Health stages division and remaining useful life prediction of rolling element bearings based on hidden semi-Markov model

Hongwei Wu[1,2], Zhenxing Liu[1,2], Yong Zhang[1,2], Ying Zheng[3], Cong Tang[1,2]

1. School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China.

2. Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan 430081, China.

3. School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China.

Abstract: Health stages division and Remaining Useful Life (RUL) prediction are two important parts in safety study of rolling element bearings. In this paper, the Hidden Semi-Markov Model (HSMM) is proposed to divide the degradation stages of rolling element bearings. Firstly, we extract the root mean square feature from the original vibration signal, then utilize Viterbi algorithm to divide the degradation stages. Secondly, Fault occurrence time is determined according to the degradation stage and RUL is predicted with HSMM. In order to verify the effectiveness of this method, IEE-PHM-2012 challenge data sets are adopted and the comparison with the existing methods is carried out.

Key Words: Health stages division; Fault occurrence time detection; Remaining useful life prediction; Hidden semi-Markov model.

2021 33rd Chinese Control and Decision Conference (CCDC) | 978-1-6654-4089-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CCDC52312.2021.9601903

1 Introduction

Accurate prediction of remaining useful life (RUL) of mechanical equipment is beneficial to improve equipment reliability and reduce equipment failure rate [1]. However, with the continuous operation of mechanical equipment, there will be a gradual failure, and generally with the continuous operation of the equipment presents an obvious degradation trend [2]. At this time, the traditional operation and maintenance methods and concepts are gradually not competent for the normal operation of the equipment. Recently, there are a variety of different equipment maintenance methods, Prognostics and Health Management (PHM) is one of them [3].

Generally, the equipment health prediction program usually consists of four steps, namely data collection, health indicator (HI) construction, health stage division and RUL prediction [4]. In fact, it is not necessary to survey the entire life of the equipment process of equipment for RUL. Therefore, it is inevitable to monitor the health stage of the equipment. According to the different degradation tendency, the equipment can be divided into two diverse stages and multiple stages [5].

For the division of health stage, most of the literatures divide degradation stage artificially according to the trend change of vibration signal characteristics. Ben Ali et al. [6] combined EMD with artificial neural networks to monitor the degradation state of the full-life bearing and classify defects. Soualhi et al. [7] took the change trend of RMS as an index to divide the degradation stage, and sets an empirical threshold by observing the change trend of RMS, and then divides the degradation stage. However, manual setting of experience threshold is not universal. Soualhi et al. [8]suggested the use of points to divide the health status through a method based on reinforcement learning through change, but it is better only for small samples.

For RUL prediction, Li et al. [9] construct a hazard model to describe the time-varying and conditional adaptive state transition probabilities and using multi-layer perception (MLP) that approximates a nonlinear function to calculate the observation probability based on the HMM model, which are integrated to estimate the tool wear state and predict the RUL online using forward algorithm on different cutting conditions.It has a more versatile and effective tool wear monitoring function. Zhu et al. [10] used the HMM model to automatically detect state changes to accurately locate FOT, and then using a novel transfer learning method based on multiple layer perceptron to accurately predict the remaining bearing life after the fault. However, these articles are all RUL predictions under the hidden Markov model. The disadvantage is that they do not con-

This work was supported in part by the National Natural Science Foundation of China [grant numbers 61873197,61873102].

sider the duration of state transition to the next state.

In order to solve the above problems, this paper is mainly devoted to solve health stages division and remaining useful life prediction. The main contributions are divided into the following parts: 1), taking RMS of the original signal as the degrade feature, health stages can be divided and fault occurrence time can be detected with the help of Viterbi algorithm; and 2) based on the information of FOT, RUL can be predicted by hidden semi-markov model, both upper bound and lower bound of the predicted RUL can be obtained simultaneously.

2 Health stages division

2.1 Hidden semi-markov model

Standard HSMM requires A, B, π and P. For the convenience, the whole can be abbreviated as $\lambda = (A, B, \pi, P)$. In order to estimate these parameters, an iterative expectation maximization (EM) algorithm also called Baum-Welch (BW) algorithm is proposed. BW algorithm is a customary method for parameter estimation of HMMs. The revaluation formulas of HSMM [11] are as follows:

Step 1: Define $\xi_t(i, j)$ as the probability that given observation sequence $[O_1, O_2, ..., O_T]$, the duration unit of the system in state i is d, and the next state is j. Then the revaluation formula of the state transition matrix $A = \{\overline{a}_{ij}\}$

$$\overline{a}_{ij} = \frac{\sum_{t=1}^{T} \xi_t(i, j)}{\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} \xi_t(i, j)} \tag{1}$$

Step 2 Denote $\gamma_t^d(j, g)$ as the output probability of the g-th Gaussian mixture of an observation sequence in state j at time t. Then the observation probability distribution $B = \overline{b_j}(O_k)$

$$\begin{cases} \overline{\omega}_{jg} = \frac{\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)}{\sum_{g=1}^{G}\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)} \\ \overline{\mu}_{jg} = \frac{\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)\sum_{s=t-d+1}^{t} O_s}{\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)} \\ \overline{\Sigma}_{jg} = \frac{\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)\sum_{s=t-d+1}^{t}[O_s-\mu_{jg}][O_s-\mu_{jg}]^T}{\sum_{t=1}^{T}\sum_{d=1}^{t-1}\gamma_t^d(j,g)} \\ \overline{b_j}(O_k) = \sum_{g=1}^{G}\overline{\omega}_{jg} \times N(O_k, \overline{\mu}_{jg}, \overline{\Sigma}_{jg}) \end{cases} \tag{2}$$

An initial state probability distribution $\overline{\pi}_i$

$$\overline{\pi}_i = \gamma_1(i) \tag{3}$$

Step 3: Let the forward variable $\alpha_t(j)$ to characterize the probability that the model produces an observation sequence and ends in state j

$$\begin{cases} \alpha_t(j) = P(O_1, O_2...O_t, q_t = S_j | q_{t+1} \neq S_j, \lambda) \\ 1 \leq t \leq T-1, 1 \leq j \leq N \\ P(O|\lambda) = \Sigma_{j=1}^{N}\alpha_T(j) \end{cases} \tag{4}$$

Step 4: Setting a backward variable $\beta_t(i)$ similar to the forward variable, which represents the probability that the model produces the observation sequence and starts at the state $j (j \neq i)$

$$\begin{cases} \beta_t(i) = P(O_1, O_2...O_t, q_t = S_j | q_{t+1} \neq S_j, \lambda) \\ 1 \leq t \leq T-1, 1 \leq j \leq N, \beta_T(j) = 1 \\ P(O|\lambda) = \Sigma_{i=1}^{N}\beta_1(i) \end{cases} \tag{5}$$

Define an intermediate variable

$$X_{t,t'}(j) = \frac{\Sigma_{i\neq j}^{N}\alpha_{t-1}(i)a_{ij}\prod_t^{t'} b_j(O_s)P(j,d)\beta_{t'}(j)}{P(O|\lambda)} \tag{6}$$

A state duration probability distribution P

$$\begin{cases} \overline{m}_j = \frac{\sum_{t=2}^{T}\sum_{t'=t}^{T} X_{t,t'}(j)(t'-t+1)}{\sum_{t=2}^{T}\sum_{t'=t}^{T} X_{t,t'}(j)} \\ \overline{\sigma}_j^2 = \frac{\sum_{t=1}^{T}\sum_{t'=t}^{T} X_{t,t'}(j)(t'-t+1)^2}{\sum_{t=1}^{T}\sum_{t'=t}^{T} X_{t,t'}(j)} - \overline{m}_j^2 \\ \overline{P}(j, d) = N(d|\overline{m}_j, \overline{\sigma}_j) \end{cases} \tag{7}$$

Bring the initial HSMM model parameter λ into the re-estimation formula, and continuously iterate to get a new HSMM model $\overline{\lambda}$ until $P(O|\overline{\lambda})$ satisfies the convergence condition, then the final training model can be obtained.

2.2 Viterbi algorithm

Viterbi algorithm is a method based on dynamic programming to solve the shortest path of an observation sequence, so that the obtained observation sequence has the best interpretability. The so-called optimal here refers to the state sequence obtained when the probability is maximum under the known model λ and observation sequence [11]. Before introducing the process of the degradation stage division of the Viterbi algorithm, here we first define an intermediate variable $\delta_t(i)$.

$$\delta_t(i) = max_{h_1...h_{t-1}} P(h_1...h_t = S_i, O_1, O_2...O_t \mid \lambda) \tag{8}$$

where $\delta_t(i)$ characterizes the maximum probability of generating observation sequence $O = O(O_1, O_2...O_t)$ when $h_t = S_i$ along the state sequence $(h_1, h_2...h_t)$ at time t. Viterbi algorithm to divide the health stage can be represented as follow.

1. Extract RMS feature of the training set as the input of the Hidden Semi-Markov Model, and obtain $\overline{\lambda}$.
2. Extract RMS feature of the test set, and get the observation sequence $O = O(O_1, O_2...O_T)$.
3. Initialization parameters:$\delta_1(i) = \pi_i b_i(O_1)$, $\psi_1(i) = 0$,$1 \leq i \leq N$,where N is the number of states, π and b are model parameters.
4. Recursive process:Calculate $\delta_t(j)$ from $j = 1$ to $j = N$, Calculate $\psi_t(j)$ from $j = 1$ to $j = N$.
5. Terminate iteration:$P^* = max_{1\leq i\leq N}[\psi_t(i)]$, $q_T^* = argmax_{1\leq i\leq N}[\psi_t(i)]$.

6. $q_t^* = \psi_{t+1}(q_{t+1}^*), t = T-1, T-2, ...2, 1$ is the optimal hidden state sequence.

3 HSMM for RUL prediction

Before performing RUL prediction, we first need to solve the problem of identifying the current degradation state, that is, given the sample to be tested to determine which degradation stage it is in. The samples to be tested are input into N states classifiers for fault state identification, and the output probability $P(O|\lambda_i)(1 < i < N)$ of the test samples in the N classifiers is calculated using the Viterbi algorithm. Comparing the size of each output probability, the degradation state corresponding to the model with the largest output probability is the current state of the component. Once the degradation state of the current observation sequence is determined, it can be matched with our prediction model and the model with the highest similarity probability is selected for RUL prediction.

For the duration parameter $P(j, d)$ of each state, the mean $\mu(h_i)$ and variance $\sigma^2(h_i)$ corresponding to the state can be determined by the Viterbi algorithm, and then the maximum duration of each state $D(h_i)$can be estimated according to the following formula:

$$\begin{cases} D(h_i) = \mu(h_i) + \rho\sigma^2(h_i) \\ \rho = (T - \sum_{i=1}^{N} \mu(h_i)) / \sum_{i=1}^{N} \sigma^2(h_i) \end{cases} \tag{9}$$

where T is the full life of the bearings.

Three related indicators are mainly calculated: the upper limit of RUL, the lower limit of RUL, and the average value of RUL.

$$\begin{cases} R_{upper}(t) = \sum_{i=k}^{N} [\mu(D(h_i)) + \sigma(D(h_i))] - t \\ R_{mean}(t) = \sum_{i=k}^{N} \mu(D(h_i)) - t \\ R_{lower}(t) = \sum_{i=k}^{N} [\mu(D(h_i)) - \sigma(D(h_i))] - t \end{cases} \tag{10}$$

where R is the remaining usefui life, k is the current state of the observation sequence.

Therefore, the method can be used for real-time prediction, that is, every time an observation value is obtained, a full life model can be selected to predict the remaining life. The forecast includes the above three indicators, depicting the fit between actual degradation and forecast.

To quantitatively evaluate the prediction effect of the proposed method, the following three evaluation criteria are introduced:

1) Mean absolute percentage error (MAPE): MAPE is calculated as follows

$$E_{MAPE} = \frac{1}{n} \sum_{k=1}^{n} |\frac{\hat{y}_k - y_k}{y_k}| \times 100\% \tag{11}$$

2) Mean absolute error (MAE): It is the average absolute error between the true and predicted values, which is usually used to measure the closeness of the predictions with the actual results.

$$E_{MAE} = \frac{1}{n} \sum_{k=1}^{n} |y_k - \hat{y}_k| \tag{12}$$

3) Normalized root mean square error (NRMSE): N-RMSE can be described as

$$E_{NRMSE} = \frac{\sqrt{\frac{1}{n} \sum_{k=1}^{n} (y_k - \hat{y}_k)^2}}{(\frac{1}{n} \sum_{k=1}^{n} \hat{y}_k)} \tag{13}$$

As far as NRMSE, MAPE and MAE are concerned, if these indicators are close to 0, the accuracy of RUL prediction is higher.

4 Simulation experiment

4.1 Experimental platform

The PRONOSTIA experimental platform is mainly used to verify methods related to bearing health assessment, diagnosis and prognosis [12]. It is an accelerated test platform that describes the natural degradation process of the bearing throughout its life. It imposes additional load on the bearing or increases the speed to accelerate the failure. The structure of the experimental platform comes from [12], in which two accelerators are installed perpendicular to each other on the bearing to collect horizontal and vertical vibration signals. The accelerator is sampled every 10 seconds, the sampling frequency is 25.6 kHZ, and the sampling time is 0.1s. Because the information used to track the bearing degradation provided by the vertical vibration signal is less than the horizontal vibration signal [13]. Therefore, only the horizontal signal is used as the research data set in this article.The platform includes three different working conditions, in which radial load and speed are variable. The detailed information of different conditions is shown in Table1.

Table 1: PRONOSTIA bearing dataset information.

	Condition1	Condition2	Condition3
Load(N)	4000	4200	5000
Speed(rpm)	1800	1650	1500
Training dataset	Bearing1−1	Bearing2−1	Bearing3−1
	Bearing1−2	Bearing2−2	Bearing3−2
Testing dataset	Bearing1−3	Bearing2−3	Bearing3−3
	Bearing1−4	Bearing2−4	
	Bearing1−5	Bearing2−5	
	Bearing1−6	Bearing2−6	
	Bearing1−7	Bearing2−7	

4.2 Health stage division

RMS feature is extracted from the original vibration signal. Taking Bearings under working condition 1 as an example, the result is shown in Fig.1. From Fig.1 we can find that the degradation trends of Bearing $1-1$ and Bearing $1-3$ are relatively close, while Bearing $1-2$ and Bearing $1-4$ are both in the rapid degradation situation, and the remaining three bearings are in the slow degradation stage.

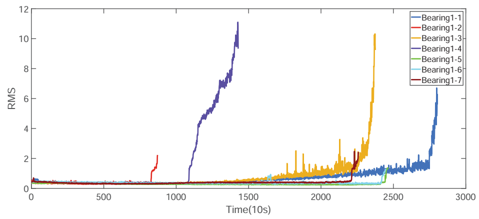


Figure 1: RMS of Seven Bearings in Condition 1.

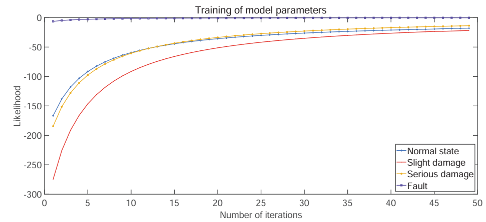


Figure 2: Convergence curve of model training.

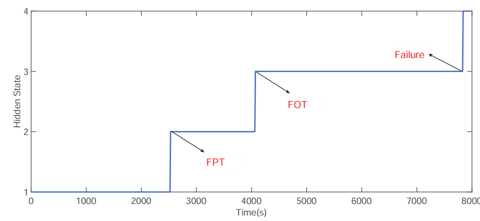


Figure 3: Illustration of different points of change

Once the feature vector is extracted, we can use the feature vector as the input to the model to train HSMM. For the initialization of model parameters, the selection of the initial values of π and A does not affect the final result, so we can initialize randomly. But for the initialization of B, the choice of initial value has a great influence on the calculation result. In this paper, the mixed Gaussian distribution is used as the observation probability density function, and the number of mixed Gaussian is set to 4. The general situation is to directly divide or cluster according to the number of states, and then calculate the corresponding distribution as the result of initialization. One of the more effective ones is to use the K-means clustering method for initialization. After initializing the parameters, the maximum number of iteration steps is set to 50, the algorithm convergence error is set to 0.0001, and the model training curve is shown in Fig.2.

After the training is completed, given the current observation sequence, the current observation sequence can be divided into health stages. There are $N-1$ change points for N states, and these change points are first predicting time(FPT), fault occurrence time (FOT) and failure point time. It can be seen from Fig.3 that FPT is the moment when the bearing begins to degenerate, so we pay more attention to FOT detection. The results of the health stage division under working condition 1 are shown in Fig.6, and the results of FOT are listed in Table2.

It is not difficult to see from Fig.4 that the method proposed in the seven bearings failed to detect the FOT of the bearings $1-5$, the bearings $1-6$, and the bearings $1-7$. The remaining bearings have been fully tested. And compared with the existing results, the results of the rapidly degraded bearing $1-2$ and the bearing $1-4$ are surprisingly similar. Bearings $1-1$ and $1-3$ are more accurate than the other methods, which shows the effectiveness of the proposed method.

Furthermore, the degradation stage division and FOT detection results are given under different working conditions in Table3. It can be concluded from Table3 that the proposed method is not only applicable to working condition 1, but also applicable to working condition 2 and working condition 3. It shows that this method can effectively divide the health stage of the bearing and seize the FOT.

4.3 RUL prediction

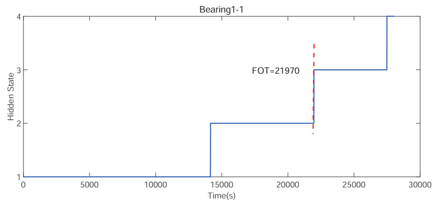
Once FOT is determined, the next step is to predict RUL. Every time a characteristic value of the tested bearing is obtained, the Viterbi algorithm is used to identify the current state. After the recognition, the characteristic value is matched with similar probability respectively, and the mod-

Table 2: FOT detection results.

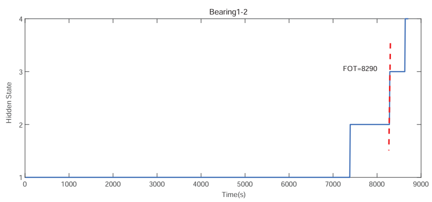
Bearing	Failure time	Method[13]	Method[14]	Method[15]	HSMM
$1-1$	2803	1490	2118	2174	**2197**
$1-2$	871	827	826	784	**829**
$1-3$	2375	1684	1613	1842	**2124**
$1-4$	1428	1083	1082	1108	**1094**

Table 3: FOT detection under different working conditions.

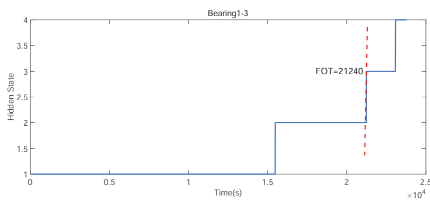
Bearings	$2-1$	$2-2$	$2-4$	$2-6$	$2-7$	$3-2$	$3-3$
FOT(s)	8750	4070	5590	6890	2250	16150	4160



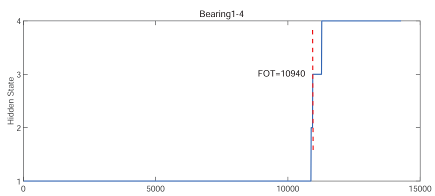
(a) Bearing1-1



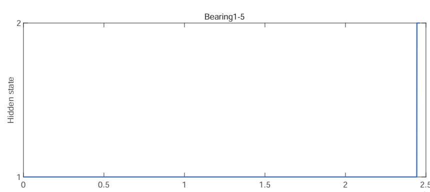
(b) Bearing1-2



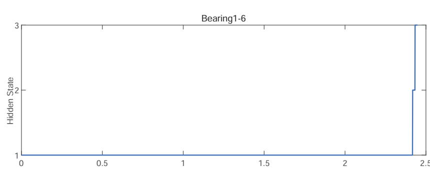
(c) Bearing1-3



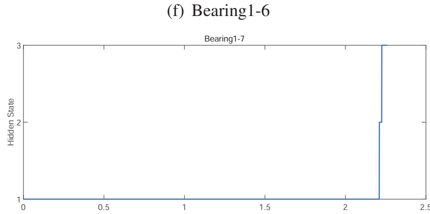
(d) Bearing1-4



(e) Bearing1-5



(f) Bearing1-6

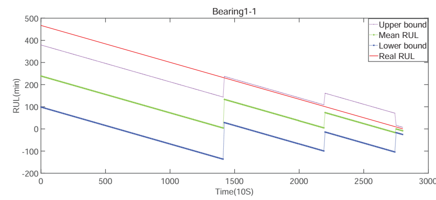


(g) Bearing1-7

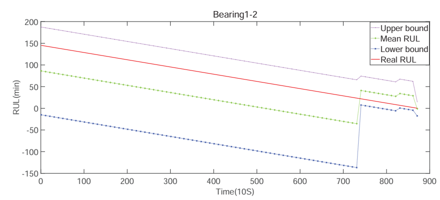
Figure 4: Health stage division.

el with the greatest similarity is selected. According to the current state and(10), the remaining life value is judged, and finally the model prediction and the actual degradation curve of the bearing are drawn.

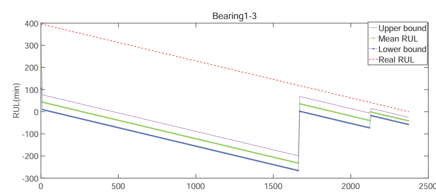
The mean $\mu(h_i)$ and standard deviation $\sigma(h_i)$ of the residence time of each state are obtained by the Viterbi algorithm, and the maximum residence time of each state is calculated according to (9). After obtaining $D(h_i)$, we can predict the remaining useful life of the bearings in real time. Then, the RUL prediction of the bearings test sets under working condition 1 are shown in Fig.5.



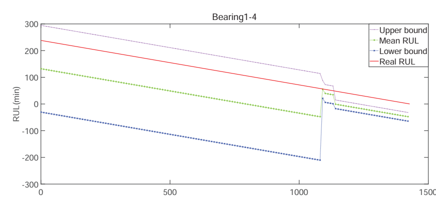
(a) Bearing1-1



(b) Bearing1-2



(c) Bearing1-3



(d) Bearing1-4

Figure 5: RUL prediction.

It can be found from Fig.5 that every time the bearing runs to the change point, there will be a step mutation, so that the FOT point can be easily found. Moreover, the confidence interval after FOT is significantly narrower, which is more accurate for RUL prediction. The error results of real-time prediction from FOT to failure are listed in Table 4.

Table 4: Performance of RUL prediction.

Bearings	Failure	FOT	$E_{MAPE}(\%)$	E_{MAE}	E_{NRMSE}
1−1	28030	21970	**1.0519**	**25.2695**	**0.0001**
1−2	8710	8290	**6.6808**	**22.3184**	**0.0005**
1−3	23750	21240	**6.1086**	**42.0000**	**-0.0001**
1−4	14280	10940	**5.5714**	**45.2682**	**-0.0001**
2−1	9110	8750	**37.2828**	**32.0014**	**0.0029**
2−2	7970	4070	**3.3361**	**32.1959**	**-0.0003**
2−4	7510	5590	**7.1745**	**38.4751**	**-0.0001**
2−6	7010	6890	**63.6083**	**25.9626**	**0.0272**
2−7	2300	2250	**52.5961**	**8.7660**	**0.9813**
3−2	16370	16150	**51.3922**	**61.7927**	**0.0275**
3−3	4340	4160	**27.0124**	**15.6197**	**0.0095**

5 Conclusion

The health stage division is a very meaningful part for PHM. In this paper, a method for evaluating the degradation state of rolling bearings based on degradation stage division and RUL prediction is proposed. The Viterbi algorithm is used to divide the degradation stages and obtain the FOT. Once the FOT is obtained, remaining life of bearings at the FOT can be predicted. Experiments show that this method can effectively obtain the bearing degradation stage and detect the FOT.

REFERENCES

[1] Lee J, Wu F J, Zhao W Y, et al. Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications, *Mechanical Systems and Signal Processing*, Vol. 42, No. 1–2, pp. 314–334, 2014.

[2] L. Xiao, Z. X. Liu, Y. Zhang, Y. Zheng, and C. Cheng, "Degradation assessment of bearings with trend-reconstruct-based features selection and gated recurrent unit network," *Measurement*,Vol.165, Article 108064, 2020.

[3] Liu Z, Zhen J, Vong C, Han J, Yan C and Pecht M, A patent analysis of Prognostics and Health Management (PHM) innovations for electrical systems, *IEEE Access*, DOI:10.1109/ACCESS.2018.2818114.

[4] Y.J. Cheng, J. Peng, X. Gu, X.Y. Zhang, W.R. Liu, Y.Z. Yang and Z.W. Huang, A Reinforcement Learning Method for Health Stage Division Using Change Points, *IEEE International Conference on Prognostics and Health Management*, DOI:10.1109/ICPHM.2018.8448499.

[5] Y. Lei, N. Li, L. Guo, N. Li, T. Yan and J. Lin, Machinery health prognostics, A systematic review from data acquisition to RUL prediction,*Mechanical Systems and Signal Processin*, Vol. 104, pp. 799–834, 2018.

[6] Saidi L, Ben Ali J, and Fnaiech F, Application of higher order spectral features and support vector machines for bearing faults classification, *Isa Transactions*, Vol. 54, pp. 193–206, 2015.

[7] A. Soualhi, H. Razik, G. Clerc and D. D. Doan, Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System,*IEEE Transactions on Industrial Electronics*,Vol. 61, No. 6, pp. 2864-2874, 2014.

[8] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilberthuang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, Vol. 64, No. 1, pp. 52-62, Jul, 2015.

[9] W. Li, T. Liu, Time varying and condition adaptive hidden markov model for tool wear state estimation and remaining useful life prediction in micro-milling, *Mechanical Systems and Signal Processing*, Vol. 131, pp. 689-702, 2019.

[10] J. Zhu, N. Chen, C.-Q. Shen, A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions, *Mechanical Systems and Signal Processing*, Vol. 139, Art, 106602, 2020.

[11] M. Dong, D. He, A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology, *Mechanical Systems and Signal Processing*, DOI:10.1016/J.YMSSP.2006.10.001.

[12] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, C. Varnier, PRONOSTIA: an experimental platform for bearings accelerated degradation tests, *Proceedings of the IEEE Conference on Prognostics and Health Management (PHM)*, pp. 1–8, 2012.

[13] X. Li, W. Zhang, Q. Ding, Deep learning-based remaining useful life estimation of bearings using multiscale feature extraction, *Reliability Engineering & System Safety*, Vol.182, pp. 208 - 218, 2019.

[14] J. Zhu, N. Chen and C.Q. Shen, A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions, *Mechanical Systems and Signal Processing*, DOI:10.1016/J.YMSSP.2019.106602.

[15] X.H. Jin, Y. Sun, Z.J. Que, Y. Wang and T.W.S. Chow, Anomaly Detection and Fault Prognosis for Bearings, *IEEE Transactions on Instrumentation and Measurement*, DOI:10.1109/TIM.2016.2570398.

[16] B. Zhang, S.H. Zhang and W.H. Li, Bearing performance degradation assessment using long short-term memory recurrent network,*Computers in Industry*, DOI:10.1016/J.COMPIND.2018.12.016.