SCIENCE CHINA
Information Sciences

• LETTER •

September 2020, Vol. 63 199202:1–199202:3
https://doi.org/10.1007/s11432-018-9734-9

Principal component analysis and belief-rule-base aided health monitoring method for running gears of high-speed train

Chao CHENG[1,2,3], Xinyu QIAO[1], Wanxiu TENG[2], Mingliang GAO[2,3], Bangcheng ZHANG[4*], Xiaojing YIN[4] & Hao LUO[5]

[1]*School of Computer Science and Engineering, Changchun University of Technology, Changchun* 130012*, China;*
[2]*National Engineering Laboratory, CRRC Changchun Railway Vehicles Co., Ltd., Changchun* 130062*, China;*
[3]*Department of Automation, Tsinghua University, Beijing* 100084*, China;*
[4]*School of Mechatronic Engineering, Changchun University of Technology, Changchun* 130012*, China;*
[5]*Academy of Astronautics, Harbin Institute of Technology, Harbin* 150001*, China*

Received 18 October 2018/Accepted 12 December 2018/Published online 26 March 2020

Citation Cheng C, Qiao X Y, Teng W X, et al. Principal component analysis and belief-rule-base aided health monitoring method for running gears of high-speed train. Sci China Inf Sci, 2020, 63(9): 199202, https://doi.org/10.1007/s11432-018-9734-9

Dear editor,
We propose a data-driven health monitoring method for running gears of a high-speed train. In the proposed approach, the principal component analysis (PCA) method is used to perform feature extraction while the belief-rule-base (BRB) method is used to perform health monitoring tasks. To improve the monitoring performance, constraints of the covariance matrix adaptive evolution strategy (CMA-ES) is used to optimize the monitoring parameters of the BRB method. Similarly, to verify the effectiveness of the proposed health monitoring approach, a set of real data of a high-speed train is used in the case studies.

The running gears of a high-speed train are a complex electromechanical system. At present, the common methods for the health estimation of complex electromechanical systems can be divided into qualitative analysis methods [1,2], data-driven methods [3–5], and analytical model-based methods [6,7]. The data-driven methods solve the gear health estimation problem of complex electromechanical systems using the monitoring data; however, less amount of effective fault data are available, which leads to the generation of inaccurate results by the estimation model. The analytical model-based methods have a high prospect for use in the mechanism analysis of an object. A complete research-analysis model is not applicable to the running gear model discussed in this study.

Thus, this study presents a PCA-BRB aided health monitoring method for running gears of a high-speed train that combines both the expert knowledge and monitoring data to establish the model. First, the key antecedent attributes in the PCA screening for the BRB model helps to simplify and improve the BRB model calculation. Subsequently, a nonlinear model was established because BRB has a good modeling ability for nonlinear data with fuzzy uncertainty or probability uncertainty [8]. The failure data of running gears provide a better solution. Finally, the parameters were updated using a CMA-ES algorithm in the initial BRB model. Furthermore, to improve the accuracy of the model, an effective health estimation model of the running gears was established.

The proposed running gear health monitoring method for a high-speed train comprises three major stages, which are described as follows.

Stage 1. Filtering out the key features using

* Corresponding author (email: zhangbangcheng@ccut.edu.cn)

PCA. PCA was performed to obtain the first principal component of the running gears. To ensure the key feature of physical meaning, the "maximum balance value" factor was used for rotation in this study. The key feature of the principal component contribution was selected at "cumulative principal component greater than 0.70" and "eigenvalue greater than 1".

Stage 2. The system of health estimation for the running gears comprises many of the belief rules of the BRB model. The BRB model can be described as follows:

$$\begin{aligned} R_k: &\text{ If } \quad x_1 \text{ is } D_1^k \bigwedge x_2 \text{ is } D_2^k \cdots \bigwedge x_M \text{ is } D_M^k, \\ &\text{then } \{(Y_1,\theta_{1,k}),\ldots,(Y_N,\theta_{N,k})\}, \text{with a rule} \\ &\text{weight } \alpha_k \text{ and attribute weight } \eta_{1,k},\eta_{2,k}, \\ &\ldots,\eta_{M,k}, \end{aligned} \tag{1}$$

where x_i $(i=1,2,\ldots,M)$ is the i-th antecedent attribute, D_i^k $(i=1,2,\ldots,M; k=1,2,\ldots,L)$ is the value of the x_i antecedent attribute in the k-th rule, L is the total number of rules in the BRB, $\theta_{j,k}$ is the belief degree of conclusion part in the k-th rule for the j-th estimation result Y_j, α_k is the rule weight of the k-th rule, and $\eta_{i,k}$ is the weight of the i-th antecedent attribute.

The final expected utility of the running gears of the high-speed train can be obtained using the following three steps.

Step 1: Calculate the antecedent attribute matching degree $\varphi_{i,j}^k$, which is the degree of matching between the antecedent attribute and the rule:

$$\varphi_{i,j}^k = \begin{cases} 1-\varphi_{i,j}^k, & k=l+1, \\ \frac{D_i^{l+1}-x_i}{D_i^{l+1}-D_i^l}, & k=l\ (D_i^l \leqslant x_i \leqslant D_i^{l+1}), \\ 0, & k=1,2,\ldots,N\ (k\neq l,l+1). \end{cases} \tag{2}$$

Step 2: Calculate the belief rules activation weight ψ_k and the antecedent attribute that can activate some belief rules in the BRB:

$$\psi_k = \frac{\alpha_k \prod_{i=1}^M (\varphi_i^k)^{\overline{\eta_i}}}{\sum_{i=1}^L \alpha_l \prod_{i=1}^M (\varphi_i^l)^{\overline{\eta_i}}}, \tag{3}$$

$$\overline{\eta_i} = \frac{\eta_i}{\max\limits_{i=1,2,\ldots,M}\{\eta_i\}}, \tag{4}$$

where $\psi_k \in [0,1]$, $k=1,2,\ldots,l$, and $\overline{\eta_i}$ denote attribute weights.

Step 3: Using the ER algorithm for reasoning [8], the final output $S(x)$ of the BRB can be obtained as

$$S(x) = (Y_j, \hat{\theta}_j),\ j=1,2,\ldots,N, \tag{5}$$

where $\hat{\theta}_j$ is relative to the belief degree of estimation result Y_j.

In the above three steps, the final expected utility of the running gears of a high-speed train $S(x)$ can be expressed as

$$y = \mu(S(x)) = \sum_{j=1}^N \mu(Y_j)\theta(j). \tag{6}$$

Stage 3. We establish the objective function $\xi(Q)$ as

$$\xi(Q) = \frac{1}{U}\sum_{n=1}^U (\tilde{y}_n - y_n)^2, \tag{7}$$

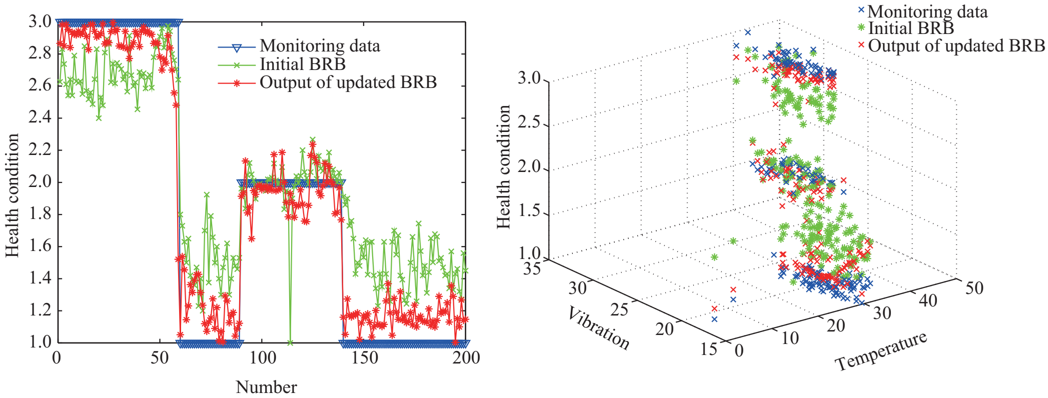
where $Q=[\alpha_k,\eta_i,\theta_{j,k},\mu(Y_j)]^{\mathrm{T}}$ is a column vector of the BRB parameters, U is the number of data of key features, the $\tilde{y}_n$ and y_n values are similar, and the constructed objective function takes the minimum value $\min_Q\{\xi(Q)\}$.

The objective function problem can be solved using the CMA-ES algorithm. The algorithm controls the optimization direction of the whole parameters by controlling the covariance matrix and helps in getting a fast convergence of small population to obtain an optimal solution. It can also be used to solve unconstrained optimization problems as well as boundary constraints.

Experiments. By analyzing the running gear system and using PCA, temperature and vibration were selected as key features. Based on the expert knowledge, the temperature and vibration reference levels, including Normal, Medium, and High, are expressed as H, U, and F, respectively. The health condition reference levels including Normal, General, and Fault, are expressed as H, G, and F, respectively. The temperature and vibration reference values were divided into three levels forming a total of nine belief rules. The health estimation model can be expressed as

$$\begin{aligned} R_k: &\text{ If temperature is } D_1^k \bigwedge \text{vibration is } D_2^k, \\ &\text{then the condition is } \{(1,\theta_{1,k}),(2,\theta_{2,k}), \\ &(3,\theta_{3,k})\}, \left(\sum_{i=1}^N \theta_{i,k} \leqslant 1\right), \\ &k \in 1,2,\ldots,9. \end{aligned} \tag{8}$$

According to expert knowledge, the estimated value was not well adapted for the monitoring data. Hence, it is necessary to optimize the parameters by CMA-ES optimization method as shown in Figure 1(c). In the first rule H AND H, it means that the temperature and vibration are in a normal state, and the health condition of the running gear system should also be kept in a normal state. The initial health condition of the running gear system was set to $\{(S_1,0.9),(S_2,0.1),(S_3,0)\}$, and $\{(S_1,0.8172),(S_2,0.1828),(S_3,0)\}$ is obtained



(a) (b)

Rule number	Vibration and temperature	Health condition distribution $\{S_1, S_2, S_3\}=\{1, 2, 3\}$
1	H AND H	$\{(S_1, 0.8172), (S_2, 0.1828), (S_3, 0)\}$
2	H AND U	$\{(S_1, 0.2664), (S_2, 0.5959), (S_3, 0.1377)\}$
3	H AND F	$\{(S_1, 0.4048), (S_2, 0.5952), (S_3, 0)\}$
4	U AND H	$\{(S_1, 0.5681), (S_2, 0.0941), (S_3, 0.3379)\}$
5	U AND U	$\{(S_1, 0.4838), (S_2, 0), (S_3, 0.5162)\}$
6	U AND F	$\{(S_1, 0.0307), (S_2, 0), (S_3, 0.9693)\}$
7	F AND H	$\{(S_1, 0.0959), (S_2, 0.6776), (S_3, 0.2265)\}$
8	F AND U	$\{(S_1, 0.0683), (S_2, 0.1072), (S_3, 0.8246)\}$
9	F AND F	$\{(S_1, 0.3371), (S_2, 0.1193), (S_3, 0.5436)\}$

(c)

Figure 1 (Color online) (a) Result of the updated BRB estimation of health status of the running gear system; (b) temperature and vibration data distribution status results; (c) CMA-ES optimized BRB parameters.

by updating and optimizing. Figures 1(a) and (b) show that the BRB model can accurately assess the state of health of the running gear system of high-speed trains.

Conclusion. In this study, we proposed a PCA-BRB aided method for the health estimation of the running gear system for high-speed trains. PCA was used to transform the features of the running gear system into the main component and the key contribution of the principal component was reversed. The proposed method was realized using the relation among BRB model, health estimation, quantitative monitoring data, and expertise based on CMA-ES algorithm for optimization and utilization of ER algorithm to estimate the reliability of the running gear system. An instance based on the health estimation of running gear system of a high-speed train shows that the BRB model can effectively reflect the health estimation of the running gear system and can be applied to practical engineering problems.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61751304, 61803044) and Jilin Scientific and Technological Development Program (Grant Nos. 20180201125GX, 20160520020JH).

References

1 Zhou D H, Ji H Q, He X, et al. Fault detection and isolation of the brake cylinder system for electric multiple units. IEEE Trans Contr Syst Technol, 2018, 26: 1744–1757
2 Ning B, Dong H R, Gao S G, et al. Distributed co-operative control of multiple high-speed trains under a moving block system by nonlinear mapping-based feedback. Sci China Inf Sci, 2018, 61: 120202
3 Miao Q, Zhang X, Liu Z, et al. Condition multi-classification and evaluation of system degradation process using an improved support vector machine. Microelectron Reliab, 2017, 75: 223–232
4 Huang D R, Ke L Y, Chu X Y, et al. Fault diagnosis for the motor drive system of urban transit based on improved hidden Markov model. Microelectron Reliab, 2018, 82: 179–189
5 Shebani A, Iwnicki S. Prediction of wheel and rail wear under different contact conditions using artificial neural networks. Wear, 2018, 406–407: 173–184
6 Swetapadma A, Yadav A. Data-mining-based fault during power swing identification in power transmission system. IET Sci Meas Tech, 2016, 10: 130–139
7 Shen T L, Eriksson L. Special focus on learning and real-time optimization of automotive powertrain systems. Sci China Inf Sci, 2018, 61: 070200
8 Yang J B, Liu J, Wang J, et al. Belief rule-base inference methodology using the evidential reasoning Approach-RIMER. IEEE Trans Syst Man Cybern A, 2006, 36: 266–285