An Improved Method for Constructing Health Factors of Rolling Bearing

Tao Shu\textsuperscript{1, a}, Yichi Zhang\textsuperscript{2, b\ast}, Yan Xu\textsuperscript{3, c}, Pengxiang Zhang\textsuperscript{4, d}

\textsuperscript{1}Air and Missile Defense College, Air Force Engineering University, Xi’an Shaan Xi 710051, China

\textsuperscript{2}Air and Missile Defense College, Air Force Engineering University, Xi’an Shaan Xi 710051, China;

\textsuperscript{3}Mathematics and Statistics Science College, Shaanxi

\textsuperscript{4}Air and Missile Defense College, Air Force Engineering University, Xi’an Shaan Xi 710051, China

e-mail: kjgcdx008@163.com, e-mail: 3332159839@qq.com,
de-mail: Zlin121212@163.com
\textsuperscript{\ast}Corresponding author’s e-mail: Zyc412181588@ahlctl.com

\textbf{Abstract.} Aiming at the problems of low performance index and redundant information of multi-class physical health factors, a method of building rolling bearing health factors based on improved restricted Boltzmann machine is proposed. By fusing multi-class physical health factors, the virtual health factor of rolling bearings is constructed. Extracting physical health factors in time domain and frequency domain of rolling bearing vibration monitoring signals as inputs; the performance degradation mechanism model is added to the regularization term of the Restricted Boltzmann Machine, the performance degradation information contained in the input data is mined, the virtual health factor construction model of rolling bearings is built, and the model parameters are adjusted by the health factor evaluation criteria to improve the performance of the virtual health factor. The life cycle test of rolling bearings shows that compared with the principal component analysis method, the monotonicity of virtual health factors based on the improved restricted Boltzmann mechanism is increased by 147.6\% and 38.5\%, the trend is increased by 113.8\% and 16.1\%, and the robustness is increased by 60\% and 8.42\%, respectively.

1. INTRODUCTION
Rolling bearing is the key supporting part of rotating machinery. The running state is directly related to the reliability and safety of mechanical equipment. Therefore, monitoring its condition and evaluating its health status are helpful to formulate accurate maintenance plans, realize the transformation from post-maintenance, regular maintenance to preventive maintenance of mechanical equipment, and reduce potential safety hazards and economic losses caused by failure shutdown. It has important engineering application significance [1]. The health assessment methods of mechanical equipment are mainly divided into two categories: health assessment methods based on mechanism model and data-driven health assessment methods. The health status assessment method based on mechanism model
describes the performance degradation process by establishing the mathematical model of mechanical equipment failure mechanism. Among them, the parameters of the mechanism model are related to material characteristics and pressure level, which are usually determined by multi-body dynamic analysis, finite element analysis and other technical means [2-3]. Paris-Erdogan model is proposed in the literature to describe the development process of mechanical cracks. Since then, the model has been widely used in the field of mechanical equipment health assessment. In addition to PE model, Zhiling Peng [4] and others analyzed the force distribution of high-speed transmission mechanism, determined the factors affecting the wear prediction model, and then established a mechanism model to determine the wear amount.

In order to avoid the fluctuation and one-sidedness of a single physical health factor, multiple physical health factors in time domain, frequency domain and time-frequency domain of mechanical equipment are often extracted in practical applications to fully reflect the performance degradation process of mechanical equipment. However, too many physical health factors often lead to information redundancy, which reduces the calculation efficiency and affects the prediction accuracy. Principal component analysis is one of the most widely used data dimension reduction algorithms, which transforms high-dimensional fault features into low-dimensional feature input by reconstructing k-dimensional features. Li et al. [5] extracted five-time domain features, three frequency domain features and three time-frequency domain features of the rolling bearing to form a physical health factor set. Kernel principal component analysis is used to reduce the dimension of physical health factor set, a virtual health factor set composed of three components is obtained, and the first principal component is selected as the virtual health factor of rolling bearings, and its performance index is higher than 11 physical health factors. Meng Wenjun and others used PCA to fuse multiple rolling bearing performance indexes monitored in real time to obtain virtual health factors after dimension reduction and establish a dynamic life prediction model for rolling bearings. Although the above virtual health factor construction method based on PCA can effectively extract performance degradation information from multiple types of monitoring data and reduce information redundancy, PCA, as a data-driven method, only starts from the perspective of data and belongs to a linear dimension reduction model, and its adaptability to some nonlinear and non-stationary performance degradation processes needs to be enhanced.

Restricted Boltzmann Machine is a special kind of Markov random field, which has a network structure composed of visible layer and hidden layer. The constrained Boltzmann machine uses the training data to adjust the node connection weights and bias parameters, so that the probability distribution represented by the corresponding constrained Boltzmann machine under the parameters is as consistent as possible with the training data, thus achieving the purpose of feature fusion and dimension reduction. However, the virtual health factor curve constructed by the above method still has large random fluctuations. A restricted Boltzmann machine model with fitting regularization term is proposed. The rolling bearing performance degradation mechanism model is organically integrated into the restricted Boltzmann machine regularization term, the performance degradation information contained in the rolling bearing monitoring data is deeply mined, and various physical health factors are fused to effectively improve the performance of virtual health factors.

2. THEORETICAL INTRODUCTION

2.1 Restricted Boltzmann Machine
Restricted Boltzmann machine is a linear logarithmic Markov random field with a network structure consisting of a visible layer and a hidden layer, as shown in Figure 1. Among them, the visible layer can also be called the input layer and is represented by \( v \); the hidden layer is the output layer and is represented by \( h \). The constrained Boltzmann machine is an energy-based model. For a given set of states \((v, h)\), assuming that both the visible layer and the hidden layer obey Gaussian distribution, its energy function is:
\[ E(v,h) = \sum_{i=1}^{n_v} \frac{(v_i - a_i)^2}{2\sigma_i^2} + \sum_{j=1}^{n_h} \frac{(h_j - b_j)^2}{2\sigma_j^2} - \sum_{i,j} \sigma_i \sigma_j \omega_{ij} \]

Wherein the \( v_i \) and \( h_j \) represents the state of the visible layer \( i \) and the hidden layer \( j \), \( n_v \) and \( n_h \) are the number of cells of the visible layer and the hidden layer, \( a_i \) and \( b_j \) are the Gaussian mean value of the visible layer and the hidden layer, \( \sigma_i \) and \( \sigma_j \) are the standard deviation of the visible layer and the hidden layer, and \( \omega_{ij} \) is the connection weight between the visible layer and the hidden layer.

Fig.1 RBM model structure

According to the energy function defined by equation (1), the joint probability distribution of the state \((v,h)\) can be obtained:

\[ P_\theta(v,h) = \frac{1}{Z_\theta} e^{-E_\theta(v,h)} \quad (2) \]

Among them, \( Z_\theta = \sum_{v,h} e^{-E_\theta(v,h)} \) is a normalization factor.

The probability distribution \( P_\theta(v) \) of the observed data \( v \) can be obtained, which corresponds to the edge distribution of the joint probability density distribution \( P_\theta(v,h) \), also known as the likelihood function, and its expression is:

\[ P_\theta(v) = \frac{1}{Z_\theta} \sum_{h} e^{-E_\theta(v,h)} \quad (3) \]

Given the training samples, the training process of the restricted Boltzmann machine is to adjust the parameter \( \theta \) so that the probability distribution represented by the corresponding restricted Boltzmann machine under the parameters is as consistent as possible with the training data. The goal of training restricted Boltzmann machines is to maximize the following likelihood functions:

\[ L_\theta = \prod_{i=1}^{n_v} P(v_i) \quad (4) \]

For the convenience of processing, use \( \ln L_\theta \) instead \( L_\theta \), therefore, the goal of training restricted Boltzmann machines is to maximize the following logarithmic function:

\[ \ln L_\theta = \ln \prod_{i=1}^{n_v} P(v_i) = \sum_{i=1}^{n_v} \ln P(v_i) \quad (5) \]

In this paper, the gradient rise method is used to maximize equation (5), and the iterative scheme is:

\[ \theta := \theta + \eta \frac{\partial \ln L_\theta}{\partial \theta} \quad (6) \]

Among them, \( \eta > 0 \) is the learning rate.
2.2 Improved Restricted Boltzmann Machine

In the training set composed of multiple life cycle health features of the restricted Boltzmann machine, there is a lot of noise interference in the health feature curve, and the trend of the curve is not strong, which leads to the defects of poor trend and poor discreteness in the health factor feature set after dimension reduction by the restricted Boltzmann machine. To solve the above problems, this paper intends to add a new regularization term to the objective function to improve the performance of health factor curve. Usually, the trend of health factor curve is mainly manifested in the linearization degree between health factor value and operation cycle. Based on this idea, a new regularization term is added to the objective function, as shown in Equation (7).

\[ L_R(\theta) = L(\theta) + \alpha R(\theta) \]  

In the above equation, \( L_R(\theta) \) represents the objective function for adding regularization terms, \( A \) represents the original objective function; \( R(\theta) \) represents a regularization term; \( \alpha \) represents the regularization constant. In order to improve the ability of constrained Boltzmann machine to learn linear trend features, it is assumed that all hidden layer node values change linearly with time, and the slope of its change is taken as the regularization term, and the expression of the regularization term is:

\[ R(\theta) = \sum_j \gamma_j^2 \]  

Among them, \( \gamma_j \) represents the slope of the value of the \( j \) hidden cell changing with time, and the expression is:

\[ \gamma_j = \frac{\sum_{k=1}^{N} (t_k - T)(y_k - \bar{y})}{\sum_{k=1}^{N} (t_k - T)^2} (k = 1, ..., N) \]  

In Equation (9), \( N \) is the total number of training samples; \( t_k \) represents the \( k \) run cycle; \( T = \sum_{k=1}^{N} t_k \) represents the average value of the total operation cycle; \( y_k \) represents the hidden node value corresponding to the \( k \) operation cycle; \( \bar{y} = \frac{\sum_{k=1}^{N} y_k}{N} \) represents the average value of all hidden nodes.

Assuming that the number of operation cycles with equal time intervals is taken, that is \( t_k = k \), Equation (9) can be expressed as:

\[ \gamma'_j = \frac{Y_{1,j} + 2Y_{2,j} + 3Y_{3,j} + \cdots + NY_{N,j} - N\bar{y}1 + N}{1^2 + 2^2 + \cdots + N^2 - N\left(\frac{1 + N}{2}\right)^2} \]  

\[ Y_{k,j} = \sigma\{\sum_i \omega_{ij}^k Y_{i,j} + C_j\} \]  

Represents the \( k \) hidden node value corresponding to the \( j \) run cycle. \( \sigma\{\} \) represents the sigmoid function, and its expression is:

\[ \sigma(x) = \frac{1}{1 + e^{-x}} \]

According to Equations (10) and (11), the regularization term can be updated by calculating the gradient, and the gradient calculation equation is:
After the regularization term is added, the regularization term is updated by gradient updating method, and the original objective function is updated by contrast divergence algorithm. Finally, the updated regularization term parameters and the updated original objective function parameters are added to obtain the updated value of the new objective function parameters.

Adding regularization terms can improve the ability of constrained Boltzmann machines to learn trend features, but it cannot reduce the problem of large divergence of health factor curves caused by noise interference, which leads to large confidence intervals obtained from residual life prediction, which is not conducive to the formulation of subsequent maintenance plans.

This paper fits the curve of each hidden layer node value changing with time in the process of updating the regularization term, and updates the hidden layer node value by using the fitted function value. The fitting function adopts the exponential function model commonly used in the process of rolling bearing performance degradation, as shown in Equation (13).

$$f(t) = a \exp(bt) + c$$

Wherein, $t$ is the running time; $a$, $b$, and $c$ represent model parameters.

3. CONSTRUCTION MODEL OF HEALTH FACTORS BASED ON IMPROVED RBM

3.1 Model's construction Evaluation criteria for health factors.

In order to fully extract the performance degradation information contained in the monitoring signal, a variety of physical health factors (time domain, frequency domain and multi-band information entropy) are constructed. Different types of physical health factors have different scale ranges, so max-min normalization method is used to normalize the remaining physical health factors. The normalization equation is:

$$x_i' = \frac{x_i - x_{i,\text{min}}}{x_{i,\text{max}} - x_{i,\text{min}}}$$

Wherein $x_{i,\text{min}}$ and $x_{i,\text{max}}$ represent the minimum value and the maximum value of the i-th physical health factor in the monitoring period, respectively. After normalization, all kinds of physical health factors are transformed into the normalized scale range. Then, the super parameters of the improved RBM network are set, and the improved RBM network is trained by using the normalized training set. When the reconstruction error is less than the set threshold, the training is completed. The Test set data is input into the trained network, and the virtual health factor of test data is obtained. Using the health factor evaluation criterion in step 2 to evaluate the constructed virtual health factor, adjusting network parameters according to the evaluation result, repeating steps 4 and 5 until the optimal evaluation result is obtained, and stopping to obtain the virtual health factor of the test device.
3.2 Evaluation criteria for health factors

3.2.1 Monotonicity Trend

The performance degradation process of rolling bearings is irreversible. Therefore, the health factors reflecting the performance degradation degree of rolling bearings should also have monotonous increasing or monotonous decreasing characteristics, which is called monotonicity. The equation (2) for calculating the monotonicity of health factors is:

$$\text{Mon}(X) = \frac{1}{K-1} \left| \text{No.of } d/dx > 0 - \text{No.of } d/dx < 0 \right|$$

In the above equation, $X = \{x_i\}_{i=1}^K$ is a sequence composed of health factors, wherein $x_i$ represents the value of health factors at $t_i$; $K$ indicates the number of all health factors in the sequence; $d/dx = x_{i+1} - x_i$ is a differential value representing a sequence of health factors; $\text{No.of } d/dx > 0$ and $\text{No.of } d/dx < 0$ are positive and negative differential values, respectively. The monotonicity value $\text{Mon}(X)$ of the health factor is located in the interval $[0,1]$, and the closer the value is to 1, the better the monotonicity of the health factor is.

3.2.2 Trend

The performance of rolling bearings will gradually deteriorate with the increase of running time. Therefore, the health factor value representing the degradation process should have correlation with running time, which is called trend. The equation for calculating the trend of health factors is:

$$\text{Tre}(H,T) = \frac{\left| \sum_{i=1}^K (h_i - \bar{H})(t_i - \bar{T}) \right|}{\sqrt{\sum_{i=1}^K (h_i - \bar{H})^2} \sum_{i=1}^K (t_i - \bar{T})^2}$$

Among them, $h_i$ is the value of the health factor at $t_i$; $\bar{H} = (1/K)\sum_{i=1}^K h_i$ is the average value of health factors in the whole life cycle; $\bar{T} = (1/K)\sum_{i=1}^K t_i$ is the average value of each time period in the whole life cycle. The trend value $\text{Tre}(H,T)$ of the health factor is located in the interval $[0,1]$, and the closer the value is to 1, the better the trend of the health factor is.

3.2.3 Robustness

The robustness of health factor curve is defined according to the fluctuation degree of performance degradation sequence, which describes the robustness of health factor to interference. The equation for calculating the robustness of the health factor curve is defined as:

$$\text{Rob}(H) = \frac{1}{K} \sum_{i=1}^K \exp \left( -\frac{h_i - \bar{h}}{h_i} \right)$$

Wherein, $H = (h_1,h_2,...,h_K)$ is a characteristic sequence of health factor; $\bar{H} = (\bar{h}_1,\bar{h}_2,...,\bar{h}_K)$ is trend sequence characteristic sequence of corresponding health factors. The range of robustness is $[0,1]$. The smoother the curve of health factor changing with life (running time), the greater the robustness index and the better the robustness.

4. EXPERIMENTAL VERIFICATION

This section uses the 2012 IEEE PHM data set to verify the proposed model. The whole experimental platform is mainly composed of four parts: data acquisition module, power plant, test bearing and load module. In this experiment, 17 rolling bearings under three different working conditions were tested. The acceleration sensor is installed on the outer ring of the rolling bearing and can obtain vibration acceleration signals. The sampling frequency is 25.6 kHz, each sampling includes 2560 points (for example, 0.1 s), and the sampling is repeated every 10s as a sampling period.
The life cycle data of the first rolling bearing under working condition 1 is selected for analysis. The vibration acceleration signal in the horizontal direction is shown in Figure 2. As can be seen from Figure 4, with the continuous increase of the experimental period, the vibration amplitude increases continuously, but it is difficult to accurately judge the early fault occurrence point directly through the vibration acceleration signal, thus it is difficult to carry out the remaining life prediction work.

31 physical health factors (11-time domain health factors, 12 frequency domain health factors, and 8 information entropy health factors in different frequency bands) are extracted from the vibration acceleration signal to form a physical health factor set, and the health factor set is input to the limited Boltzmann machine added with fitting regularization terms for training, and the reduced dimension virtual health factor set is obtained. Taking Bearing1_1 rolling bearing as an example, the performance degradation curves characterized by different health factors are shown in Figure 5. In order to facilitate visualization, only 9 kinds of physical health factor curves with better performance were selected for comparison. Among them, X axis numbered 1 is a virtual health factor performance degradation curve based on improved RBM, numbered 2 ~ 5 are time domain health factors, 6 ~ 8 are frequency domain health factors and 9 are information entropy health factors. As can be seen from Figure 3, the improved RBM virtual health factor construction model has learned the performance degradation information contained in the physical health factor, and the monotonicity and trend of the constructed performance degradation curve are obviously enhanced.

In order to further analyze the ability of different virtual health factors to construct models to extract performance degradation characteristics, the improved RBM model was compared with the original RBM model and principal component analysis model. The input data of the three models are the same, all of which are 31 kinds of physical health factors (11 kinds of time domain physical health factors, 12 kinds of frequency domain physical health factors and 8 kinds of information entropy physical health factors). Principal Component Analysis is an unsupervised linear dimensionality reduction method, which is the most widely used data dimensionality reduction method at present. The main idea of PCA is to map n-dimensional features to k-dimensional features, which are brand-new orthogonal features.
and also become principal components. They are k-dimensional features reconstructed on the basis of the original n-dimensional features. The principle of PCA is to find a set of mutually orthogonal coordinate axes sequentially from the original space. The selection of new coordinate axes is closely related to the data. Among them, the first coordinate axis is to select the direction with the largest variance in the original data; the second coordinate axis is to select the direction with the largest variance in the plane orthogonal to the first coordinate axis; the third coordinate axis is to choose the direction with the largest variance in the plane orthogonal to the first and second axes. By analogy, a total of new axes is obtained. Among them, most of the variance is contained in the front K coordinate axes, and the variance contained in the back coordinate axes is almost 0. Therefore, the feature dimension with almost zero variance can be ignored, thus realizing the dimension reduction of data features. In order to facilitate analysis and comparison, the number of dimension reduction of principal component analysis method is also set to 3, and the performance characteristic curve with the best performance is selected for comparison. The result is shown in Figure 4.

Fig.4 Comparison of performance degradation curves of different health indicators

Comparing Figure 4 (a), (b) and (c), it can be seen that the virtual health factor curve based on principal component analysis and the original restricted Boltzmann mechanism has more overall noise. As can be seen from Figure 4 (a), the virtual health factor performance degradation curve constructed by principal component analysis method basically cannot effectively reflect the health status of rolling bearings. The main reason is that the vibration acceleration data of rolling bearings collected by sensors have complex characteristics such as nonlinearity and unsteady state, while the principal component analysis method belongs to a linear dimension reduction model and cannot effectively extract nonlinear features, which leads to the virtual health factor after dimension reduction not effectively reflecting the nonlinear performance degradation process. As can be seen from Figure 4 (b), the virtual health factor constructed by the original restricted Boltzmann machine model can generally reflect the degradation trend of rolling bearings, but due to the absence of the constraint of regularization penalty term, the learned health characteristics are not strong in trend, which is not conducive to the subsequent prediction of remaining life. As can be seen from Figure 4 (c), the restricted Boltzmann machine model added with fitting regularization terms can learn health features with strong trend and consistent change trend through the constraints of regularization terms. The virtual health factor curves constructed by the three models are quantitatively evaluated by using the health factor evaluation criteria, and the results are shown in Figure 5.
Fig. 5 Characteristic comparison of virtual health indicator

As can be seen from Figure 5, the virtual health factors constructed based on principal component analysis have poor monotonicity, trend and robustness. Compared with the health characteristic curve constructed by principal component analysis method, the constrained Boltzmann machine model effectively improves the trend and robustness of the health characteristic curve, which is due to the ability of learning nonlinear characteristics of the constrained Boltzmann machine. The constrained Boltzmann machine model based on the addition of fitting regularization terms has the best monotonicity, robustness and trend. By adding regularization terms, the ability of extracting trend features of restricted Boltzmann machine can be effectively improved, and the interference of external disturbances on health characteristic curves can be overcome to a certain extent, thus improving the robustness of health characteristic curves. The life cycle test of rolling bearings shows that compared with the principal component analysis method, the monotonicity of virtual health factors based on the improved restricted Boltzmann mechanism is increased by 147.6% and 38.5%, the trend is increased by 113.8% and 16.1%, and the robustness is increased by 60% and 8.42%, respectively.

5. CONCLUSION

In this paper, a virtual health factor construction method of rolling bearings based on improved restricted Boltzmann machine is proposed, which effectively improves the performance of health factors. The performance degradation mechanism model of rolling bearings is integrated into the regularization term of restricted Boltzmann machine, and the performance degradation information contained in multiple physical health factors in time domain and frequency domain is extracted to construct virtual health factors, which reduces the redundant information in multiple physical health factors and improves the performance of health factors. The experimental results show that the restricted Boltzmann machine model with fitting regularization term can effectively improve the performance of health factors.

REFERENCES
[1] Marcia L B, Elsa M P H, Kai G. More effective prognostics with elbow point detection and deep learning[J]. Mechanical Systems and Signal Processing, 2021, 146: 1-22.
[2] Wang J J, Gao R X, Yuan Z, et al. A joint particle filter and expectation maximization approach to machine condition prognosis[J]. Journal of Intelligent Manufacturing, 2019, 30(2): 605-621.
[3] Lei Y, Li N, Gontarz S, et al. A model-based method for remaining useful life prediction of machinery[J]. IEEE Transactions on Reliability, 2016, 65: 1314-1326.
[4] PENG Z L, ZHANG Y, DING M J, et al. Analysis and prediction model of wear mechanism for high-speed transmission mechanism of a missile system[J]. Journal of North University of China (Natural Science Edition), 2018, 39(2): 155-158.
[5] LI X Q, JIANG H K, XIONG X, et al. Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network[J]. Mechanism and Machine Theory, 2019, 133: 229-249.