



HAL
open science

Hidden Markov Models for diagnostics and prognostics of systems under multiple deterioration modes

Thanh Trung Le, Florent Chatelain, Christophe Bérenguer

► **To cite this version:**

Thanh Trung Le, Florent Chatelain, Christophe Bérenguer. Hidden Markov Models for diagnostics and prognostics of systems under multiple deterioration modes. ESREL 2014 - 24th European Safety and Reliability Conference, Sep 2014, Wroclaw, Poland. pp.1197-1204. hal-01027509

HAL Id: hal-01027509

<https://hal.science/hal-01027509>

Submitted on 22 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hidden Markov Models for diagnostics and prognostics of systems under multiple deterioration modes

Thanh Trung LE, Florent CHATELAIN & Christophe BERENGUER

Univ. Grenoble Alpes, GIPSA-lab, F-38000 Grenoble, France

CNRS, GIPSA-lab, F-38000 Grenoble, France

ABSTRACT: Multi-state systems have recently attracted a great deal of interest with regards to reliability and maintenance. Since most mechanical equipment operates under some sorts of stress or load, it tends to degrade over time, thus possibly resulting in discrete degradation states (damage degrees), ranging from perfect functioning to complete failure. Over recent years, Hidden Markov Models (HMMs) have been applied to model these discrete degradation states for diagnostic and prognostic purposes. However, most of the reported researches on HMMs for multi-state equipment in the literature consider only one degradation mechanism of degradation processes. The present paper proposes a novel model called multi-branch HMM (MB-HMM) to deal with deterioration processes modeling under multiple competing modes. To illustrate the proposed approach, a numerical study is given.

1 INTRODUCTION

Condition-based Maintenance (CBM) plays nowadays an important role in the maintenance of engineering systems since it has the ability to recommend in advance proper maintenance actions by basing on the information collected from condition monitoring (Jardine et al. 2006). Within a CBM program, diagnostics and prognostics are two important aspects where the former deals with fault detection, isolation and identification when it occurs and the latter helps to assess the current health state of the system and predict the remaining time before a failure occurs. An accurate prediction of the remaining useful life (RUL) could provide ample time for maintenance engineers to schedule a repair, and to acquire replacement components before they actually fail (Zhang et al. 2005). Due to the stochastic nature of the deterioration, stochastic processes have been adapted to model the deterioration processes and have shown the promising results within the CBM framework (Van Noortwijk 2009, Le Son et al. 2012, Wang and Wang 2012).

As a stochastic process, Hidden Markov Model (HMM) has been successfully applied in many domains, such as Speech recognition, genes and Deoxyribonucleic acid (DNA) analysis (Rabiner 1989) thanks to its strong mathematical basis. In the CBM context, HMMs can divide equipment conditions into several meaningful states, such as "good", "minor defect", "maintenance required" and "failure" and

therefore easy to understand (Si et al. 2011). They can also characterize the stochastic relationship between the features extracted from condition monitoring data and the actual health states of the equipment. For these reasons, HMMs, in the recent year, are being more and more investigated to be used as an efficient tool for the deterioration modeling and the RUL estimation (Wang and Wang 2012). For example, Baruah *et al.* employed HMMs for carrying out both the diagnostics and the prognostics for metal cutting tools (Baruah and Chinnam 2005). In (Zhang et al. 2005), the authors used the HMMs to construct a health/degradation index representing the current system health status. This index was then fed into an adaptive prognostic scheme in order to estimate the RUL of a bearing. In order to better represent the real degradation mechanism, several extensions of HMMs have been adopted in the literature, such as Mixture of Gaussian HMM (Tobon-Mejia et al. 2012), age-based HMM (Peng and Dong 2011), segmented HMM (Geramifard et al. 2012), segmented Hidden Semi-Markov Model (Dong and He 2007), etc.

To our best knowledge, almost all of the HMM-based frameworks in the literature consider only the mono-mode case, meaning that there is only one degradation mode occurring at one time. In practice, however, multiple degradation modes could exist in competition. In this paper, we propose a new HMM-based structure called multi-branch Hidden Markov Model (MB-HMM) to deal with this problem. In order to evaluate the effectiveness of the

proposed model with respect to a traditional mono-branch HMM, a simulation study is implemented.

The remainder of the paper is organized as follows. Section 2 is devoted to review some basic theories of HMMs. Section 3 describes the extension to MB-HMM and its application for diagnostics and prognostics. To illustrate the methodology, a comparative study is introduced in the section 4. Conclusion follows in Section 5.

2 METHODOLOGY

2.1 Basic theory of Hidden Markov Model

Hidden Markov Model is an extension of the Markov chain in which the state process are latent and can be only revealed through an observation process. This is where the word "hidden" comes from. In the deterioration modeling framework, the hidden state process represents the health states of the equipment, while the observations can be measurable signals such as the vibration signals or the features extracted from condition monitoring data. The relation between these two processes is represented by a probabilistic model. Figure 1 illustrates an example of an HMM model.

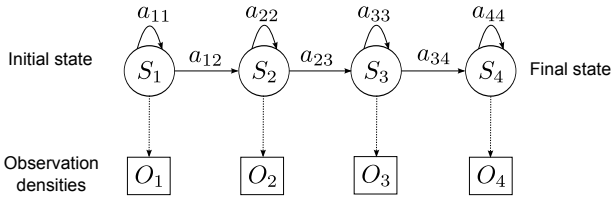


Figure 1: A five state left-right HMM

Mathematically, the complete specification of an HMM consists of the following elements (Rabiner 1989)

- N , the number of states in the model
- M , the number of distinct observation symbols per state, i.e., the discrete alphabet size
- A finite set of hidden states, i.e., $S = \{S_1, S_2, \dots, S_N\}$
- A state transition probability distribution, i.e., $A = \{a_{ij}\}$, where $a_{ij} = P(q_{t+1} = S_j | q_t = S_i)$
- An observation symbol probability distribution, i.e., $B = \{b_j(k)\}$ where $b_j(k) = P(Y_k | q_t = S_j)$, $1 \leq k \leq M$
- An initial state probability distribution, i.e., $\pi = \{\pi_i\}$ where $\pi_i = P\{q_t = S_i\}$, $1 \leq i \leq N$

For convenience, in this study, an HMM is represented by a compact notation $\lambda = (A, B, \pi)$.

2.2 HMM topology selection

In order to apply the HMM model for the deterioration modeling framework, one question is naturally raised: Given the observations, how one can choose/select a proper topology of the model. For instance, how can we choose the type of the model, how many states should be used or which observation probability distribution is suitable? These questions will be answered in this sub-section.

2.2.1 HMM type selection

There are several types of HMMs existing in the literature, such as the ergodic HMM or the left-right (Bakis) model (Rabiner 1989). Since the deterioration processes are often irreversible in reality, components cannot come back from a bad health state to a better one as time progresses. For this reason, the left-right HMM model is chosen to be investigated. More precisely, the used left-right model in this study has the characteristic that the state index can either stay the same or increase by one. In other words, the components can only stay in the current health state or move to the next one. In this case, the coefficients of the transition matrix of the model satisfy the following properties:

$$\begin{aligned} a_{ij} &= 0, & j < i \\ a_{ji} &= 0, & j > i + 1 \\ a_{NN} &= 1, \\ a_{Ni} &= 0, & i < N \end{aligned}$$

The state S_N is considered as the final (failure) state.

2.2.2 Observation model selection

The observations are usually considered as discrete symbols, hence discrete probability densities are used to model the observation probabilities. However, in the framework of machine condition monitoring, the observations obtained are typically continuous signals (Tobon-Mejia et al. 2012). In this case, the continuous HMM possessing a continuous probability density function (pdf) to model the emission probabilities could give a better accuracy in diagnostics, compared to the discrete HMM (Wang and Wang 2012). However, to ensure that the parameters of the HMM can be estimated in a convenient way, the observation pdf should be elliptically symmetric densities or mixtures of elliptically symmetric pdf (Rabiner et al. 1985). For this reason, the Gaussian pdf is used in this study as the observation model:

$$b_j(O) = \mathcal{N}(O, \mu_j, \Sigma_j), \quad 1 \leq j \leq N$$

where μ_j and Σ_j are the mean vector and covariance matrix in state S_j .

2.2.3 Number of states selection

Selecting the number of states in an HMM model is always an interesting but difficult question. It exists several methods as well as criteria for this selection purpose, such as Akaike information criterion (AIC), Bayes information criterion (BIC), Cross-validation, etc. (Knoblauch 2004). In this study, we investigate the use of BIC for selecting the number of state N .

The Bayes information criterion (BIC) is defined as follow

$$BIC = -2\mathcal{L} + k\log N \quad (1)$$

where \mathcal{L} is the log-likelihood function; k is the number of free parameters to be estimated and N is the number of data points in observation sequences \mathbf{O} . The chosen model will be the one having the minimum BIC value.

3 EXTENSION TO MULTI-BRANCH HIDDEN MARKOV MODEL

3.1 General principles

The idea of the multi-branch HMM have been appeared in the literature and have mainly been applied for the applications of handwriting recognition (Lee, Kim, & Kim 2001, Wang, Brakensiek, Kosmala, & Rigoll 2001). In this section, we will present the general principle of MB-HMM for modeling degradation processes under different modes. An example of a two-branch HMM is presented in the figure 2.

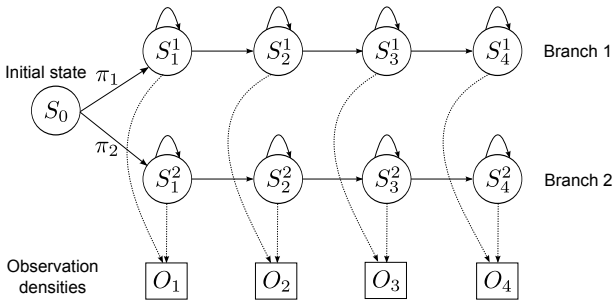


Figure 2: Example of a two-branch HMM

Such a multi-branch model is suitable to model the equipment deterioration process which could follow different modes due to different changes of the operating condition. Considering the propagation of the crack appearing in a bearing. In a normal condition, the bearing operates normally. However, when more stress is applied on it, the crack could initiate and then propagate with different rates. The more stress applied, the more quickly the crack propagates. In the MB-HMM model, each constituent branch is dedicated to model a different deterioration mode. It is worth noting that the observations emitted from the corresponding states in different constituent branches could be similar (e.g. having the same probability distributions). The only thing that distinguishes the

branches is the rate of deterioration which is represented by the transition probabilities.

There are some important assumptions that must be made in constructing the model. Firstly, once the equipment has degraded following one mode, it will follow that mode until it reaches the end of its life. There is hence no connection between the branches. The second assumption is that, the normal condition can last very long with no special characteristics in the measurement. We are interested in the deterioration process, with the purpose of assessing the deterioration level of equipment as well as predict its residual lifetime when it has suffered a problem. For this reason, the initial state can be assumed as a dummy state with no emission of observation. Each branch can therefore be seen as a single model and is assigned an *a priori* probability π_k as shown in the figure 2. These *a priori* probabilities satisfy:

$$\pi_k = P(\lambda_k) = K_k/K, \quad \sum_{k=1}^M \pi_k = 1 \quad (2)$$

where λ_k denotes the constituent HMM for branch k , K_k is the number of training data sequences corresponding to mode k , K is the total number of training data sequences and M is the total number of branches of the MB-HMM.

3.2 MB-HMM for diagnostics and prognostics

The application of MB-HMM for diagnostics and prognostics can be implemented in two phases, as shown in figure 3.

In the off-line phase, the training data is classified into M different groups and then, each constituent HMM model is trained individually with the data in the corresponding group. At this step, the well-known Baum-Welch algorithm can be used without any modifications. In the next step, the number of data in each group is used to assign the *a priori* probabilities for each model (c.f equation (2)). Finally, a trained MB-HMM model is obtained.

In the online phase, the diagnostics involves two sessions. In the first session, the mode that the equipment is following must to be identified. This goal can be achieved by using the decision theory. For instance, the model that has the maximum *a posteriori* probability, given the data, can be chosen in order to minimize the misclassification rate, (Bishop et al. 2006). That is:

$$\hat{k} = \arg \max_k P(\lambda_k | \mathbf{O}) \quad (3)$$

Note that, using Bayes' theorem, these *a posteriori* probabilities can be calculated as follows:

$$P(\lambda_k | \mathbf{O}) = \frac{P(\mathbf{O} | \lambda_k) P(\lambda_k)}{\sum_{k=1}^M P(\mathbf{O} | \lambda_k) P(\lambda_k)} \quad (4)$$

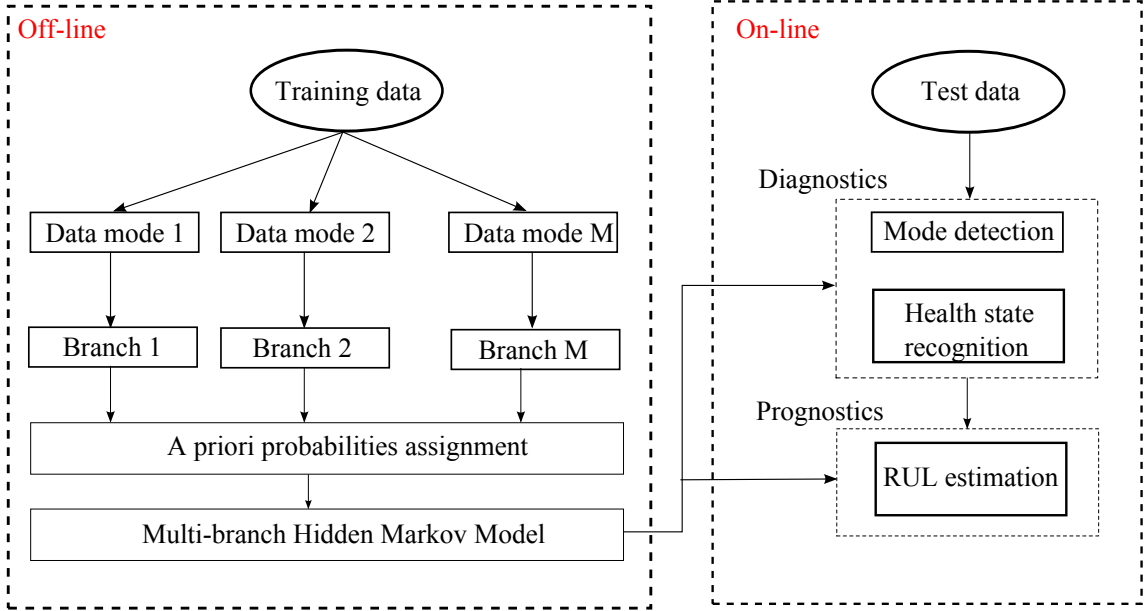


Figure 3: Proposed MB-HMM framework for diagnostics and prognostics

where $P(\mathbf{O} | \lambda_k)$ is the likelihood of the model λ_k given the data \mathbf{O} and can be calculated thanks to the popular forward-backward algorithm. $P(\lambda_k)$ is calculated from equation (2).

After having identified the deterioration mode, the next stage of the diagnostics is to recognize the actual health state of the equipment. This is carried out in two steps consisting of finding the single state sequence (path) which is the most probable one representing the given test data, and considering the final state as the actual health state. The first step is done thanks to the well-known Viterbi algorithm. Let q_t denote the equipment state at time t and $Q = (q_1, q_2, \dots, q_t)$ denote a path till the current time. Given the estimated mode, the most probable path is the one that gives the maximum joint probability of the path and the observation:

$$Q^* = \arg \max_{Q_k} P(\mathbf{O}, Q_k | \lambda_k) \quad (5)$$

where Q_i represents a possible path within the mode i . The actual health state of the equipment is the final one of the path Q^* .

3.2.1 RUL calculation

Given the deterioration mode and the current health state estimated from the diagnostic stage, the RUL estimation can be conducted the same as in a single HMM case and is straightforward. In the MB-HMM framework, this estimation, however, strongly depends on the mode detection. Indeed, a wrong deterioration mode detection may result in a large bias in the RUL estimation. For example, if a crack propagates with a high rate, the detection of a low rate mode may lead to an extremely larger value of the RUL than the actual one. This phenomenon often occurs at the beginning of a defect propagation since the observations are insufficient to show an obvious

deterioration trend. In this study, the Bayesian Model Averaging (BMA) technique (Hoeting et al. 1999) is implemented to tackle this problem. The RUL distribution is calculated as an average of the posterior distributions under each constituent model, weighted by their posterior model probability:

$$P(\text{RUL} | \mathbf{O}) = \sum_{k=1}^M P(\text{RUL} | \lambda_k, \mathbf{O}) P(\lambda_k | \mathbf{O}) \quad (6)$$

where $P(\lambda_k | \mathbf{O})$ is calculated from the equation (4).

The RUL estimation now turns to the estimation of the individual RULs under each constituent HMM, which is straightforward and can be computed in a recursive way as follows. Given the test data, the health state of equipment is identified thanks to the Viterbi algorithm. The RUL can be defined as the necessary time steps to reach for the first time the final state S_N from the current state S_i ($i < N$):

$$\text{RUL} = \min \{n \geq 0 : q_{t+n} = S_N | q_t = S_i\} \quad (7)$$

where q_t denote the health state at the time t . By this definition, the RUL can be seen as a discrete variable and its probability mass function is given by

$$P(\text{RUL} = n | q_t = S_i) = P(q_{t+n} = S_N, q_{t+n-1} \neq S_N, \dots, q_{t+1} \neq S_N | q_t = S_i) \quad (8)$$

Denote $h_i^{(n)} = P(\text{RUL} = n | q_t = S_i)$. Since the model is strictly one-order left-right topology HMM, the RUL can be calculated by the following backward recursive equations:

At state S_{N-1} :

$$h_{N-1}^{(1)} = a_{(N-1)N} \quad (9)$$

$$h_{N-1}^{(n)} = a_{(N-1)(N-1)} h_{N-1}^{(n-1)} \quad (10)$$

At state S_{N-2} :

$$h_{N-2}^{(1)} = a_{(N-2)N} \quad (11)$$

$$h_{N-2}^{(n)} = a_{(N-2)(N-2)} h_{N-2}^{(n-1)} + a_{(N-2)(N-1)} h_{N-1}^{(n-1)} \quad (12)$$

.....

At state S_i

$$h_i^{(1)} = a_{iN} \quad (13)$$

$$h_i^{(n)} = a_{ii} h_i^{(n-1)} + a_{i(i+1)} h_{i+1}^{(n-1)} \quad (14)$$

where a_{ij} is the transition probability from state S_i to state S_j and is obtained from the transition matrix A .

After calculating all the individual components, the RUL estimation under the MB-HMM can be accomplished by the equation (6).

4 NUMERICAL EXAMPLES

4.1 Fatigue Crack Growth model for data generation

To evaluate the proposed MB-HMM-based diagnostic and prognostic framework performance, it is necessary to have the data which represent several different deterioration modes. In this section, the Fatigue Crack Growth (FCG) model is used to model the crack appearing and propagating within a bearing and to generate both the training as well as the testing data. The FCG is chosen since it has been widely used to describe the crack propagation in the literature (Huynh et al. 2012, Myotyri et al. 2006). The bearing is supposed to operate under different operation conditions so that the crack will propagate with different rates which represent different deterioration modes for our purpose.

4.1.1 Fatigue Crack Model

The FCG model is constructed based on the popular Paris-Erdogan equation to express the crack growth rate:

$$\frac{dx}{dt} = C (\Delta K)^n \quad (15)$$

where x is the crack depth, t is the time, ΔK is the stress intensity amplitude and C and n are constant depending on the material property. It is shown that ΔK is roughly proportional to square root of crack depth x by a factor β : $\Delta K = \beta \sqrt{x}$ (Myotyri et al. 2006).

To take into account the stochastic aspect of the crack propagation, equation (15) could be discretized and then randomized by adding a multiplicative factor following a log-normal distribution as follows:

$$x_{t_i} = x_{t_{i-1}} + e^{w_{t_i}} C (\beta e \sqrt{x_{t_{i-1}}})^n \Delta t \quad (16)$$

where w_{t_i} are independent and identically distributed according to a Normal distribution $\mathcal{N}(0, \sigma_w^2)$, $0 < x_{t_{i-1}} < x_{t_i} < \infty$. This discretized and randomized model allows us to determine recursively the crack depth x_{t_i} at time t_i given the previous depth $x_{t_{i-1}}$ and the model parameters. By this way, the deterioration data can be generated.

For modeling different deterioration modes, the factor β is considered to be a function $\beta(e)$ of the operating environment state e as below (Huynh et al. 2012):

$$\beta(e) = \beta_b \cdot e^{\gamma_e} \quad (17)$$

where β_b is the base stress level of system, the values $\gamma_e \geq 0$, $e = 1, 2, \dots, M$ determine the level of extra stress linking with the state e of the environment. Obviously, from the equations (16) and (17), the propagation rate is directly proportional to the parameters γ_e : the greater the value of γ_e is, the more quickly the crack propagates.

4.1.2 Observation model

The actual crack depth is difficult to be directly and accurately measured because of several different reasons (Huynh et al. 2012). Hence, we need an observation model to represent the relationship between these actual values and the measure ones. In this study, the measurement is assumed to be the sum of the actual crack depth and a zero-mean Gaussian noise. Let X denote the crack depth and Y denote the corresponding measurement, the observation model is:

$$Y = X + \xi \quad (18)$$

where $\xi \sim N(0, \sigma_\xi^2)$ is the measurement error.

4.2 Application of MB-HMM for FCG data

In this section, the bearing is supposed to operate under two different environments and therefore the crack propagates in two different rates. A two-branch HMM is chosen to represent these two deterioration modes. The following parameters of the FCG model are chosen: $C = 0.005$, $n = 1.3$, $\beta_b = 1$, $\sigma_w = 1.7$, $\Delta t = 1$, $\gamma_e = [0 \ 0.75]^T$, $\sigma_\xi^2 = [5 \ 10]^T$ where the subscript T denotes the transpose of a vector. The different variance values of σ_ξ^2 mean that the measurement error may depend slightly on the operating condition.

The bearing is supposed to be failed once the crack depth reaches a critical level $d = 100$. 100 data curves have been generated by using the FCG model with the above parameters and used as the training data. Figure 4 illustrates this training data set.

In this study, the number of deterioration modes is assumed to be known beforehand. Therefore, "kmeans" tool is used to classify the failure times into 2 "clusters". Based on this, the training data can be

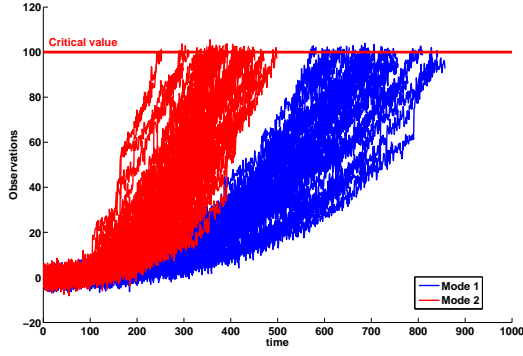


Figure 4: Two-mode training data

separated into two groups, which is then used to train different branches of the MB-HMM.

In order to train each constituent model, it is necessary to determine the number of states of each model. This task is done by using the BIC criterion. Figure 5 illustrates the BIC values with respect to different state numbers for a branch of the MB-HMM.

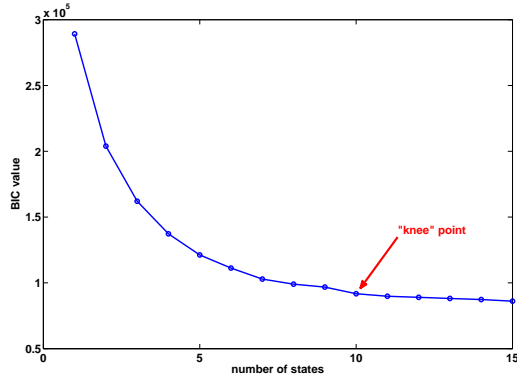


Figure 5: BIC for number of state selection

Since the crack propagation herein is a continuous process, the increase in the states number N will obviously reduce the BIC value. To avoid the overfitting situation, it is sufficient to take the "knee" point on the BIC curve. The value $N_{statesmode1} = 10$ is hence chosen for this example.

After having determined the number of states, each model is trained thanks to the Baum-Welch algorithm. The next step is to assign the *a priori* probability for each branch (c.f. equation (2)) and the result are: $\pi_1 = 0.41$ and $\pi_2 = 0.59$. An MB-HMM is obtained and the offline phase is terminated.

We now move to the on-line phase. One data curve representing the propagation of the crack depth from the beginning to the end of bearing life is also generated by the FCG model. This curve serves as the test data and is shown in the figure 6 (upper sub-figure). Only the red part till the actual time $t_{act} = 100h$ are assumed to be observable. The estimated RUL pdf is illustrated in the lower sub-figure of the figure 6.

To simulate the on-line RUL estimation problem, the actual time t_{act} is gradually replaced by $30h$ towards the failure. After each replacement, more information about the deterioration is achieved and

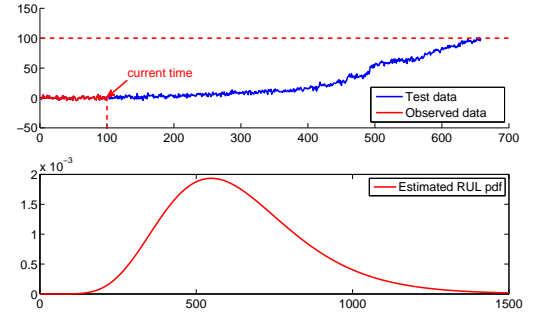


Figure 6: Observations and RUL estimation at $t_{act} = 100h$

the diagnostic and prognostic procedures are re-implemented. Figure 7 represents the mean value of the estimated RUL associated with the 95% confidence interval. It can be realized that, at the early instances, the lack of information due to the limited observations results in the bias and large variance in the RUL estimation. However, the actual values always lie within the 95% confidence interval. As the time passes, the length of the confidence interval significantly decreases and the estimated RUL converges to the real value. This demonstrates demonstrates the accuracy of the proposed RUL estimation method.

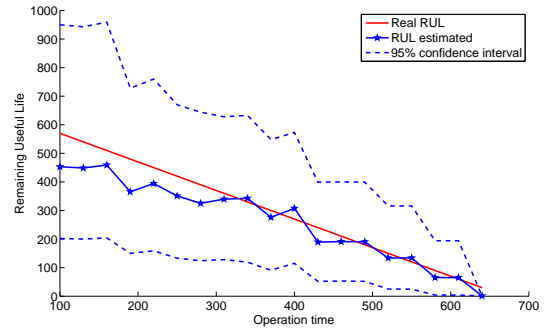


Figure 7: RUL estimation at different time instances

4.2.1 MB-HMM vs average HMM

Investigating the advantage of the proposed approach naturally raises a question: given the same data set, which method between the one using the proposed MB-HMM and the one using only an "average" HMM will give the better performance in RUL prediction. To answer such question, a comparative study is given in this section with 100 test data curves. The prognostic performance of each method is characterized by the root mean squared error (RMSE) metric.

For notation convenience, we denote the average HMM by AVG-HMM. In the training phase, the training data set in the previous section is not grouped and is used as a whole to train only one AVG-HMM. This model is then used in the on-line phase to estimate the RUL. At $t_{act} = 100h$, we obtain $RMSE_{AVG-HMM} = 74.4$. Compared to the corresponding result of the MB-HMM: $RMSE_{MB-HMM} = 52.5$, it can be concluded that the MB-HMM give a better performance

in RUL estimation compared with the AVG-HMM for this case.

Another interesting question raised in evaluating the advantages of the proposed MB-HMM is: does the above conclusion still hold in case that the two propagation rates are not quite different, or in case that there are two obvious trends in the training data? To answer this question, it is necessary to define a "distance" between two deterioration modes. In this study, the "mode distance" is defined by the gap between the two propagation rates and can be changed through the parameter γ_e of the FCG model. Indeed, as indicated in the section 4.1.1, the propagation rate of a crack is directly proportional to γ_e . The greater the value of γ_e is, the more quickly the crack propagates. Therefore, by fixing γ_1 and do varying γ_2 , the mode distance is changed. Figure 8 shows that, when the distance between modes increases, the RMSE given by MB-HMM approach decreases while with AVG-HMM, the RUL prediction error increases. It can be concluded that the larger the distance between deterioration modes is, the better performance in RUL estimation it is given by the proposed MB-HMM approach, in comparison to an "average" HMM.

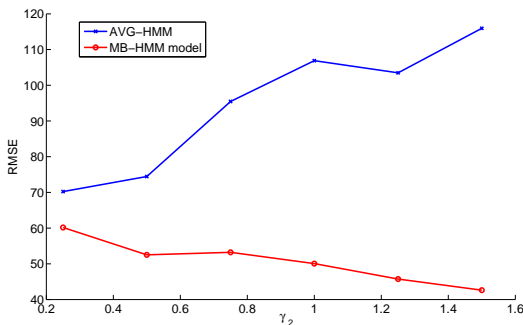


Figure 8: Prediction RMSE vs γ_2

5 CONCLUSION

The present paper proposes a novel model called multi-branch Hidden Markov Model (MB-HMM) to deal with the deterioration modeling problem in case that there are several competing modes existing in deterioration mechanism of equipment. A diagnostic and prognostic framework based on the proposed model is also given. The actual health state recognition and the RUL estimation can be implemented in the same way as the traditional HMM. However, by using different HMM models for different modes, together with the *a priori* probabilities of each model, the MB-HMM shows a very promising result in dealing with the multi-mode deterioration mechanism.

Future research will be focused on the extension of the MB-HMM to a multi-branch Hidden semi-Markov Model (MB-HSMM) due to the fact that the state sojourn time of an HMM follows an exponential distribution, which may not be hold in practice. Another possible extension of the model is the ability of

transition between the states of the different branches. The proposed model and the extension should also be validated on the data of real systems.

ACKNOWLEDGEMENT

This work is a part of the Ph.D research work of Thanh Trung LE in the framework of SUPREME ("Sustainable Predictive Maintenance for manufacturing Equipment") project, funded by the European Commission in FP7 Programme, under the Factories of the Future PPP.

REFERENCES

- Baruah, P. & R. B. Chinnam (2005). Hmms for diagnostics and prognostics in machining processes. *International Journal of Production Research* 43(6), 1275–1293.
- Bishop, C. M. et al. (2006). *Pattern recognition and machine learning*, Volume 1. springer New York.
- Dong, M. & D. He (2007). A segmental hidden semi-markov model (hsmm)-based diagnostics and prognostics framework and methodology. *Mechanical Systems and Signal Processing* 21(5), 2248–2266.
- Geramifard, O., J.-X. Xu, J.-H. Zhou, & X. Li (2012). A physically segmented hidden markov model approach for continuous tool condition monitoring: Diagnostics and prognostics. *Industrial Informatics, IEEE Transactions on* 8(4), 964–973.
- Hoeting, J. A., D. Madigan, A. E. Raftery, & C. T. Volinsky (1999). Bayesian model averaging: a tutorial. *Statistical science*, 382–401.
- Huynh, K. T., A. Barros, & C. Bérenguer (2012). Adaptive condition-based maintenance decision framework for deteriorating systems operating under variable environment and uncertain condition monitoring. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 226(6), 602–623.
- Jardine, A. K., 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.
- Knoblauch, D. (2004). Data driven number-of-states selection in hmm topologies. In *INTERSPEECH*.
- Le Son, K., M. Fouladirad, A. Barros, E. Levrat, & B. Iung (2012). Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety*.
- Lee, J. J., J. Kim, & J. H. Kim (2001). Data-driven design of hmm topology for online handwriting recognition. *International journal of pattern recognition and artificial intelligence* 15(01), 107–121.
- Myotyrki, E., U. Pulkkinen, & K. Simola (2006). Application of stochastic filtering for lifetime prediction. *Reliability Engineering & System Safety* 91(2), 200–208.
- Peng, Y. & M. Dong (2011). A prognosis method using age-dependent hidden semi-markov model for equipment health prediction. *Mechanical Systems and Signal Processing* 25(1), 237–252.
- Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE* 77(2), 257–286.
- Rabiner, L. R., B.-H. Juang, S. Levinson, & M. Sondhi (1985). Some properties of continuous hidden markov model representations. *At&T Technical Journal* 64(6), 1251–1270.
- Si, X.-S., W. Wang, C.-H. Hu, & D.-H. Zhou (2011). Remaining useful life estimation—a review on the statistical data

- driven approaches. *European Journal of Operational Research* 213(1), 1–14.
- Tobon-Mejia, D. A., K. Medjaher, N. Zerhouni, & G. Tripot (2012). A data-driven failure prognostics method based on mixture of gaussians hidden markov models. *Reliability, IEEE Transactions on* 61(2), 491–503.
- Van Noortwijk, J. (2009). A survey of the application of gamma processes in maintenance. *Reliability Engineering & System Safety* 94(1), 2–21.
- Wang, M. & J. Wang (2012). Chmm for tool condition monitoring and remaining useful life prediction. *The International Journal of Advanced Manufacturing Technology* 59(5-8), 463–471.
- Wang, W., A. Brakensiek, A. Kosmala, & G. Rigoll (2001). Multi-branch and two-pass hmm modeling approaches for off-line cursive handwriting recognition. In *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, pp. 231–235. IEEE.
- Zhang, X., R. Xu, C. Kwan, S. Y. Liang, Q. Xie, & L. Haynes (2005). An integrated approach to bearing fault diagnostics and prognostics. In *American Control Conference, 2005. Proceedings of the 2005*, pp. 2750–2755. IEEE.