RUL Prediction of Bearings Based on Mixture of Gaussians Bayesian Belief Network and Support Vector Data Description

Qianhui Wu, Yu Feng, and Biqing Huang(✉)

Department of Automation, Tsinghua University, Beijing 100084, China
hbq@tsinghua.edu.cn

Abstract. This paper presents a method to predict the remaining useful life of bearings based on theories of Mixture of Gaussians Bayesian Belief Network (MoG-BBN) and Support Vector Data Description (SVDD). Our method extracts feature vectors from raw sensor data using wavelet packet decomposition (WPD). The features are then used to train the corresponding MoG-BBN and SVDD model. Genetic algorithm is employed to determine the initial value of training algorithm and enhance the stability of our model. The two models are combined to acquire a good generalization ability. The effectiveness of the proposed method is verified by actual bearing datasets from the NASA prognostic data repository.

Keywords: Wavelet packet decomposition · Mixture of Gaussians Bayesian Belief Network · Genetic algorithm · Support Vector Data Description · Remaining useful life

1 Introduction

Bearings are one of the most commonly used components in mechanical equipment. Due to their high failure rate, the working condition of bearings directly affects the safety of the whole equipment. RUL prediction of bearings plays a key role in condition based maintenance (CBM) [1], as it can effectively anticipate bearing failure, reduce the maintenance cost as well as increase the productivity.

Since the condition monitoring data is available, this paper focuses on the data-driven methods [2]. Under this framework, a variety of previous researches about bearings prognostic and RUL prediction has been conducted, including artificial networks [3–5], hidden Markov models (HMM) [6, 7], support vector machines [8–10], etc. Huang *et al.* [5] trained the back propagation neural networks which focus on ball bearings' degradation periods by the MQE indicator obtained from SOM, then applied WAFT technology to make RUL prediction. Tobon-Mejia *et al.* [6] proposed a method based on the Mixture of Gaussian Hidden Markov Models, in which hidden states are used to represent the failure modes of bearings. The RUL can be estimated straightly by the stay durations in each state. Shen *et al.* [8] took the fuzziness of degradation into account and proposed a damage severity index (DSI) based on fuzzy support vector data description (FSVDD), which can indicate the growth of degradation with running time.

L. Zhang et al. (Eds.): AsiaSim 2016/SCS AutumnSim 2016, Part II, CCIS 644, pp. 118–130, 2016.
DOI: 10.1007/978-981-10-2666-9_13

In addition, Zhang *et al.* [11] constructed a Mixture of Gaussians Bayesian Belief Network (MoG-BBN) to characterize the degradation state by the condition monitoring data from sensors. However, the initial values of the parameters used in model training have a great impact on the accuracy of the RUL prediction. Hence, the stability and generalization of the model are reduced. To overcome this deficiency, this paper proposed a RUL prediction method based on the MoG-BBN and SVDD. The novelty of this paper lies in two respects. First, genetic algorithm is used to find the optimal initial value when training the MoG-BBN model so that the stability of the model enhances significantly. Second, a method based on SVDD are presented to estimate the RUL when the MoG-BBN model does not work well, which improve the generalization capability and the prediction accuracy.

The remainder of the paper is organized as follows: Sect. 2 introduces the methodology proposed for remaining useful life prediction of bearings. Section 3 carries out experiments on actual bearing data from NASA to examine the effectiveness of the proposed method. Section 4 concludes the work.

2 Methodology

The framework of the methodology is shown in Fig. 1. It can be divided into two phases: the off-line phase and the on-line phase. During the off-line phase, the raw data is processed to extract features, and then these features are used to train the MoG-BBN model and the SVDD model. During the on-line phase, the processing of the real-time vibration signal from sensors to extract features remains the same. Then, the features

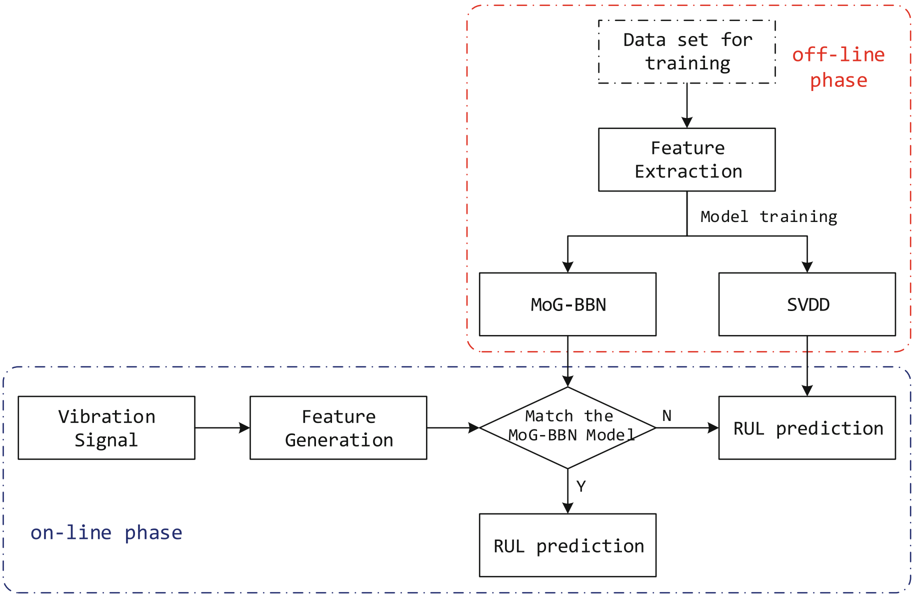


Fig. 1. Framework of the proposed method

are fed into the MoG-BBN model to characterize the degradation state. Some measures will be taken to determine whether the trained MoG-BBN model matches well with current component. If matching well, it comes to the RUL prediction step directly. Otherwise, the features will be fed into the SVDD model to get a better model performance and robustness. Finally, the RUL estimation will be implemented.

2.1 Feature Extraction

Wavelet packet decomposition (WPD) is an effective technique in signal analysis. It has sufficient high-frequency resolution, which contains the most useful fault information of bearings [12]. WPD can be considered as a tree, and its root is the original signal. By recursively applying the wavelet transform, WPD can automatically choose the appropriate frequency scale according to the characteristics of the analyzed signal, further decompose the high and low frequency data, and divide the spectrum band into several levels [7].

Based on excellent properties described above, WPD method is used to extract features from the raw vibration data of bearings in this paper. Assume that the decomposition level is l, then there will be $L = 2^l$ nodes on the last level. Let f_{it} represent the ith node of the last level, the feature vector at time t can be described as following.

$$\mathbf{f} = (f_{1t}, f_{2t}, \ldots, f_{Lt})^T \tag{1}$$

Note that a normalization process based on the mean and standard deviation of the raw data should be applied to the result of WPD before training the model in order to improve the generalization capability.

2.2 The Mixture of Gaussians Bayesian Belief Network

Structure. Figure 2 illustrates the MoG-BBN structure where D and M are discrete variables, and O is a continuous variable. In this structure, D represents the degradation states which cannot be directly observed, $D \in \{1, 2, \ldots a\}$. a is the maximum degradation state number. In this paper, a is set to 3, representing healthy, sub-healthy, and faulty states. M represents the distinctive Gaussian distributions for each state D, $M \in \{1, 2, \ldots b\}$, where b is the number of components in mixed Gaussian distribution. O represents the observation vector corresponding to a degradation state.

Note that M is the connection of the degradation state D and the observation O, which makes MoG-BBN a suitable tool, because it transforms continuous observations from monitoring sensors to discrete degradation states of physical components. From the above definition, once the probability $P(D|O)$ is known, the degradation state can be recognized, and then RUL of bearings can be estimated.

Let $\mathbf{o}$ be a realization of O. According to Fig. 2, the definition of conditional probability and the total probability formula, the value of $P(D|O)$ can be calculated as following.

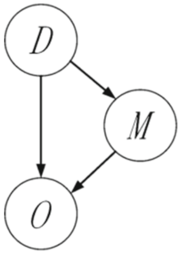


Fig. 2. Mixture of Gaussian Bayesian belief network.

$$P(D|O=\mathbf{o}) = \frac{P(D)\sum_{M} P(M|D)P(O=\mathbf{o}|D,M)}{\sum_{X} P(D)\sum_{M} P(M|D)P(O=\mathbf{o}|D,M)} \tag{2}$$

Note that, the values of $P(D)$, $P(M|D)$ and $P(O|D,M)$ must be inferred by EM algorithm shown below.

EM Algorithm. A declaration of the variables used in the estimation procedure is given first.

- π_d: the initial distribution of degradation state D, and $\pi_d = P(D=d)$ for $d \in \{1,2,\ldots,a\}$.
- C_{dm}: mixture coefficient of the mth Gaussian distribution for degradation state d, where $C_{dm} = P(M=m|D=d)$, for $d \in \{1,2,\ldots,a\}$, $m \in \{1,2,\ldots,b\}$.
- $\boldsymbol{\mu}_{dm}$: mean vector of the mth Gaussian distribution for degradation state d.
- $\boldsymbol{\Sigma}_{dm}$: covariance matrix of the mth Gaussian distribution for degradation state d.

To compute $P(O|D,M)$, suppose O_d is the distribution of an observation O generated by degradation state d, then

$$O_d = \sum_{m=1}^{b} C_{dm} N(\mu_{dm}, \Sigma_{dm}), 1 \le d \le a \tag{3}$$

The statistical values of the above parameters can be inferred via training data sets and the EM algorithm. Given an observation sequence $\mathbf{o} = \{\mathbf{o}^{(1)}, \mathbf{o}^{(2)}, \ldots, \mathbf{o}^{(N)}\}$, where N denotes the length of the sequence, the EM algorithm is divided into the Expectation-Step and the Maximization-Step.

- Expectation-Step: for each pair of (d,m,n), with $n \in \{1,2,\ldots,N\}$,

$$\begin{aligned}\omega_{dm}^{(n)} &= P(D^{(n)}=d, M^{(n)}=m|\mathbf{o}^{(n)}, \pi_d, C_{dm}, \boldsymbol{\mu}_{dm}, \boldsymbol{\Sigma}_{dm}) \\ &= \frac{P(\mathbf{o}^{(n)}|D^{(n)}=d, M^{(n)}=m, \boldsymbol{\mu}_{dm}, \boldsymbol{\Sigma}_{dm})P(D^{(n)}=d, M^{(n)}=m, \pi_d, C_{dm})}{\sum_{d=1}^{a}\sum_{m=1}^{b} P(\mathbf{o}^{(n)}|D^{(n)}=d, M^{(n)}=m, \boldsymbol{\mu}_{dm}, \boldsymbol{\Sigma}_{dm})P(D^{(n)}=d, M^{(n)}=m, \pi_d, C_{dm})}\end{aligned} \tag{4}$$

- Maximization-Step: update the above parameters as follows,

$$
\begin{aligned}
\pi_d &= \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{b}\omega_{dm}^{(n)} \\
C_{dm} &= \frac{1}{N\pi_d}\sum_{n=1}^{N}\omega_{dm}^{(n)} \\
\boldsymbol{\mu}_{dm} &= \frac{\sum_{n=1}^{N}\omega_{dm}^{(n)}\mathbf{o}^{(n)}}{\sum_{n=1}^{N}\omega_{dm}^{(n)}} \\
\boldsymbol{\Sigma}_{dm} &= \frac{\sum_{n=1}^{N}\omega_{dm}^{(n)}(\mathbf{o}^{(n)}-\boldsymbol{\mu}_{dm})(\mathbf{o}^{(n)}-\boldsymbol{\mu}_{dm})^T}{\sum_{n=1}^{N}\omega_{dm}^{(n)}}
\end{aligned}
\tag{5}
$$

Let $\lambda = (\boldsymbol{\pi}, \mathbf{C}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. The process is repeated until $\left|P(\mathbf{o}|\lambda^t) - P(\mathbf{o}|\lambda^{t-1})\right| < \xi$ or iteration number exceeds the maximum value set in advance. The threshold $\xi = 10^{-4}$ is used in the experiment section of this paper.

Initial Value Optimization Based on Genetic Algorithm. The initial value of $\boldsymbol{\pi}$ and $\mathbf{C}$ can be generated randomly from a uniform distribution, while $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be acquired through clustering methods for training samples. Due to the fact that the initial values of $\boldsymbol{\pi}$ and $\mathbf{C}$ have a great impact on the model obtained by training, in this paper, an initial value optimization method based on genetic algorithm is proposed.

Encoding. According to what is proposed in [13], encoding methods of GA can be classified into two main approaches: binary encoding and float encoding. Taking into account the convenience and accuracy, floating encoding is adopted in this paper. Suppose that the maximum degradation state number in the MoG-BBN model is a, and mixed Gaussian distribution number for each state is b. The numbers of total parameters to optimize related to $\boldsymbol{\pi}$ and $\mathbf{C}$ are a and $a \cdot b$, respectively. Then there are $a \cdot (b + 1)$ decimal floating numbers constituting an individual.

Fitness function. Corresponding to the fitness of the individuals to the environment, the value of this function reflects the fitness of the individuals in terms of measurement indicators. In this paper, the fitness function depicts the accuracy of the models generated by different individuals when identifying the degradation states.

Selection. According to each individual's fitness, the strategy of RWS is to calculate the probability that the gene is inherited from the individual by the offspring, and based on that probability, the offspring population is randomly selected in the parent generation. The higher the fitness of the parent is, the greater the probability that its genes will be selected to be inherited to the offspring will be. Let the fitness of the ith individual be f_i, and the population quantity be pop. The probability that this individual will be selected is

$$P_i = \frac{f_i}{\sum_{i=1}^{pop} f_i} \tag{6}$$

A method based on Roulette Wheel Selection (RWS) is used in this paper: let the first to the kth individuals be a part of the next generation directly on the basis of the fitness sorting (from large to small). And the rest $pop - k$ individuals can be generated by genic recombination and mutation of $pop - k$ pairs that are selected from the current population by RWS method.

Crossover. This paper use Arithmetic crossover to produce a new individual. Assume that two individuals of the parent are X_A and X_B, then the new individual generated by arithmetic crossover operator is as following, where γ is the parameter, which is commonly set to 0.5.

$$\begin{cases} X_A' = \gamma X_B + (1-\gamma)X_A \\ X_B' = \gamma X_A + (1-\gamma)X_B \end{cases} \tag{7}$$

Mutation. Mutation is a genetic operator which is used to maintain genetic diversity. To ensure the convergence, the mutation operator used in this paper adds or subtracts a small random number to the original floating number. This random number is called step width. Bigger step width leads to faster evolution speed at the beginning. However, it will be more difficult to converge at the end. In order to speed up the evolution and ensure that the genetic algorithm can be more accurate when converging to the optimal solution at the same time, a method that the step width changes dynamically is taken.

The steps of float coding genetic algorithm for the optimization of the initial values of $\boldsymbol{\pi}$ and $\mathbf{C}$ can be summarized as following.

(a) Randomly generate the initial population consisting of pop individuals.
(b) Calculate and sort the fitness of each individual in the current population, from large to small.
(c) Let the first to the kth individuals be a part of the next generation directly on the basis of the fitness sorting. Generate the rest $pop - k$ individuals by genic crossover and mutation of $pop - k$ pairs that are selected from the current population by RWS method. Note that the step width in the mutation process decreases gradually with the iteration. Besides, if necessary, additional measures should be taken to ensure that the values of float genes related to $\boldsymbol{\pi}$ and $\mathbf{C}$ are significant after crossover and mutation, respectively.
(d) If the resulting solution tends to be stable in a certain range or the number of iterations reach the maximum value, exit and the optimal solution is gotten. Otherwise, turn to step (b).

2.3 Support Vector Data Description

Assume a training set containing n vectors of objects $\{\mathbf{x}_i, i = 1, 2, \ldots, N\}$. The optimization objective of the SVDD method is to find the minimum-volume hypersphere containing all or most possible target data points in feature space, and it can be described as following:

$$\begin{aligned} &\min_{R,\mathbf{c},\xi} R^2 + C\sum_{i=1}^{N}\xi_i \\ &s.t.(x_i - c)^T(x_i - c) \leq R^2 + \xi_i,\ \xi_i \geq 0, i = 1, 2, \ldots, N \end{aligned} \tag{8}$$

where c is the center of the hypersphere, and ξ_i is slack variable, cooperating with the penalty constant C to make the trade-off between the radius R and the number of data points that lie out of the hypersphere [14]. Construct the Lagrangian:

$$L(R, c, \alpha_i, \xi_i) = R^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left\{R^2 + \xi_i - (x_i^2 - 2cx + c^2)\right\} - \sum_{i=1}^{N}\gamma_i\xi_i \tag{9}$$

Where $\alpha_i \geq 0$ and $\gamma_i \geq 0$. Set the partial derivatives of Eq. (9) to 0.

$$\begin{aligned} &\frac{\partial L}{\partial R} = 0, \quad \therefore \sum_{i=1}^{N}\alpha_i = 1 \\ &\frac{\partial L}{\partial c} = 0, \quad \therefore c = \sum_{i=1}^{N}\alpha_i x_i \\ &\frac{\partial L}{\partial \xi_i} = 0, \quad \therefore C - \alpha_i - \gamma_i = 0 \end{aligned} \tag{10}$$

With Eqs. (9) and (10), the objective function can be reconstructed:

$$\begin{aligned} &\max L(\alpha) = \sum_{i=1}^{N}\alpha_i(x_i \cdot x_i) - \sum_{i=1,j=1}^{N}\alpha_i\alpha_j(x_i \cdot x_j) \\ &s.t.\ \sum_{i=1}^{N}\alpha_i = 1, 0 \leq \alpha_i \leq C \end{aligned} \tag{11}$$

In practice, the inner product $(x_i \cdot x_j)$ is replaced by a kernel function $K(x_i \cdot x_j)$ that satisfies Mercer's theorem. In this paper, the RBF kernel function is used, and then Eq. (11) is transformed to the following form:

$$\max L(\alpha) = \sum_{i=1}^{N} \alpha_i K(x_i \cdot x_i) - \sum_{i=1,j=1}^{N} \alpha_i \alpha_j K(x_i \cdot x_j)$$
$$s.t. \sum_{i=1}^{N} \alpha_i = 1, 0 < \alpha_i < C \tag{12}$$

The value of α_i can be obtained by solving Eq. (12). Target data points corresponding with $0 < \alpha_i < C$ are support vectors. Then the radius R of the hypersphere can be acquired by any support vector x_{sv}.

$$R^2 = K(x_{sv} \cdot x_{sv}) - 2\sum_{i=1}^{N} \alpha_i K(x_i \cdot x_{sv}) + \sum_{i=1,j=1}^{N} \alpha_i \alpha_j K(x_i \cdot x_j) \tag{13}$$

2.4 RUL Prediction

To predict the RUL in the on-line phase, two curves (the degradation state curve and the radius curve) must be obtained first in the off-line phase.

Degradation State Curve. By EM algorithm, the parameters $\boldsymbol{\pi}$, $\mathbf{C}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ of the MoG-BBN are estimated, which allows us to obtain the state sequence of the training data through Eq. (2), and the degradation state curve (Fig. 3). The time duration for which the component in the off-line phase has been in each state can be computed based on the curve, as Eq. (14),

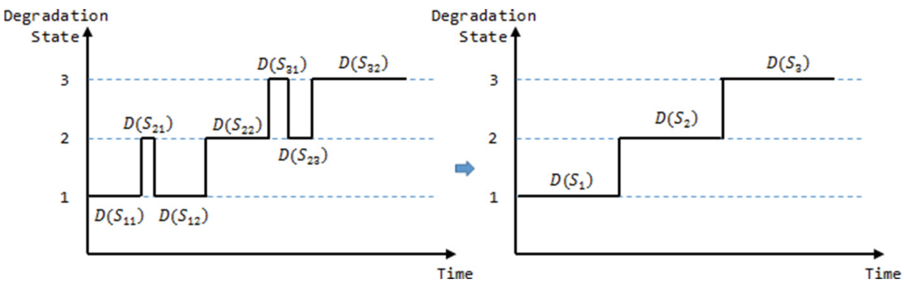


Fig. 3. The degradation state curve of the MoG-BBN.

$$T(S_d) = \sum_{\omega=1}^{\Omega} T(S_{d\omega}) \tag{14}$$

where $T(S_d)$ stands for the total time duration of the state d, Ω represents the number of consecutive visits.

Radius Curve. Suppose that the training set contains N vectors of objects $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$, and the objects $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n_0}\}$ correspond to the healthy state at the beginning. The SVDD models are trained sequentially with each sub data set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, $n \in \{n_0, n_1, \ldots, N\}$, to get the corresponding SVDD hypersphere's radius R_n. Then we obtain the radius change $\mathbf{R} = \{R_{n_0}, R_{n_1}, \ldots, R_N\}$ with the time evolution, as shown in Fig. 4.

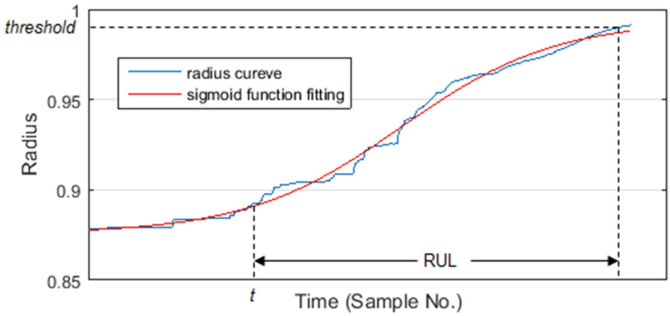


Fig. 4. Radius curve and the Sigmoid function fitting.

RUL Prediction. To predict the remaining useful life of the component in the real-time signal in the on-line phase, first we characterize the degradation state of the component by Eq. (2), then we determine whether current state of the component match the trained MoG-BBN model. Concretely, supposed that the characterized degradation state is s at time t while the degradation state is $\bar{s}$ according to the right part of Fig. 3 (obtained by training data). If s is equal to $\bar{s}$, it matches the model, otherwise it doesn't.

Denote $L(t)$ as the prediction of remaining useful life at time t. If s is equal to $\bar{s}$, $L(t)$ can be computed as following.

$$\begin{aligned} L(t) &= \sum_{d=1}^{a} P(D=d) L_d(t) \\ &= \sum_{d=1}^{a} P(D=d) \left(\sum_{i=d}^{a} T(S_i) - \bar{T}_d(t) \right) \end{aligned} \tag{15}$$

In Eq. (15), $L_d(t)$ denotes the RUL when the component is in the degradation state d, and $\bar{T}_d(t)$ denotes the past time for which the component has been in the degradation state d.

When s is not equal to $\bar{s}$, the Sigmoid function curve fitting will be implemented by the radius curve obtained by SVDD. Figure 4 illustrates how the RUL at time t is estimated.

3 Experiments and Discussions

The RUL prediction method of bearings proposed previously is tested by the condition monitoring data from NASA's prognostics data repository [2]. During the experiments, raw data of bearing 1 and bearing 4 in the test #2 are used, and both of them can be considered failed at the end.

First, wavelet packet decomposition was applied to the data, where the number of decomposition levels is set to 3 and the wavelet base is db4 (Fig. 5). Then two experiments were implemented: in experiment #1, the raw data collected from bearing 1 was divided into training data and test data, in experiment #2 the dataset of bearing 1 was used to as training data and the dataset of bearing 4 was used as test data. A detailed discussion of the experiment #1, and the results of both experiments are given as following.

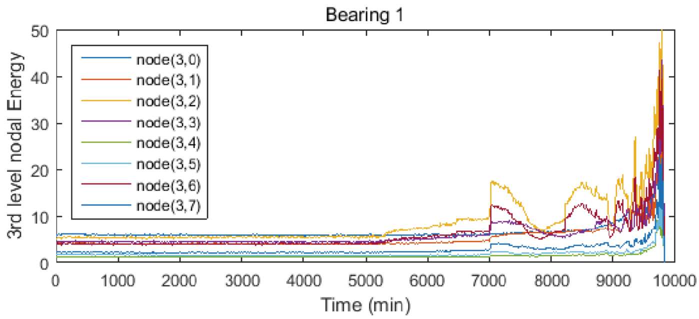


Fig. 5. WPD is applied to raw signal of bearing 1.

In the off-line phase, the sojourn time of each state of the trained MoG-BBN model is shown in Table 1, the radius curve is shown in Fig. 6. The reason why the radius curve starts at 4500 min is that the former data was collected when the bearing was in the healthy state. And the SVDD model was first trained at 4500 min with all the former data. Sigmoid function was used in the curve fitting of the radius curve, for the bearing has been in the healthy state for a long time according to Table 1, and in this period, the radius growth of the hypersphere is slow, while at the end of bearings' life cycle, the slack variable ξ_i in the objective function of SVDD makes the radius grow slow, too. From Fig. 6, we can observe that the radius curve fits well with the Sigmoid function.

Table 1. Sojourn time of each degradation state in Expt. #1

Degradation state	1	2	3
Sojourn time (min)	5090	1890	2870

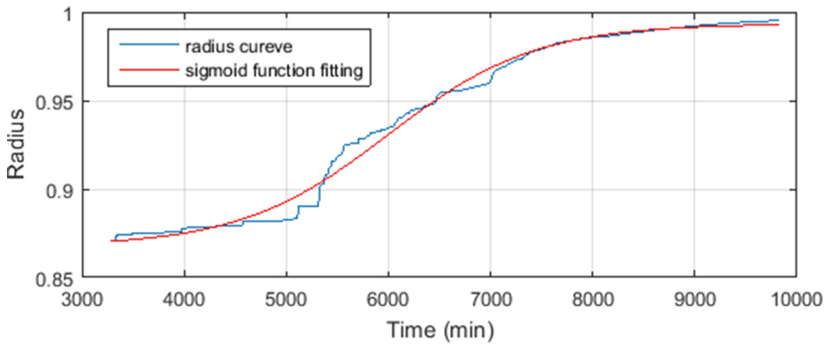
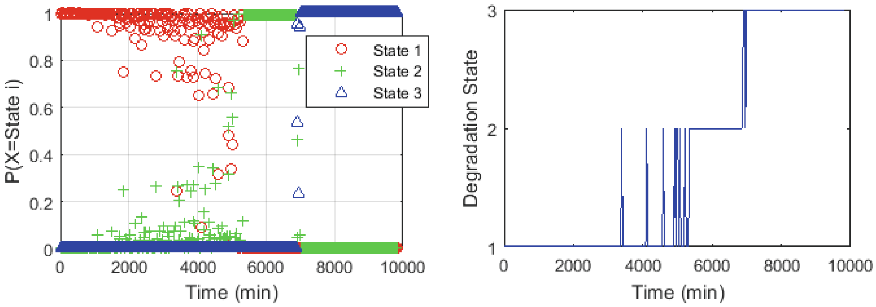


Fig. 6. Radius curve generated by the training data.



(a) Probability of each degradation state. (b) Degradation state characterization result.

Fig. 7. Degradation state characterization in Expt. #1.

In the on-line phase, Fig. 7 depicts how the degradation states are characterized after the validate data are fed into the MoG-BBN model. We could observe that the bearing had been in the healthy state for almost 50 % of its whole lifetime, which matches the result in Table 1. It can be explained that the training data and validate data is collected from the same bearing.

The RUL prediction results of the method based on MoG-BBN without GA and the method proposed by this paper are compared in Figs. 8 and 9.

From Figs. 8 and 9, we can observe that the predicted remaining useful life converges at the end of predictions. With the optimization of the initial value by genetic algorithms and the combination of the SVDD model, our method has an excellent performance in the RUL predictions of bearings. It has higher prediction accuracy, and better generalization ability and robustness as well.

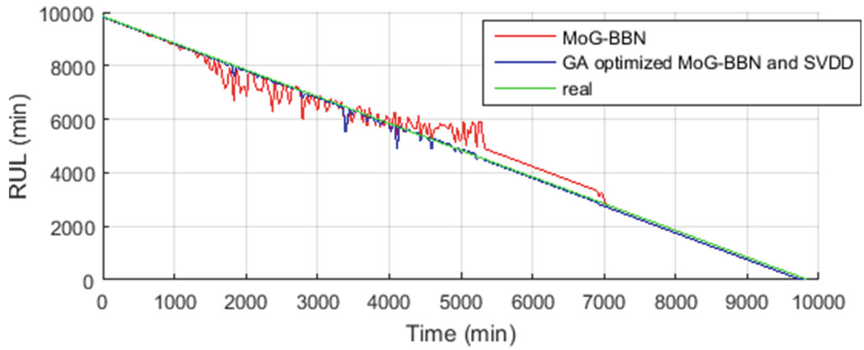


Fig. 8. RUL prediction result of experiment #1.

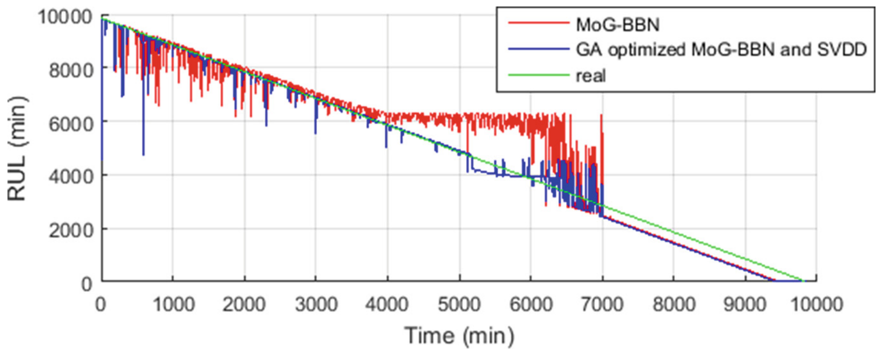


Fig. 9. RUL prediction result of experiment #2.

4 Conclusions

A method based on the MoG-BBN and SVDD for RUL prediction of bearings is proposed in this paper. WPD is chosen to extract features because it has sufficient high-frequency resolution, which contains the most useful fault information of bearings. The MoG-BBN model is a useful tool to predict the RUL with high accuracy when the work condition of bearings in on-line phase is very similar to the training data. However, different initial value may lead to different performance, the genetic algorithm is used to overcome this deficiency and acquire good stability with randomly generated initial parameters. And, an appropriate combination of the MoG-BBN and the SVDD model could improve the generalization capability and ensure the accuracy of the prediction at the same time.

Acknowledgement. This work was supported by the National Hig-Tech. R&D (863) Program (No. 2015AA042102) in China.

References

1. Jammu, N.S., Kankar, P.K.: A review on prognosis of rolling element bearings. Int. J. Eng. Sci. Technol. **3**(10), 7497–7503 (2011)
2. Lee, J., Qiu, H., Yu, G., Lin, J.: Rexnord Technical Services. IMS, University of Cincinnati. "Bearing Data Set", NASA Ames Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA (2007). http://ti.arc.nasa.gov/project/prognostic-data-repository
3. Gebraeel, N.Z., Lawley, M.A.: A neural network degradation model for computing and updating residual life distributions. IEEE Trans. Autom. Sci. Eng. **5**(1), 154–163 (2008)
4. Tian, Z., Wong, L., Safaei, N.: A neural network approach for remaining useful life prediction utilizing both failure and suspension histories. Mech. Syst. Signal Process. **24**(5), 1542–1555 (2010)
5. Huang, R., Xi, L., Li, X., et al.: Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. Mech. Syst. Signal Process. **21**(1), 193–207 (2007)
6. Tobon-Mejia, D.A., Medjaher, K., Zerhouni, N., et al.: Hidden Markov models for failure diagnostic and prognostic. In: Prognostics and System Health Management Conference (PHM-Shenzhen), pp. 1–8. IEEE (2011)
7. Tobon-Mejia, D.A., Medjaher, K., Zerhouni, N., et al.: A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models. IEEE Trans. Reliab. **61**(2), 491–503 (2012)
8. Shen, Z., He, Z., Chen, X., et al.: A monotonic degradation assessment index of rolling bearings using fuzzy support vector data description and running time. Sensors **12**(8), 10109–10135 (2012)
9. Wang, H., Chen, J.: Performance degradation assessment of rolling bearing based on bispectrum and support vector data description. J. Vibr. Control **20**(13), 2032–2041 (2014)
10. Sloukia, F., El Aroussi, M., Medromi, H., et al.: Bearings prognostic using mixture of gaussians hidden Markov model and support vector machine. In: 2013 ACS International Conference on Computer Systems and Applications (AICCSA), pp. 1–4. IEEE (2013)
11. Zhang, X., Kang, J., Jin, T.: Degradation modeling and maintenance decisions based on Bayesian belief networks. IEEE Trans. Reliab. **63**(2), 620–633 (2014)
12. Wald, R., Khoshgoftaar, T.M., Sloan, J.C.: Using feature selection to determine optimal depth for wavelet packet decomposition of vibration signals for ocean system reliability. In: 2011 IEEE 13th International Symposium on High-Assurance Systems Engineering (HASE), pp. 236–243. IEEE (2011)
13. Zhang, T., Zhang, H., Wang, Z.: Float encoding genetic algorithm and its application. J. Harbin Inst. Technol. **32**(4), 59–61 (2000)
14. Tax, D.M.J., Duin, R.P.W.: Support vector data description. Mach. Learn. **54**(1), 45–66 (2004)