment lifetime will be fully exploited (e.g., Zio and Compare, 2013).
A number of prognostics approaches have been proposed in the literature in support of PrM (Zio, 2013). Among these, PF is emerging as a powerful model-driven technique, capable of robustly predicting the future behavior of the distribution that describes the uncertainty in the actual degradation state of the equipment (e.g., the crack depth of a mechanical component, Cadini et al., 2009). From the prediction of the future evolution of the degradation and knowledge of the failure threshold (i.e., the degradation value beyond which the equipment loses its function), one can infer the equipment RUL.
Although “PF-based prognostic algorithms have been established as the de facto state of the art in failure prognostics” (Orchard et al. 2009), their application may be impaired in contexts where precise and/or conservative estimations of the RUL are mandatory, because these estimations require large computational times. Indeed, the RUL is related to a failure event, which is generally associated to particles located at the tails of the predicted distributions of the degradation state. This entails that the estimation of the failure probability is more sensitive to the imprecision due to the approximate particle representation. The large number of particles that PF needs to trace for providing robust estimates, heavily affects the computational expenses.
A solution to this problem is proposed in this work, based on the RSPF algorithm (Thrun et al. 2001). It is exemplified on a case study taken from literature, concerning crack growth degradation.
The proposed scheme of RSPF algorithm extends that proposed by Orchard et al. 2009, which is acts just on the sampling distribution without biasing the posterior one, describing the uncertainty in the degradation state.
The remainder of the paper is organized as follows: Section 2 briefly describes the main characteristics of PF-based algorithms in PrM applications, and the idea beyond the RSPF. Section 3 introduces the

Please cite this article as: Compare M., Zio E., 2013, Risk sensitive particle filtering in support of predictive maintenance, Chemical Engineering Transactions, 33, 307-312 DOI: 10.3303/CET1333052

reference example; the application of the RSPF to such example is discussed in Section 4; Section 5 concludes the work.

2. PF-based prognostics form PrM

PF for prognostics is based on (e.g., Orchard 2007, Vachtsevanos et al. 2006):

1. a degradation model describing the stochastic evolution of the equipment degradation x in discretized time instants $k = 1, 2, ...$:

$$x_{k+1} = g(x_k, \omega_k) \tag{1}$$

where g is a possibly non-linear function and ω_k is a possibly non-Gaussian noise.

1) A set of measures $z_1,...,z_k$ of past and present values of some physical quantities z related to the equipment degradation x.
2) A probabilistic measurement model that links z with the equipment degradation x:

$$z_k = h(x_k, \nu_k) \tag{2}$$

where h is a possibly non-linear function and ν_k is the measurement noise.

Briefly, in PF a set $X_k = \{x_k^i\}$, $i = 1,..., N_s$, $k = 1,2,...$ of weighted particles is considered, which evolve independently on each other, according to the probabilistic degradation model of Eq. 1. The basic idea is that such set of weighted random samples constitutes a discrete approximation of the true pdf of the system state at time k. When a new measurement z_k is collected, the predicted pdf $P(x_k | x_{k-1}, z_1,..., z_{k-1})$ is adjusted through the modification of the weights of the particles, thus yielding the posterior distribution $P(x_k | z_1,..., z_k)$. This step requires the knowledge of the probabilistic law which links the state of the equipment to the gathered measure (Eq. 2). From this model, the probability distribution $P(z | x)$ of observing the sensor output z given the true degradation state x is derived (measurement distribution), and used to update the weights of the particles upon a new measurement collection (see Appendix). Roughly speaking, the smaller the probability of encountering the acquired measurement value, when the actual equipment state is that of the particle, the larger the reduction of the particle weight in the posterior distribution. On the contrary, a good match between the acquired measure and the particle state results in an increase of the particle importance (for further details, see Arulampalam et al. 2002).

Notice that in this work, PF will always refer to its Sampling Importance Resampling (SIR) version (see Arulampalam et al. 2002). This algorithm allows avoiding the degeneracy phenomenon (i.e., after few iterations, all but few particles would have negligible weights), which is typical of the standard version of PF (i.e., Sequantial Importance Sampling, SIS).

Finally, in the present work the PF-based prognostic model is embedded within the PrM scheme as follows:

- Degradation state can be known only via measurements, which are periodically taken with period *II* (e.g., *II*=20 h).
- When a measure is acquired, the estimation of the distribution of the current degradation state is updated, and the future evolution of such distribution is simulated. This allows to estimate the RUL, defined as the difference between time instant t_{PT} at which the failure probability reaches a pre-fixed threshold value P_T (e.g., P_T = 1%) and the current time instant.
- The preventive replacement action is performed either when the estimated RUL is elapsed or at the acquisition of a measurement if the updated distribution leads the equipment failure with a probability larger than P_T.

2.1 The RSPF algorithm

In the considered PrM setting, we are asked to estimate at a given time instant k, the equipment failure probability at the time instants $k^* = k+1, k+2,....$ To this aim, the stochastic behavior of the N_s particles from instant k is simulated according to the stochastic model in Eq. 1, thus providing an estimation of the distribution $P(x_{k^*} | z_1,..., z_k)$. The failure probability P_T is given by the area of the tail of $P(x_{k^*} | z_1,..., z_k)$ that crosses the failure threshold d (filled area in Figure 1). When we are interested in estimating small values of P_T (e.g., P_T = 1%, 5%), PF necessitates handling an appropriate number N_s of particles, otherwise it gives non conservative values. This is due to the fact that PF provides a particle-based, approximate pdf representing the uncertainty in the state x; the smaller the number of particles N_s, the worse the approximation.

To clarify this issue, let us assume that P_T = 1% (this may be the case where failures are associated with very high costs); then, PF needs tracing at least N_s =100 particles to (roughly) estimate $t_{0.01}$ as the time instant at which there is at least 1 particle above d. If N_s is smaller than 100, then the failure probability value is inevitably underestimated. For example, if we consider N_s =20, the time instant $t_{0.01}$ corresponds at least to $t_{0.05}$. Obviously, considering N_s larger than 100 (e.g., 500), would provide a more robust estimation of $t_{0.01}$, because a larger number of particles would exceed the threshold (e.g., 5).

The larger computational times needed to run the PF algorithm with greater values of N_s, may be in contrast with the requirement of estimating the equipment RUL in due time to anticipate reaction to failures to solve problems in advance.

RSPF offers a solution for this. The idea is to generate particles in such a way that the risk associated to the errors coming from the approximate particle representation is embedded in the representation itself. In details, instead of the posterior distribution $P(x_k \mid z_1,...,z_k)$ the RSPF tracks:

$$\gamma_k r(x_k) P(x_k \mid z_1,...,z_k) \qquad (3)$$

where $r(x_k)$ is a positive and finite (almost everywhere) risk function, whereas γ_k is a normalization constant ensuring that $\gamma_k r(x_k) P(x_k \mid z_1,...,z_k)$ is a probability distribution.

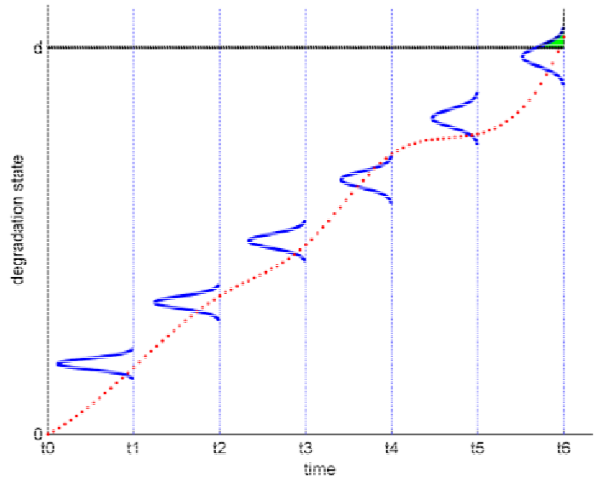

Figure 1: Estimation of the failure probability; dotted line: true equipment degradation behaviour (in arbitrary unit); continuous lines: pdfs describing the uncertainty on x at the corresponding time instants (in arbitrary unit of time). Dashed line: failure threshold

From Eq. 3, it clearly appears that the probability that a particle x_k^i belongs to the set X_k depends not only on the posterior distribution, but also on the risk associated with that sample. In turn, the effect of the risk function is to force the algorithm to sample from the regions of the sample space with higher risk values.

From these considerations, it emerges that not all risk functions will be equally useful to successfully apply RSPF, and identifying a proper shape of $r(x_k)$ is a fundamental issue. In this work, we derive the expression of $r(x_k)$ on the basis of the following consideration: the risk associated with the generic weighted particle x_k^i is strictly related to two main aspects: the closeness of x_k^i to the threshold and the uncertainty in its next *ll* future steps. The larger the distance of x_k^i from the threshold d, the smaller the probability of achieving it within the next *ll* steps, whereas the closer the particle to d, the smaller the probability of surviving *ll* steps. These two factors influence the RUL estimation, and the corresponding maintenance decision, which comes from the rules:

i) If RUL<=*ll*, then a preventive replacement of the equipment is done. The risk associated to this choice is that the equipment lifetime is not fully exploited, since it would be removed at least *ll* steps before the optimal replacement time. Such preventive action is supposed to last D_P hours.

ii) If RUL >II, then delay the decision to the next measurement acquisition. Thus, the risk associated to this choice is that the equipment fails before the next *II* steps. In this case, a corrective replacement is performed, which lasts D_C hours.

In this work, we assume that the risk associated to decision i) is negligible with respect to that relevant to ii). Thus, the aim of RSPF is to generate samples from a distribution that factors-in the risk $r(x_k)$ of not removing the equipment before failure.

The shape of $r(x_k)$ is derived by resorting to the Monte Carlo method. Namely, the sample space [0,d] is discretized into small bins. A large number M of particles are positioned in every bin, and their behaviors are simulated and traced for II steps. The portion of the M particles that cross the threshold provides an estimate of the probability for a particle located at x^i of exceeding the threshold d. Such probability value $r(x^i)$ is indicative of the risk associated to the starting degradation state x^i: multiplying it by the failure cost yields the expected cost associated with the decision of leaving the equipment working when its degradation state is x^i. Notice that from the pseudo-code given in Appendix, it clearly appears that accounting for the cost factor in $r(x_k)$ is useless, being the biased sampling dependent on the ratio $r(x_k^i)/r(x_{k-1}^i)$.

3. Reference Example

In this Section, the example concerning the fatigue degradation process of a mechanical component is presented. The randomized Paris – Erdogan model (e.g., Cadini et al., 2009) is used to describe this degradation mechanism. Its most widely used expression is:

$$x_k = x_{k-1} + e^{\xi_k} \cdot C \cdot (\beta\sqrt{x_{k-1}})^n \cdot \Delta t \quad (4)$$

where x_k is the crack depth at the (discretized) k^{th} time instant, β, C and n are constant parameters; $\xi_k \sim N(0,\sigma_\xi)$, $k = 1,\ 2,...$, are independent and identically distributed random variables, whereas Δt is a sufficiently small time interval (in our calculations, Δt =1h).

The model (1) makes x_k, $k = 1,\ 2,...$, a Markov process with independent increments, which are log-normally distributed. The values of the model parameters, taken from (Cadini et al. 2009), are reported in Table 1.

With regards to the measurement model, the uncertainty about the observations z_k is described by a white Gaussian noise $\upsilon \approx N(0,\sigma_\upsilon^2)$; this enters the physical law that links z_k to the depth x_k, leading to the following conditional pdf:

Table 1: Parameters of the reference example

Paris Erdogan Model		Measurement Model		Maintenance Model	
Parameters	Values	Parameters	Values	Parameters	Values
C	0.005	β_0	0.06	II	20h
β	1	β_1	1.25	D_P	100h
n	1.3	σ_υ	0.47	D_C	200h
Δt	1	d	100 mm	C_P	$\propto D_P$
σ_ξ	1.7			C_C	$\propto D_C$

$$p(z_k \mid x_k) = \frac{1}{\sqrt{2\pi\sigma_\upsilon^2}} \cdot e^{-\frac{\left(\ln\left(\frac{z_k}{d-z_k}\right)-\mu_k\right)^2}{2\sigma_\upsilon^2}} \cdot \frac{d}{(d-z_k)z_k} \quad (5)$$

where:

$$\mu_k = \beta_0 + \beta_1 \cdot \ln\left(\frac{x_k}{d-x_k}\right) \quad (6)$$

where β_0 and β_1 are parameters to be estimated from experimental data, and d is the component thickness. The values of the variables of both the measurement model (Cadini et al., 2009) and

maintenance model are listed in Table 1. In regard to the costs of the maintenance actions, these are assumed proportional to the component unavailability. This situation is typical of the plants where the main maintenance costs are related to the business interruption (e.g., energy production plants).
Finally, notice that the reference mission time *T* does not end with the failure of the component; it is set to *T* = 10,000 h. This entails that the performances of the maintenance strategies are assessed with regards to the transient period, rather than to the steady state (Zio and Compare 2013).

4. Application of RSPF

In this Section, the PrM policy described in Section 2 is applied to the case study at hand. First, Figure 2 shows the risk function $r(x_k)$, derived according to the procedure proposed in Section 2.1. This function is used to sample more particles with larger crack length (i.e., in the risky region of the interval [0, d]).

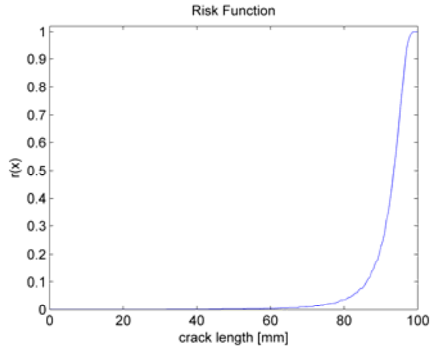

Figure 2: Risk function relevant to the case of the crack growth mechanism

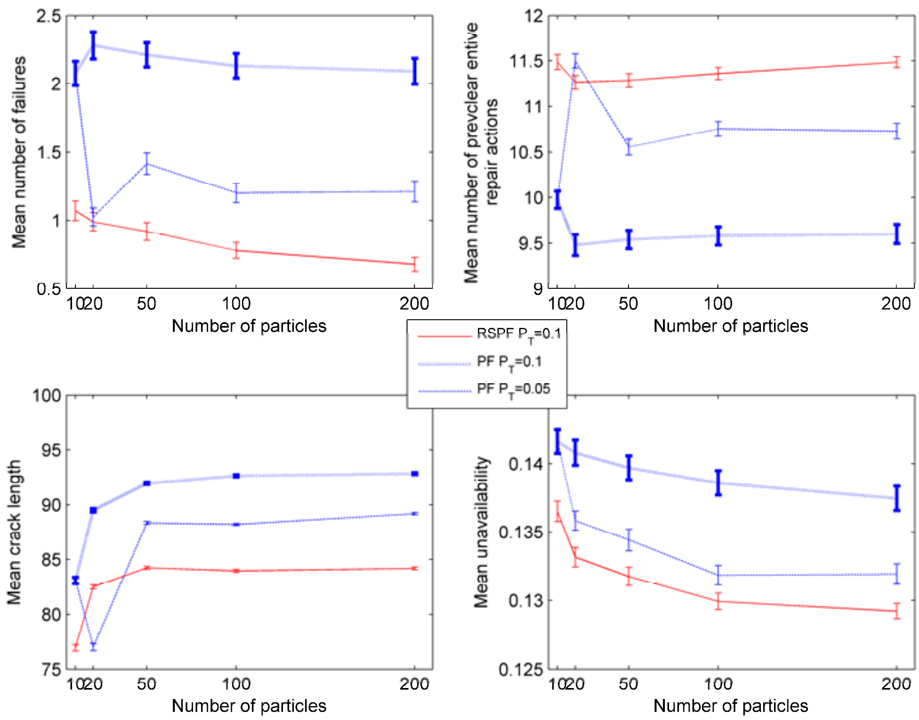

Figure 3: Results of the study

Figure 3 summarizes the results of the study. With regards to the performance of the RSPF when PT =0.1 (continuous line), it can be noted that larger values of N_s correspond to smaller values of the component unavailability. This is due to the fact that a more precise approximation of the biased distribution in Eq. 3 allows effectively avoiding component failures (Figure 3 (a) and (b)). This effectiveness in avoiding failures is counterbalanced by the poor exploitation of the component life (i.e., the mean crack length of the preventively replaced components decreases (Figure 3(c)). For comparison, the performance of the classical PF algorithm are also reported in Figure 3 (dotted lines). These are worse than those of the RSPF.

Finally, Figure 3 shows also the results relevant to the application of the classical PF when PT = 0.05. The maintenance performance in this setting is similar to that of the classical PF when N_S=10. This confirms that when the PT value is small, then PF tends to overestimate the RUL, unless a proper number of particles is tracked. Thus, reducing the value of PT is not a feasible way to reduce the number of failures when a small number of particles is handled by PF.

5. Conclusions

In this work, PrM maintenance has been considered, in combination with RSPF based prognostics. It has been shown that this combination allows considering a number of particles smaller than that required by the classical PF. The risk function considered accounts for the cost associated to the component failure only, whereas it neglects that of under-exploitation of the component preventively maintained before the optimal time. In this respect, the investigation of more complete risk functions is an issue worth of investigations in future works. Moreover, the performance of the PrM policy has been assessed in the 'nominal' setting, only. A further analysis needs to be carried out to understand how and to which extent the performance values are sensitive to the model parameters (e.g., II, ratio between the cost of the preventive and corrective actions, etc.).

References

Arulampalam M.S., Maskell S., Gordon, N. and Clapp T., 2002, A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, IEEE Transactions on Signal Processing, 50 (2), 174-188.

Cadini F., Zio E. Avram D., 2009, Model-based Monte Carlo state estimation for condition-based component replacement, Reliability Engineering and System Safety, 94 (3), 752-758.

Orchard M.E., 2007, A Particle Filtering-based Framework for On-line Fault Diagnosis and Failure Prognosis, PhD thesis, Department:of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. .

Orchard M.E., Tang L., Goebel K. and Vachtsevanos G., 2009, A novel RSPF approach to prediction of High-Risk, Low-Probability failure events, in proceedings of the Annual conference of the Prognostics and Health Management Society 2009, September 27-October 1, San Diego, CA, USA. Available on line at: http://www.phmsociety.org/events/conference/phm/09/proceedings.

Thrun S., Langford J., and Verma, V., 2001, Risk Sensitive Particle Filters, in Dietterich, Becker, Ghahramani (Eds.). Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada]. MIT Press 2001, 961-968.

Vachtsevanos G., Lewis F.L., Roemer M., Hess A. and Wu B., 2006, Intelligent Fault Diagnosis and Prognosis for Engineering Systems, John Wiley & Sons, Chichester, UK.

Zio E. and Compare M., 2013, Evaluating maintenance policies by quantitative modeling and analysis, Reliability Engineering and System Safety, 109, 53-65.

Zio E., 2013, Prognostics and health management of industrial equipment, in S. Kadry (Ed.), Diagnostics and Prognostics of Engineering Systems: Methods and Techniques (pp. 333-356). Hershey, PA.

Appendix

Pseudo-code of the RSPF algorithm (adapted from Thrun et al. 2001)

set $X_k = X_k^{aux} = \varnothing$

for $i = 1, ..., N_s$

pick the i-th sample $x_{k-1}^i \in X_{k-1}$; draw $x_k^i \sim p(x_k \mid x_{k-1}^i)$; set $w_k^i \sim p(z_k \mid x_k^i) \cdot r(x_k^i) / r(x_{k-1}^i)$;

add x_k^i to X_k^{aux} with weigh w_k^i

endfor

for $i = 1, ..., N_s$ %SIR algorithm

draw x_k^i from X_k^{aux} with probability proportional to w_k^i ; add x_k^i to X_k with $w_k^i = 1 / N_s$

endfor