IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 21 June 2024, accepted 23 July 2024, date of publication 29 July 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3435057

RESEARCH ARTICLE

An Interpretable Bearing Fault Diagnosis Method Based on Belief Rule Base With Interval Structured

HAIFENG WAN, MINGYUAN LIU, HAILONG ZHU, NING MA, AND WEI HE
School of Computer Science and Information Engineering, Harbin Normal University, Harbin 150025, China

Corresponding authors: Hailong Zhu (zhl.0451@163.com) and Ning Ma (maninghsd@163.com)

This work was supported in part by the Open Foundation of Key Laboratory of the Ministry of Education on Application of Artificial Intelligence in Equipment under Grant AAIE-2023-0102; in part by the Social Science Foundation of Heilongjiang Province under Grant 21GLC189; in part by Shandong Provincial Natural Science Foundation under Grant ZR2023QF010; in part by the National Natural Science Foundation of China under Grant 62227814, Grant 62203461, and Grant 62203365; in part by Shaanxi Provincial Science and Technology Innovation Team under Grant 2022TD-24; in part by China Postdoctoral Science Foundation under Grant 2023M742843; in part by the Young Talent Promotion Program of Shaanxi Association for Science and Technology under Grant 20220121 and Grant 20230125; and in part by the Natural Science Basic Research Program of Shaanxi under Grant 2022JQ-580.

ABSTRACT Rolling bearings are crucial components in industrial applications, forming the core of modern rotating machinery. The belief rule base with interval structure (IBRB-r) is a rule-based expert system designed to describe the causality of bearing faults. It is mainly applicable in situations where data within the intervals are relatively concentrated. However, in practical industrial applications, bearing operation data is often dispersed. This leads to uneven data distribution within the reference interval, affecting the model's accuracy. Additionally, the interpretability of the parameters decreases after the IBRB-r model is optimized. To address these issues, an adaptive rule matching interval structured belief rule (IBRB-Di) is proposed for bearing fault diagnosis. First, a rule matching method that can be dynamically computed within an reference interval is introduced. Also, a new computation method for the weights of rules activation within an interval is presented. These methods adapt to the dispersion of data and simplify the number of rules. Second, the projected covariance matrix adaptive evolution strategy with expert constraints (P-CMA-ES) is adopted to optimize the model. This aims to balance accuracy and interpretability. Finally, the method is experimentally validated using the Case Western Reserve University bearing dataset. This verification confirms the effectiveness of the proposed method.

INDEX TERMS Belief rule base, rolling bearing fault diagnosis, interpretable modeling, rule matching.

I. INTRODUCTION

Bearings are the key support components of mechanical rotating shafts [1]. They have a wide range of applications in many fields such as aerospace, rail transport, advanced machine tools, and construction machinery, and have a direct impact on the normal operation and performance of large equipment [2]. Given that bearing failures may lead to equipment downtime, real-time monitoring and diagnosis of bearing conditions are essential to prevent accidents and ensure safety [3]. However, unreliable prediction results and unexplained prediction processes may pose significant risks [4]. Therefore, utilising interpretable fault diagnosis methods not only helps to identify problems in a timely manner, but also further enhances the transparency and effectiveness of maintenance decisions, thus effectively reducing potential risks.

For the performance evaluation of rolling bearings, scholars at home and abroad have conducted a lot of research. In the current research, there are four main types of fault diagnosis models: data-driven models, model-driven methods, knowledge-driven models and hybrid-driven models.

(1) Data-driven modelling for bearing fault diagnosis method. This method relies heavily on a large amount of

The associate editor coordinating the review of this manuscript and approving it for publication was Juan Wang.

historical data, which is analysed to identify and predict potential fault modes. He and He proposed a bearing fault diagnosis method based on Large Memory Storage Retrieval (LAMSTAR) neural network, and the experimental results show its classification performance with high accuracy [5]. Sun and Li proposed a convolutional neural network (CNN)-based bearing fault diagnosis method using a symmetric dot pattern (SDP) to convert vibration signals into images for analysis, and tests showed the method to be highly robust [6]. Jin et al. proposed mapping and untraceable Kalman filtering (UKF). This method was able to successfully detect the bearing degradation process [7]. Li et al. proposed a transformer fault diagnosis method based on a hybrid kernel limit learning machine optimization, which improves the performance of the limit learning machine in high-dimensional nonlinear data processing [8]. This data-driven approach is also applicable to rolling bearing fault diagnosis by optimizing model parameters to improve fault identification accuracy. Zhao et al proposed a combined model based on the Kernel Extreme Learning Machine (KELM) and Weights at Failure Time (WAFT). It effectively predicts the remaining life (RUL) of rolling bearings [9]. Cui et al. proposed a three-stage learning algorithm applied to wind turbine bearing fault diagnosis, combining unsupervised learning, sensitivity analysis and regression methods to effectively improve the accuracy of fault detection [10].

(2) Model-driven fault diagnosis methods. Model-driven fault diagnosis methods focus on the use of mathematical or physical models to simulate and analyse the performance of bearings and their potential failures. Liu and Gryllias proposed an approach using a combination of physical model-based simulation and domain adversarial neural network (DANN) to successfully solve the problem of insufficient data in industrial bearing fault diagnosis [11]. Ma et al. Combining Digital Twin (DT) techniques and Enhanced Meta Migration Learning (EMTL) improves the accuracy and adaptability of bearing fault diagnosis [12]. Hou et al. analyzed interaxle bearing failures under noisy conditions via a feature extraction technique using the Laplace wavelet and an orthogonal matched tracking algorithm combined with sparse representation theory [13]. Zhang et al. assessed the severity of rolling bearing faults by developing a quantitative electrical model of stator currents. The method was validated through experiments and simulations to provide accurate fault indications under various operating conditions [14]. Pham et al. proposed a three-stage approach combining an identification model, a proportional risk model, and a support vector machine to assess machine performance degradation and predict remaining useful life (RUL). The validation results show that this approach offers significant advantages in reducing maintenance costs and improving system reliability [15].

(3) Knowledge-based approach for fault diagnosis method of bearings. Fault diagnosis based on knowledge-driven models. The model needs to draw on the practical experience and modelling knowledge of domain experts to determine the relationship between attributes and fault states over time [16]. Yin et al. successfully incorporated a priori knowledge into a deep model by means of a knowledge and data dual-driven transmission network to improve machine fault diagnosis performance [17]. Li et al. successfully improved the generalization ability of a machine fault diagnosis model trained with small sample sizes by integrating prior knowledge and combining convolutional neural networks and attention mechanisms [18].

(4) Bearing Fault Diagnosis Method for Hybrid-Drive Models. Hybrid-driven models make full use of a variety of information and effectively combine data-driven, model-driven and knowledge-driven approaches, achieving effective complementarity between the models. The data-driven approach is able to process and analyse large amounts of data and identify potential patterns and trends, but it relies on the quality and quantity of the data and may be ineffective in the face of poor data quality or changing new environments. At this point, the model-driven approach complements it by providing stable and reliable predictions based on physical and mathematical models, independent of data constraints, ensuring that effective fault analysis can still be performed despite insufficient data. However, model-driven approaches may lack sufficient flexibility and adaptability, especially under unknown or changing conditions, and this is where the introduction of knowledge-driven approaches is crucial, which utilise the experience and a priori knowledge of domain experts to fill the gap between the theoretical model and the practical application, increasing the adaptability and explanatory power of the system. Eke et al. improved the accuracy of transformer diagnosis through fuzzy C-mean clustering and fuzzy logic [19]. Similarly, Soni and Mehta combined fuzzy logic and fuzzy clustering to improve the reliability of transformer fault diagnosis [20]. This hybrid approach also has the potential to be applied in rolling bearing fault diagnosis to improve the overall performance of the diagnostic system. Hybrid drive models are more advantageous in bearing fault diagnosis because they ensure not only high accuracy but also interpretability, especially for complex systems [21]. Therefore, the hybrid drive model is more suitable for bearing fault diagnosis.

As a typical hybrid-driven approach, the belief rule base (BRB) is a complex system modeling method proposed by Yang et al. [22] in 2006. The model is developed based on a belief function and performs well in aggregating quantitative data and qualitative information [3]. Belief rule base have a wide range of applications in fault diagnosis. BRB models are able to handle various uncertainties to guarantee the accuracy of diagnostic results, and their reasoning process is transparent and traceable. Currently, BRBs are widely used in fault detection [23], safety in risk assessment [24], [25], and medical diagnosis [3], [26].

Facing the current challenge of too many attributes, the combinatorial rule explosion problem in the BRB model has

become a key challenge in complex systems. To solve this problem, He et al. [27] proposed the IBRB-r model, which not only proposes a new method for constructing interval rules with rule reliability, but also designs the corresponding inference process. In order to avoid the exponential growth in the number of rules caused by the traditional BRB combining rules in the form of cartesian product, the method of combining rules in the form of interval addition is adopted.

However, IBRB-r based bearing fault diagnosis still faces two problems. First, since bearing fault diagnosis is usually carried out in complex systems. Influenced by the environment, the measured vibration signal data fluctuates greatly and lacks continuity [28]. This makes it difficult for the model to maintain high accuracy when dealing with data fluctuations and nonlinear trends. Secondly, existing models also have difficulties in maintaining a balance between interpretability and accuracy in the diagnostic process. To address these two issues, an IBRB-Di bearing fault diagnosis method based on adaptive rule matching degree of interpretable belief rule base with interval structure is proposed. The method uses dynamic rule mapping. The system dynamically calculates the match between sample values and activation rules in the interval. This adapts to different data variations within the interval, thus mitigating the impact of data fluctuations and non-linear trends on diagnostic accuracy.In addition, reasonable interpretable constraints and optimisation objective functions are constructed, and interpretable constraints are added to the optimisation algorithm so that the parameters in the optimisation process vary around the constraints. This ensures a balance between interpretability and accuracy of the model.

The remainder of the paper is structured as follows: In Section II, problem description and IBRB-Di methodology definition. In Section III, definition of Interpretable Predictive Models.In Section IV, inference and optimisation of the IBRB-Di. In Section V, the validity of the model is tested by a case study. In Section VI, the paper is summarized, along with an outlook for future work.

II. PROBLEM DESCRIPTION AND IBRB-DI METHODOLOGY DEFINITION

This section describes the problems in the bearing fault diagnosis process of IBRB-r and the definition of the IBRB-Di method.

A. PROBLEM DESCRIPTION

There are still some issues that need to be addressed when diagnosing faults with the IBRB-r:

Problem 1. when faced with the problem of model accuracy. Especially when using the IBRB-r model, the accuracy of the model may be affected if the data distributions within the intervals are less similar. The traditional approach is to improve the model accuracy by increasing the number of dense reference intervals. However, this method increases the number of rules and is difficult to define. Therefore, this paper proposes a method to dynamically calculate the matching degree of rules in the intervals and assess the degree of activation of rules in the intervals by different samples. The accuracy and interpretability of the model are improved.

Problem 2. How to make sure the interpretability in the process of constructing, reasoning and optimising the bearing fault diagnosis model. Interpretability not only improves the user's trust and acceptance of the diagnostic results, but it also helps the user understand the cause of the fault. It realises a clear process from the indicator to the result and provides convenient decision support. Therefore, the interpretability of the bearing fault diagnosis model needs to be defined based on the BRB interpretable model definition. The basic interpretability requirements in the model modelling and reasoning process must be clarified, as shown in Eq. (1). In the process of model optimisation, the adjustment of model parameters may affect its interpretability. Therefore, we need to establish an effective mechanism to prevent the model parameters from being corrupted and ensure that the model parameters are consistent with the mechanism, as shown in the following Eq. (2).

$$Principle : \{c|c_1, c_2, \ldots, c_f\} \tag{1}$$

$$\phi_{best} = optimize\,(z, r, c) \tag{2}$$

where f represents the number of interpretable criteria. ϕ_{best} represents the best parameters after optimization, $optimize(.)$ is the optimization function, and r is defined as the set of parameters in the optimization process. z is the output of the model.

Problem 3. How to construct a reasonable model for when there are too many bearing feature attributes. The cartesian product combination rules used in traditional BRB models may lead to the generation of a large number of rules. This may significantly increase the complexity of the model and negatively affect its performance [29]. Therefore, in order to satisfy the interpretability criterion, effectively deal with the challenge of too many combining rules, and improve the accuracy, we can adopt the following nonlinear mapping relation:

$$z = h(X, c, \phi^0). \tag{3}$$

where z is the model diagnostic result. c is an interpretable group of constraint criteria. ϕ^0 is the initial parameter set by the expert. $X = \{x_1, x_2, \ldots, x_T\}$ is a prerequisite attribute of the model that serves as an input to the model. $h(.)$ denotes a nonlinear functional relationship between the system characteristics and the predicted values.

B. DEFINITION OF INTERPRETABLE BEARING FAULT DIAGNOSIS BASED ON IBRB-DI

The a priori value of the IF-THEN based rules in IBRB-Di is the fault feature x_i. When a sample falls into a certain interval it immediately activates the rules within that interval. First, the fault features are divided into multiple reference intervals (c_i, d_i). To efficiently generate reasonable rules, a rule matching degree coefficient δ_k is introduced to compute

the degree of matching with the activated rules within the interval. Further a new activation rule weight w_k is formed to calculate. Second, the fault feature is evaluated as an output term in the diagnostic result θ_l of the rule with a certain belief level $\beta_{l,k}$. Finally, the ER method with rule reliability r_k is used to implement the inference of the rule-based system.

Therefore, the IBRB-Di model in the bearing fault diagnosis model consists of a set of belief rules. The k_{th} belief rule is as follows:

$$\begin{aligned}&\text{Rule}_{\ k}:\\&\quad \text{IF } x_1 \in (c_1, d_1) \vee x_2 \in (c_2, d_2) \vee \cdots \vee x_T \in (c_T, d_T),\\&\quad \text{THEN } z \text{ is } \{(\theta_1, \beta_{1,k}), (\theta_2, \beta_{2,k}), \ldots, (\theta_N, \beta_{N,k})\},\\&\quad \text{with } r_k \text{ and } w_k,\\&\quad \text{Rule matching degree } \delta_k,\\&\quad \text{in } c_1, c_2, \ldots, c_f,\\&\quad k \in [1, S], \quad \sum_{l=1}^{N} \beta_{l,k} \leq 1 \end{aligned} \tag{4}$$

where $x_1, x_2, \ldots, x_T$ denotes a prerequisite attribute for the bearing failure characteristics to be used as input to the model. T is the number of prerequisite attributes. S denotes the total number of rules. (c_i, d_i) is the reference value interval, where $i = 1, \ldots, T$.r_k denotes the rule reliability.w_k represents the rule weights.$c_1, c_2, \ldots, c_f$ are interpretable constraints on the model. f indicates the number of interpretable guidelines.δ_k is the rule matching degree factor. The modeling process of the bearing fault diagnosis model is shown in Figure 1.

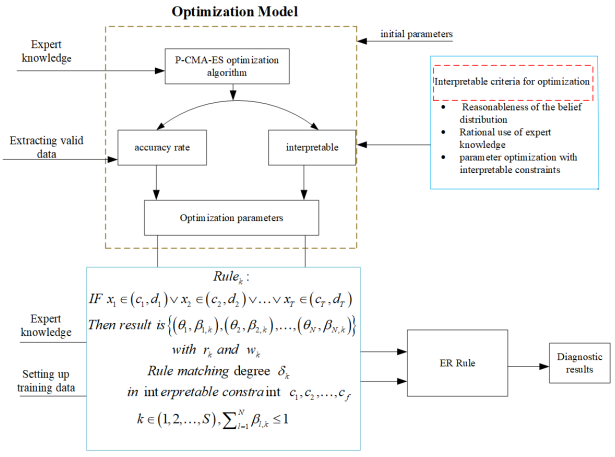


FIGURE 1. The overall structure of the IBRB-Di model.

To construct an interpretative and clear IBRB-Di model, its initial structure needs to be built reasonably first. The construction of the initial IBRB-Di model is divided into the following two steps:

Step 1: Definition of the reference interval and determination of the reference values corresponding to the results.

In the process of rule-based modeling of bearing diagnosis system, the causality of the system must first be determined, that is, the prerequisite attributes of system diagnosis and its associated results [22]. This step includes defining the fault characteristics and fault results of the bearing diagnosis system, dividing the fault characteristics into several reference intervals, and setting the reference value of the corresponding fault results as the evaluation level. Fault characteristics are used as input of IBRB-Di model, and fault results are used as output. When the fault feature sample data of the prerequisite attribute matches a certain reference interval, the relevant rules within the interval will be activated immediately [27].

Step 2: Creation of a belief table.

The process of creating a belief table begins with structuring each rule to include explicit conditions and outcomes. The conditions usually relate to fault characteristics within a previously defined reference interval, while the results are the corresponding fault results or assessment levels. The initial belief rule base contains a total number of rules S, determined by the sum of the number of reference intervals for all attributes. Next, a belief level is assigned to each rule, which is a value based on expert in-depth analyses of the bearing mechanism and accumulated experience from long-term practice. It indicates the reliability and importance of the rule in the diagnostic process. Finally, the belief table will clearly list all rules and their corresponding conditions, results and belief levels in a tabular format. Ensure that each rule can accurately map the relationship between fault characteristics and results. The process of creating the initial belief table is also the formation of the initial belief distribution.

III. DEFINITION OF INTERPRETABLE PREDICTIVE MODELS

In order to obtain reliable and interpretable models, the factors that diminish the interpretability of BRB based fault diagnosis models are first analyzed. The complexity of rolling bearing systems makes it difficult for experts to fully access and master the information of the system. The accuracy of BRB models constructed based only on limited expert knowledge may be limited. Therefore, the structure and parameters of the initial model need to be adjusted using optimization learning methods to improve the accuracy of the model. However, optimization algorithms are usually random searches for updating the initial parameters, and such unconstrained random oscillations may reduce the interpretability of the model [30]. Interpretability is expressed as the similarity between the optimized parameters and the expert knowledge, as interpretability retains valuable information from the expert insights. Therefore, ensuring the accuracy and interpretability of the model is also one of the objectives of this study.

Cao et al. developed eight criterions for BRB interpretability by analyzing the modeling, reasoning and optimization process of BRB models [30]. These criterions discuss BRB interpretability in terms of knowledge base, inference mechanism and model optimization. The IBRB-Di based bearing fault diagnosis model is constructed on the basis of

these interpretability criterions to ensure its transparency and credibility in practical applications, as shown in Figure 2.

Interpretability criterion include whether the trend of the optimized belief distribution is broadly consistent with the initial expert's belief distribution. The belief distribution quantifies the belief level of different health states through the form of probability distribution, which helps to deal with the ambiguity between health states. The expert performs mechanism analysis and trend analysis by constructing a relationship table between attributes and health states.

Mechanism analysis involves converting the bearing's operation process into a belief distribution [30]. Through the belief distribution, the system can more accurately determine the health state of the bearing. Even in the case of incomplete or noisy information, the belief distribution can still provide a reliable diagnostic result.

In bearing fault diagnosis, there are four health conditions: normal, inner ring failure, roller body failure, and outer ring failure. When the belief distribution is 0.4 for normal, 0.4 for roller body failure, and 0.1 for each of the remaining two failures, the resulting belief distribution, as shown in Figure 3, is concave. In this case, the health of the bearing may be both normal and failed, which is actually unreasonable. There is no uniform belief distribution format, but according to Cao et al.'s study, the interpretable belief distribution should be convex or monotonic.

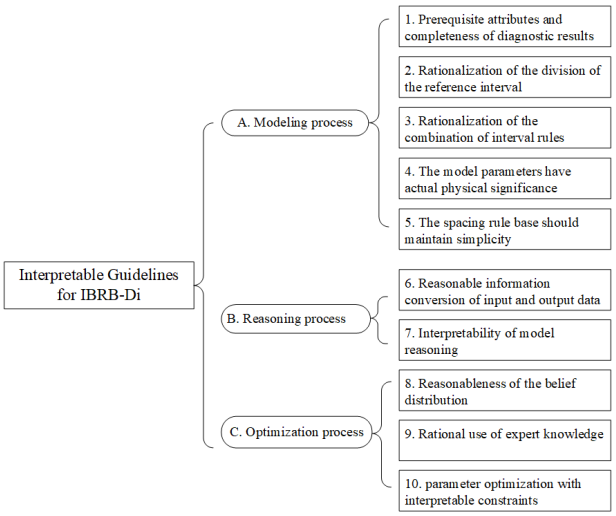


FIGURE 2. Criteria for the interpretability of models.

IV. INFERENCE AND OPTIMISATION OF THE IBRB-DI

This section describes the method of calculating the dynamic rule matching degree and rule weights within the interval. The process of inference of IBRB-Di model. The process of optimisation of IBRB-Di model.

A. RULE MATCHING FACTOR CALCULATION

In practical bearing operations, harsh environmental changes can cause discontinuities in the detected vibration signals. This results in a very broad sample data distribution, leading to a low degree of similarity in the data distribution within the intervals in the interval structure. In IBRB-r, He et al. set a large number of reference intervals densely to maintain the accuracy of the model [27]. However, this approach leads to higher model complexity. For this reason, a method to calculate from static rule matching degree to dynamic rule matching degree is proposed in IBRB-Di. When different data sample values are in the same reference interval, the rules within the interval are activated. The degree of matching of the corresponding activated rules in the interval is calculated differently by using the samples falling into the interval. Further calculations are used to compute the weights of the activation rules.

(1) Calculation of the rule match factor: When the sample value $X = \{x_1, x_2, \ldots, x_T\}$ of each attribute is input, the rule in the corresponding interval is activated immediately when it is determined to fall into the corresponding reference interval based on the judgment. The rule matching degree of the data sample is calculated dynamically to indicate the degree to which the rule is activated to understand how well the rules generated by the sample match the rules in the interval. The rule matching degree is calculated by determining the position of the sample relative to the midpoint of the interval. The following formula gives the rule matching degree of the interval corresponding to the rule:

$$\begin{aligned} &\delta_i = 1 - \frac{|x_i - med|}{(c_i - d_i)/2}, \\ &x_i \in (c_i, d_i), \delta_i \in [0, 1], \\ &\text{where med} = \frac{c_i + d_i}{2} \end{aligned} \tag{5}$$

where δ_i denotes the degree to which the sample data for i_{th} attribute falls into a certain interval of rule matching. $med = (c_i + d_i)/2$ indicates the center of the reference interval. (c_i, d_i) denotes that x_i falls into this reference interval.

The key concept of the formula is that the closer the sample is to the center of the reference interval, the better its match to the rule, and in turn the better its match to the interval rule. This indicates that the stronger the influence of the rule relative to other rules in the problem decision. Conversely, the farther the sample is from the center of the interval, the lower its match and the less influence the rule has on problem solving. The process of calculating the degree of rule matching in the interval can be divided into the following four steps, as shown in Figure 4.

(2) By calculating the rule matching coefficients within an interval, a new method for calculating the activation rule weights within an interval is designed. When the attributes of the input samples are located in a certain interval, the method for determining the activation rule weights within a specific interval is as follows:

$$w_i = [\epsilon + (1 - \epsilon) \times \delta_i] \times w, \quad \{\epsilon, w_i, w\} \in [0, 1] \tag{6}$$

where w_i denotes the value of the rule weights that are computed after the attribute sample x_i falls into the

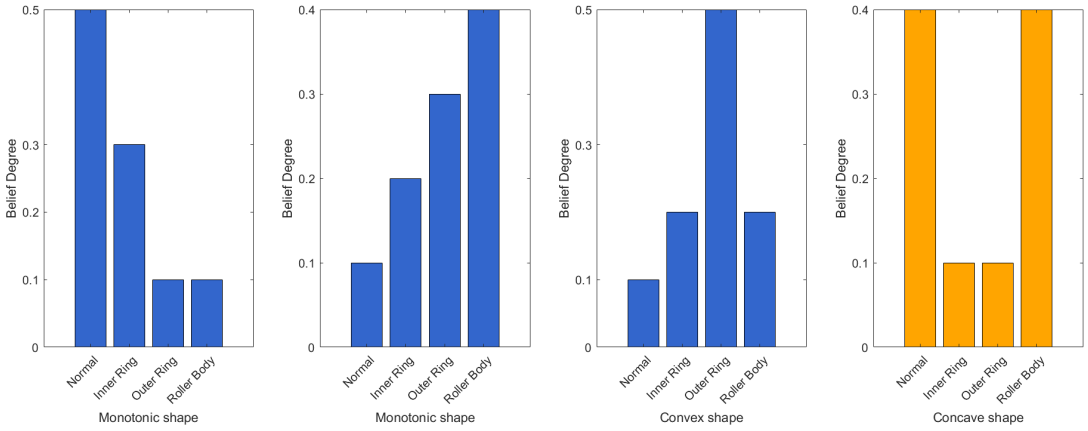


FIGURE 3. Reasonableness of belief distribution.

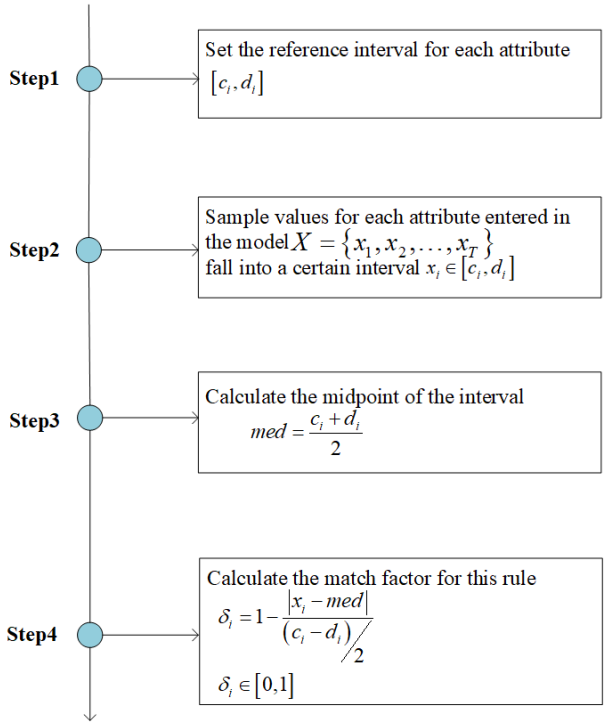


FIGURE 4. Dynamic rule matching calculation process.

i_{th} reference interval.w denotes the weight value after the initial rule weight optimisation. Because when the sample of the attribute falls into the interval, the rule in the interval will be activated immediately so ε denotes the minimum activation weight in the interval which is set by the expert.

Based on the data distribution within the interval, the expert sets the minimum activation rule weight within the interval. When the data distribution is more similar, it indicates that the sample is located in the interval with higher matching. Therefore, the rules within the interval are critical to the decision result, and the expert will increase the weight value accordingly. On the contrary, if the data distribution similarity is low, the expert will set the activation rule weights within the interval to smaller values.

B. THE REASONING PROCESS OF THE MODEL

The IBRB-Di model builds a transparent and intuitive reasoning framework by adopting the ER rules as the reasoning process for evidence. The ER rules take full account of the reliability of evidence. This framework is capable of effectively describing, transforming, and integrating multiple pieces of information under uncertainty, and ultimately forming a unified conclusion. This feature enables the IBRB-Di model to process information accurately and efficiently in complex and changing environments, providing strong support for decision making. The model's ER rule inference process is shown in Figure 5. The IBRB-Di model performs well in following the interpretability criterion proposed by Cao et al. [30]. It ensures a reasonable transformation of the model input and output information during the inference process, which not only maintains the information characteristics of the original samples, but also makes the inference results closer to the actual situation. At the same time, the reasoning process of the model maintains a high degree of interpretability, which enables users to clearly understand each step and basis of reasoning. From Figure 2 the reasoning process is required to satisfy the interpretability criterion 6 and criterion 7:

Criterion 6 (Information equivalence and reasonableness conversion): This criterion requires that input and output information be transformed into belief distributions that maintain the characteristics of the original sample information to ensure that the reasoning process is rational.

Criterion 7: (Interpretability of the reasoning engine): An reasoning engine that uses ER rules as a model ensures that the reasoning process must be transparent, traceable and logically sound.

The specific reasoning process of the ER rule is as follows:

Step 1: Converting rules into belief distributions in a bearing fault diagnosis system:

A piece of evidence ℓ_k $(k = 1, \ldots, S)$ in the bearing fault diagnosis system is converted into a belief distribution as shown below:

$$\ell_k = \left\{\left(\theta_n, \beta_{n,k}\right), n \in [1, N]; \left(\Theta, \beta_{\Theta,k}\right)\right\}$$
$$0 \le \beta_{n,k} \le 1; \sum_{n=1}^{N} \beta_{n,k} \le 1 \tag{7}$$

where $\Theta = \{\theta_1, \ldots, \theta_N\}$ denotes the identification framework.θ_n $(n = 1, \ldots, N)$ a represents the level of assessment of the results of the diagnosis of bearing faults.N indicates the number of evaluation levels.$\beta_{n,k}$ denotes the θ_n belief level for the outcome. $\beta_{\Theta,k}$ denotes that the k_{th} attribute in the recognition framework is globally ignorant.

Step 2: Calculate the basic probability mass of the rule from the evidence weights w_k $(k = 1, \ldots, S)$ and the evidence reliability r_k $(k = 1, \ldots, S)$. where $w_k \in [0, 1]$, $r_k \in [0, 1]$:

$$\tilde{m}_{n,k} = \begin{cases} 0, \theta_n = \phi \\ b_{rw,k} m_{n,k}, \theta_n \subseteq \Theta, \theta_n \ne \phi \\ b_{rw,k}\left(1 - r_k\right), \theta_n = \beta\left(\Theta\right) \end{cases} \tag{8}$$
$$b_{rw,k} = 1/\left(1 + w_k - r_k\right) \tag{9}$$
$$m_{n,k} = w_k \beta_{n,k} \tag{10}$$

where $b_{rw,k}$ is the normalization factor and satisfies the condition $\sum_{n=1}^{N} \tilde{m}_{n,k} + \tilde{m}_{\beta(\Theta),k} = 1.\tilde{m}_{n,k}$ denotes the mixed probability mass of the evaluation level θ_n under the k_{th} rule. $m_{n,k}$ denotes the probability mass under rule k_{th} at assessment level θ_n.

Step3:For S independent rules $\ell = \{\ell_1, \ell_2, \ldots, \ell_S\}$, the joint support for outcome θ is obtained by aggregating all rules through ER rule inference $\beta_{\theta,\ell(S)}$:

$$\forall \theta_n \in \Theta,$$
$$\hat{m}_{n,\ell(k)} = \left[(1 - r_k) m_{n,\ell(k-1)} + m_{\beta(\Theta),\ell(k-1)} m_{n,t}\right] + \sum_{C \cap D = \theta_n} m_{C,\ell(k-1)} m_{D,k} \tag{11}$$
$$\hat{m}_{\beta(\Theta),\ell(k)} = (1 - r_k) m_{\beta(\Theta),\ell(k-1)} \tag{12}$$
$$m_{n,\ell(k)} = \begin{cases} 0, \theta_n = \phi \\ \dfrac{\hat{m}_{n,\ell(k)}}{\sum_{C \subseteq \Theta} \hat{m}_{C,\ell(k)} + \hat{m}_{\beta(\Theta),\ell(k)}}, \theta_n \ne \phi \end{cases} \tag{13}$$
$$\beta_{\theta,\ell(k)} = \begin{cases} 0, \theta_n = \phi \\ \dfrac{\hat{m}_{n,\ell(k)}}{\sum_{C \subseteq \Theta} \mu_{C,\ell(k)}} \end{cases} \tag{14}$$

where the variable $k = 1, 2 \ldots, S.\beta_{\theta,\ell(k)}$ denotes the joint belief level of the combination of k pieces of evidence for the evaluation level θ. denotes the joint probability quality of k pieces of evidence.Ulfill the condition $m_{n,\ell(1)} = m_{n,1}$ and $m_{\beta(\Theta),\ell(1)} = m_{\beta(\Theta),1}$. The belief distribution of the relative results can be derived:

$$\ell(k) = \left\{\left(\theta_n, \beta_{n,\ell(k)}\right), n \in [1, N] \ , \left(\Theta, \beta_{\Theta,\ell(k)}\right)\right\} \tag{15}$$

Step 4: Combining the above arithmetic processes, the expected utility value obtained can be considered as the final output:

$$z = \sum_{n=1}^{N} u(\theta_n) \beta_{n,\ell(k)} + u(\Theta) \beta_{\Theta,\ell(k)} \tag{16}$$

where z denotes the final expected utility value.$u(\theta_n)$ denotes the utility of θ_n.

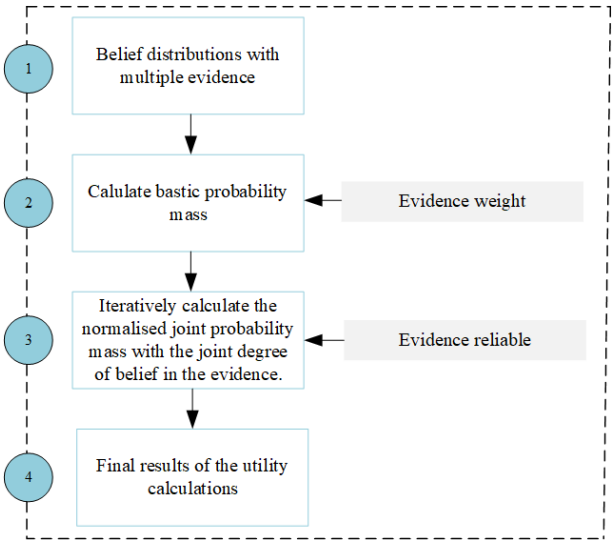


FIGURE 5. The reasoning process of the ER rule.

C. OPTIMIZATION OF MODELS

The model is difficult to accurately describe the real health condition of the bearing due to the environmental factors during bearing rotation and the limitation of expert knowledge. Searching the optimal region more efficiently through optimisation algorithms can improve the diagnostic efficiency and the accuracy of the model [32]. Therefore, in order to further improve the accuracy of the model, the optimisation of the model is also considered in this paper.

Due to its efficient parameter optimization performance, P-CMA-ES can quickly converge to the global optimal solution or local optimal solution in most cases [33], [34]. Therefore, after reasoning about the IBRB-Di model, the P-CMA-ES optimization algorithm is used to optimize the model parameters. The optimized parameters include rule reliability, rule weights, and the belief rule table. However, since the parameters of the original P-CMA-ES algorithm's optimization process randomly oscillate throughout the solution space, expert subjectivity is ignored [35]. This approach aims to achieve high model accuracy but may lead to a loss of interpretability of the optimized IBRB-Di model parameters. Therefore, interpretability constraints are added to the optimization algorithm, and the Euclidean distance

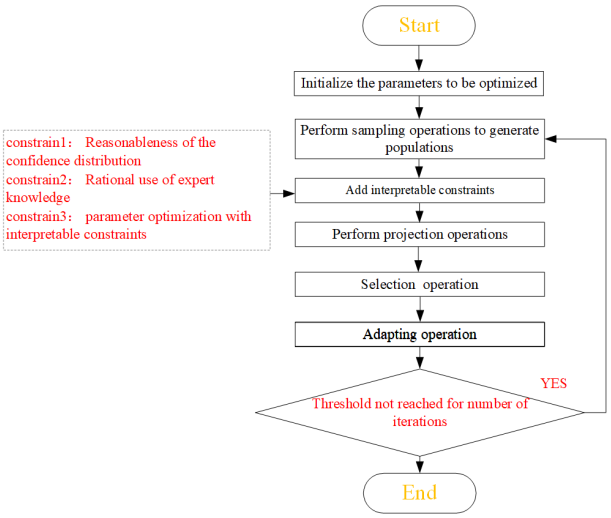


FIGURE 6. Improved P-CMA-ES algorithm optimisation process.

set by the experts is used to limit the search range of the population. The role of setting interpretability constraints in the IBRB-Di model is to balance the interpretability and accuracy of the model to prevent potential failures in complex systems. The flow of the P-CMA-ES optimization algorithm with interpretability constraints is shown in Figure 6.

To construct an optimization model, the function to be optimized needs to be specified first, and the model is expected to be more and more accurate during the optimization process. In addition, parameters such as the belief level of the rule, the weight of the rule, and the reliability of the rule should be controlled within the expert constraints. The difference between the predicted and actual values of the IBRB-Di model of the bearing fault diagnosis system is denoted by $Mse(*)$. Then, the objective function can be expressed as:

$$\begin{aligned} &Min\left\{Mse\left(\beta, w, r\right)\right\} \\ &s.t.(\beta, w, r)_{low} \leq (\beta, w, r)_{initial} \leq (\beta, w, r)_{up}, \\ &(\beta, w, r)_{low} \leq (\beta, w, r)_{optimal} \leq (\beta, w, r)_{up}, \\ &\sum_{l=1}^{N} \beta_l^k \leq 1, \beta_l^k \sim C_8 \left(l = 1, 2, \ldots, N\right) \end{aligned} \tag{17}$$

where C_8 represents interpretable guideline 8 in Figure 2. $\beta_l^k \sim C_8$ denotes that β_l^k is the belief distribution k_{th} that satisfies the interpretable criterion C_8. where $Mse(.)$ can be expressed as:

$$Mse\left(\beta, r, w\right) = \frac{1}{G}\sum_{u=1}^{G}\left(z_u - z^*{}_u\right)^2 \tag{18}$$

where G denotes the total number of training samples.z_u denotes the predicted value of the output of the IBRB-Di model, and $z^*{}_u$ denotes the true value of the whole system.

Therefore, the operational steps of the improved P-CMA-ES optimization algorithm are as follows:

Step 1: Initialize the parameters to be optimized:

$$\iota^0 = \phi^0\left(\beta, w, r\right) \tag{19}$$

where ι^0 denotes the set of initial parameters.

Step 2: Perform sampling operations to generate populations:

$$\phi_j^{s+1} = \iota^{(s)} + \upsilon^{(s)} N\left(0, B^{(s)}\right), j = 1, 2, \ldots, h \tag{20}$$

where ϕ_j^{s+1} is the j_{th} solution in the first $(s+1)_{th}$ generation optimization. $\upsilon^{(s)}$ represents the mutation strength of the s_{th} generation is also the global step size. $\iota^{(s)}$ represents the vector of search directions for the generation.$B^{(s)}$ is the covariance matrix.$N(.)$ is the normal distribution.h a represents the magnitude of the stock.

Step 3: Add interpretable constraints:

Constraint 1. Expert knowledge is an important part of interpretability, and the optimization process for local searches should be based on expert judgment. Therefore, expert knowledge is introduced in the initial population, and the introduction of the Euclidean distance for further local optimization can be described by the following:

$$\iota^{(s)} = \begin{cases} EK, ifs = 1 \\ \iota^{(s)}, ifs \neq 1 \end{cases} \tag{21}$$

where $\iota^{(s)}$ is the s_{th} generation population.EK denotes expert knowledge, and the local search is to be carried out within the scope of expert knowledge Eq.(22):

$$p\left(x_n, x'_n\right) = \sqrt{\sum_{i=1}^{n}\left(x_i - x'_i\right)^2} \leq d \tag{22}$$

where $p\left(x_n, x'_n\right)$ denotes the euclidean distance between the parameters obtained by the optimization algorithm through the population and the parameters set by the expert knowledge; d denotes the distance determined by the expert. When the value of d is set small, it means that the optimized parameters are closer to the expert knowledge. In this way, by the distance d set by the expert, limit the possibility that the parameters produced by the optimization algorithm deviate from the expert knowledge and reduce the effect of random vibrations.

Constraint 2: The initial and optimal parameters should be within the feasible region initially determined by the experts to ensure that the model parameters remain interpretable after optimisation, as in Eq. (23):

$$\begin{aligned} &(\beta, w, r)_{low} \leq (\beta, w, r)_{initial} \leq (\beta, w, r)_{up}, \\ &(\beta, w, r)_{low} \leq (\beta, w, r)_{optimal} \leq (\beta, w, r)_{up} \end{aligned} \tag{23}$$

where $(\beta, w, r)_{initial}$ denotes the initial parameter values set by the expert knowledge, while $(\beta, w, r)_{optimal}$ denotes the optimized expert knowledge parameter values. $(\beta, w, r)_{low}$ and $(\beta, w, r)_{up}$ denotes the interval of variation of the given parameter of the expert.

Constraint 3: A reasonable belief distribution can be monotonic or convex and cannot be concave:

$$\begin{aligned}&\beta_k \sim G_k(k=1,\ldots,S)\\&G_k \in \{\{\beta_1 \le \beta_2 \le \ldots \le \beta_N\},\\&\{\beta_1 \le \ldots \le \max(\beta_1,\beta_2,\ldots,) \ge \ldots \ge \beta_N\}\}\\&\{\beta_1 \ge \beta_2 \ge \ldots \ge \beta_N\}\end{aligned} \tag{24}$$

The interpretability constraint G_k is a constraint on the distribution of beliefs under rule k. There is no uniform fixed pattern for its form, but rather it is determined by experts' analysis of complex systems and long-term practical experience [36].

Step 4: Project the solution onto the hyperplane using the following:

$$\eta_e \phi_j^{(s+1)}(1+n_e(t-1):n_e t) = 1, t \in [1, N+1] \tag{25}$$

where N denotes the number of constraint variables.η_e is the set of parameters. n_e and t denote the number of variables in the equational constraints and equational constraints in $\phi_j^{(s+1)}$, respectively. The projection operation:

$$\begin{aligned}\phi_j^{(s+1)}(1+n_e(t-1):n_e t) &= \phi_j^{(s+1)}(1+n_e(t-1):n_e t)\\&- \eta_e^T\left(\eta_e \eta_e^T\right)^{-1} \times \phi_j^{(s+1)}(1+n_e(t-1):n_e t)\,\eta_e\end{aligned} \tag{26}$$

Step 5: Select the optimal solution, update the mean value, and calculate and sort the value of *Mse* according to the following Eq. (27):

$$Mse\left(\phi_{1:h}^{(s+1)}\right) \le Mse\left(\phi_{2:h}^{(s+1)}\right) \le \cdots \le Mse\left(\phi_{h:h}^{(s+1)}\right) \tag{27}$$

where $\left(\phi_{a:h}^{(s+1)}\right)$ is the a_{th} solution among the h solutions.

The subgroup density was calculated by the following Eq. (28):

$$\iota^{(s+1)} = \sum_{i=1}^{\rho} \varsigma_i \phi_{i:h}^{(s+1)} \tag{28}$$

where the weighting coefficients are denoted as ς_i and the size of the progeny population is $\rho.\phi_{i:h}^{(s+1)}$ is the i_{th} solution out of h solutions of the $(s+1)_{th}$ generation.

Step 6: Update the overall covariance matrix with:

$$\begin{aligned}B^{(s+1)} &= (1-c_1-c_2)B^{(s)} + c_1 p_c^{(s+1)}\left(p_c^{(s+1)}\right)^T\\&+ c_2 \sum_{i=1}^{\rho} \varsigma_i \left(\frac{\phi_{i:h}^{(s+1)} - m^{(s)}}{\upsilon^{(s)}}\right)\left(\frac{\phi_{i:\lambda}^{(s+1)} - \iota^{(s)}}{\upsilon^{(s)}}\right)^T\end{aligned} \tag{29}$$

$$\begin{aligned}p_c^{(s+1)} &= (1-c_c)p_c^{(s)} +\\&\sqrt{c_c(2-c_c)\left(\sum_{i=1}^{\rho}\varsigma_i^2\right)^{-1}}\left(\iota^{(s+1)} - \iota^{(s)}\right)\Big/\upsilon^{(s)}\end{aligned} \tag{30}$$

where the step size υ^{s+1} is updated by the following:

$$\upsilon^{s+1} = \upsilon^s \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\left\|p_\sigma^{(s+1)}\right\|}{E\left\|N(0,1)\right\|} - 1\right)\right) \tag{31}$$

$$\begin{aligned}p_\sigma^{(s+1)} &= (1-c_c)p_\sigma^{(s)} + \sqrt{c_c(2-c_c)\left(\sum_{i=1}^{\rho}\varsigma_i^2\right)^{-1}}\\&\left(B^{(s)}\right)^{-\frac{1}{2}}\left(\iota^{(s+1)} - \iota^{(s)}\right)\Big/\upsilon^{(s)}\end{aligned} \tag{32}$$

where c_1 and c_2 denote the learning rate and p_c denotes the evolutionary path.c_c is the evolutionary path parameter.d_σ is the damping parameter.

V. CASE STUDY

In this section, the validity of the established IBRB-Di fault diagnosis model is verified using the bearing fault diagnosis as an example.

Bearings are the most critical mechanical components in large rotating machines such as wind turbines, aircraft turbine engines, and high-speed trains [37]. Therefore, bearing failure is one of the main causes of rotating machinery failure [38]. Diagnosing bearing faults in an interpretable manner ensures that the diagnostic outcome is credible.

In addition, interpretable evaluation results are highly important, and the specification of interpretable criteria is important for the diagnosis of bearing failures. A clearer and more transparent understanding of bearing failures can be provided by users and maintenance personnel, making it more credible.

A. EXPERIMENTAL DATA SET

The monitoring data of rolling bearings under various working conditions were obtained on the simulation test bench of Electrical Engineering Laboratory of Case Western Reserve University. The rolling bearing type is SKF 6205 and the sampling frequency is 12 kHz. The running conditions of rollers, inner rings, and outer rings under normal operating conditions and three fault operating conditions (motor load = 735.49875, approximate speed = 1772 r/min) are simulated. The diameter of the fault point is 0.1778 mm. In general, the amount of experimental data is relatively large. However, the high sampling frequency results in short sampling intervals and insignificant differences between sampling points. Therefore, each sample needs to contain additional sampling points to ensure that the extracted features are valid. The sum of the three faulty and normal data collected is 1236 samples. To sum up, this experiment set 70% of them as training data and 30% as test data. The training set and test set for each fault level are shown in Table 1.

In this experiment, vibration signals of rolling bearings under different working conditions recorded in public data sets were analyzed. Identify and distinguish the operating state of bearings under normal and fault conditions, and then verify the effectiveness of vibration signal analysis for

TABLE 1. Experimental data distribution.

Category	Normal	Roller	Inner ring	Outer ring
Number	0	1	2	3
Train	170	259	261	237
Test	72	102	100	124

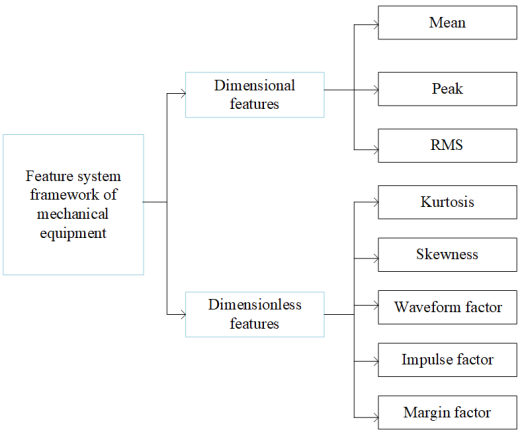


FIGURE 7. Feature system framework of mechanical equipment for vibration signals.

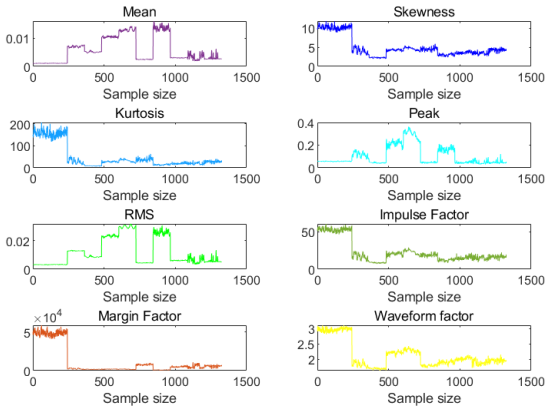


FIGURE 8. Bearing vibration signal dataset by attribute.

detecting bearing faults. We analyzed the vibration signals under different fault characteristics in the data set, and used a time-domain feature extraction method similar to that proposed in reference [39]. After time domain analysis of bearing vibration signals, the system framework of bearing vibration signal characteristics is shown in Figure 7. Bearing fault characteristics and fault results used in this experiment are shown in Figure 8 and Figure 9.

B. BEARING FAULT DIAGNOSIS USING THE IBRB-DI MODELS

The initial parameters of the model are set by the domain expert as shown in Table 2.

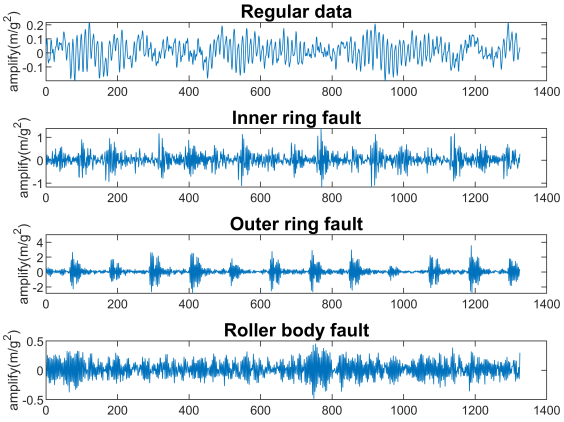


FIGURE 9. Fault status of bearing vibration.

The rule k_{th} representation of the IBRB-Di-based bearing fault diagnosis model is Eq.(33):

$$\text{Rule}_k: \text{IF} \begin{Bmatrix} \text{Mean} \in (c_1, d_1) \vee \text{Skewness} \in (c_2, d_2) \vee \\ \text{Kurtosis} \in (c_3, d_3) \vee \text{Peak} \in (c_4, d_4) \vee \\ \text{RMS} \in (c_5, d_5) \vee \\ \text{Impulse factor} \in (c_6, d_6) \vee \\ \text{Margin factor} \in (c_7, d_7) \vee \\ \text{Waveform factor} \in (c_8, d_8) \end{Bmatrix},$$

Then power fault detection results is

$\{(\theta_1, \beta_{1,k}), (\theta_2, \beta_{2,k}), (\theta_3, \beta_{3,k}), (\theta_4, \beta_{4,k})\}$,

with r_k and rule weight w_k,

rule matching degree δ_k,

in interpretable constraint $c_1, c_2, \ldots, c_f$,

$$k \in [1, S], \sum_{l=1}^{4} \beta_{l,k} \leq 1 \tag{33}$$

Description of the formula:

Where $\theta_1, \theta_2, \theta_3, \theta_4$ are the four evaluation levels of normal, inner ring fault, outer ring fault, and roller body fault, corresponding to reference values of 0, 1, 2, and 3, respectively.

As an example to validate the effectiveness of the model for bearing fault diagnosis, three evaluation metrics were used to validate it. The first one is the overall accuracy(ACR), which can be described:

$$ACR = \frac{\sum_{i=1}^{N} TP_i}{\sum_{i=1}^{N} (TP_i + FP_i + FN_i + TN_i)} \tag{34}$$

where TP_i is the number of true positives for category θ_i. FP_i is the number of false positives for category θ_i. FN_i is the number of false positives for category θ_i. TN_i is the number of true negatives for category θ_i. The second evaluation indicator is the positive predictive value (PPV):

$$PPV_i = \frac{TP_i}{TP_i + FP_i} \tag{35}$$

TABLE 2. The initial rules of the IBRB-Di.

DI	Referential interal	Rule reliability	The rule reliability constrait	Rule weight	The rule weight constrait	The initial belief
1	[0.001, 0.0027]	1	0.5-1	1	0.5-1	{0.27,0.07,0.38,0.28}
2	[0.0027,0.0043]	1	0.5-1	1	0.5-1	{0.01,0.00,0.02,0.97}
3	[0.0060,0.0077]	1	0.5-1	1	0.5-1	{0.71,0.02,0.16,0.11}
4	[0.0127,0.0144]	1	0.5-1	1	0.5-1	{0.16,0.51,0.09,0.24}
5	[0.0144,0.0160]	1	0.5-1	1	0.5-1	{0.02,0.46,0.22,0.30}
6	[1.99, 3.09]	1	0.5-1	1	0.5-1	{0.18,0.51,0.18,0.13}
7	[6.39,7.49]	1	0.5-1	1	0.5-1	{0.32,0.49,0.06,0.13}
8	[7.49,8.59]	1	0.5-1	1	0.5-1	{0.03,0.29,0.42,0.26}
...	...	...	...	...	...	...
72	[2.94,3.11]	1	0.5-1	1	0.5-1	{0.97,0.00,0.03,0.00}

TABLE 3. Optimized model parameter table.

ID	Rule reliability	Rules weights	Output results
1	0.507	0.658	{0.320, 0.040, 0.340, 0.30}
2	0.950	0.997	{0.010, 0.000, 0.00, 0.99}
3	0.950	0.985	{0.800, 0.110, 0.080, 0.01}
4	0.576	0.895	{0.100, 0.530, 0.170, 0.20}
5	0.883	0.741	{0.070, 0.450, 0.220, 0.26}
6	0.508	0.603	{0.190, 0.530, 0.180, 0.10}
7	0.779	0.798	{0.330, 0.450, 0.080, 0.14}
8	0.611	0.672	{0.100, 0.320, 0.400, 0.18}
...	...	...	...
72	0.950	0.921	{0.990, 0.000, 0.010, 0.00}

TABLE 4. Optimisation algorithm parameter settings.

The parameter	Value of initial parameter
Iteration number G	500
Population size h	24
Step size υ^0	0.2
European distance size d	2.4

is calculated separately for θ_i categories. where the PPVs for all categories are averaged to obtain them:

$$PPV_{avg} = \frac{1}{N}\sum_{i=1}^{N} PPV_i \tag{36}$$

The third is the True Positive Rate (TPR):

$$TPR_i = \frac{TP_i}{TP_i + FN_i} \tag{37}$$

is calculated separately for θ_i categories. where the TPRs for all categories are averaged to obtain the values:

$$TPR_{avg} = \frac{1}{N}\sum_{i=1}^{N} TPR_i \tag{38}$$

C. RESULTS OF THE EXPERIMENTS

The initial parameters of the model based on expert knowledge analysis are shown in Table 2. The constraints of rule reliability and rule weights are given by the experts. The parameters after optimization by the improved P-CMA-ES algorithm are shown in Table 3. The initial parameters in the optimization algorithm are set as in Table 4.

(1) Parameters of the optimized model:

The model is optimized using P-CMA-ES with constraints, constraining both rule weights and rule reliability to be optimized in the range of 0.5-1. The belief degree associated with the outcome in the belief distribution in each rule was also constrained and the optimization range fluctuated within the range of the initial values. The parameters after optimization are shown in Table 3. The observed results show that all the parameter variations remain within the constraints set by the expert, which ensures that the optimized parameters are well interpretable. This ensures that the optimized parameters of the IBRB-Di model are reasonable and feasible in real systems.

(2) Curve fitting results of predicted versus actual values of model diagnostic results:

The fitting results of the model predicted value and the real value are shown in Figure 10. The accuracy of curve

TABLE 5. Fault diagnosis performance of IBRB-Di.

Reuslt	θ_1	θ_2	θ_3	θ_4	TP+FP	PPV(%)
θ_1	72	0	0	0	72	100
θ_2	0	100	0	0	100	100
θ_3	0	2	100	7	109	91.74
θ_4	0	0	0	117	117	100
TP+FN	72	102	100	124	398	/
TPR(%)	100	98.04	100	94.35	/	/

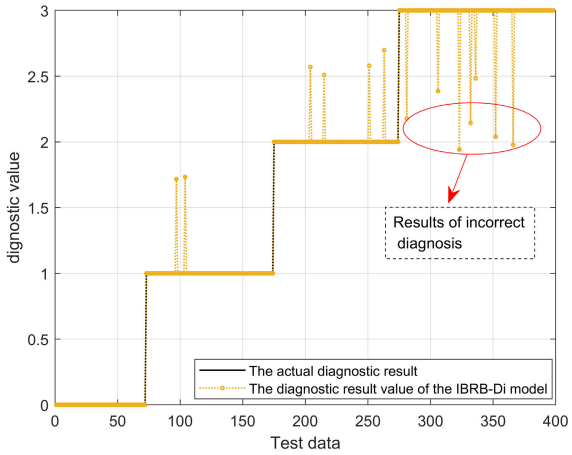


FIGURE 10. Results of the IBRB-Di diagnosis.

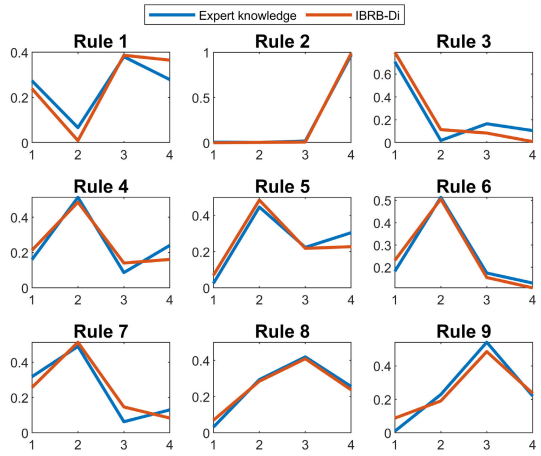


FIGURE 11. Output belief distribution of the IBRB-Di model.

TABLE 6. Comparison of different model rules and accuracy.

Method	Rule amount	ACR
IBRB-Di	72	97.74%
IBRB-D	72	98.49%
IBRB-r	160	85.18%
BRB	262144	/

fitting is 98.74%. This is due to the dynamic rule matching degree and the new rule weight calculation method used by the model, which enables it to better adapt to data fluctuations and non-linear trends. As can be seen from Figure 10, there is a good fit between the predicted value of the IBRB-Di model and the actual annotation of the real value of the system. From Table 5, we can see in detail how the model behaves under different fault types (normal, inner ring, outer ring, roller body). TP, FN and FP represent the total number of true positive, false negative and false positive results respectively. In particular, for θ_1 (normal), θ_2 (inner ring failure), and θ_4 (roller body failure), the PPV of the model reached 100%, indicating that it predicted these fault types almost perfectly. In the case of θ_3 (outer ring failure), although the PPV decreased slightly to 91.74%, The TPR remained above 94%, indicating the model's high sensitivity to failure. These data clearly demonstrate the ability of IBRB-Di model to diagnose bearing faults under various conditions, and confirm its reliability and effectiveness in industrial applications.

(3)Analysis of belief distributions for modeling rules:

As can be seen from Figure 11, with the P-CMA-ES optimisation algorithm with expert constraints, the IBRB-Di model maintains the interpretability of the parameters during the optimisation process. Thus, the belief distribution of the output results is highly consistent with the belief distribution initially set by the experts. It also shows that even after complex algorithm optimization to improve the model accuracy, the IBRB-Di model can still accurately reflect the expert's judgment and knowledge structure. The interpretability of the diagnostic results of the model is guaranteed, thus ensuring the transparency of the decision-making process and the reliability of the results.

D. COMPARATIVE EXPERIMENTS

In order to demonstrate the effectiveness of IBRB-Di in bearing fault diagnosis, it is analyzed in this subsection in a three-part comparative study.

In Part I, is a comparative experimental analysis with the traditional BRB, IBRB-r, the model IBRB-D without the addition of the improved P-CMA-ES to optimize the model, and the model IBRB-Di optimized by the improved P-CMA-ES.Next, we will compare the performance of the IBRB-Di model with other models in three key areas: rule complexity, model interpretability, and model precision.

(1) Firstly, the complexity of the model is analyzed, that is, the number of rules. According to the data in Table 6, BRB-D and IBRB-Di have the same number of rules, 72 rules for combining rules using interval addition. While the traditional BRB model has 262,144 rules using the Cartesian product combination rule, resulting in a huge number of rules. The main difference between the IBRB-r and IBRB-Di models is that they significantly reduce the number of rules and decrease the complexity of the model through the dynamic rule matching degree and the new method of calculating the rule weights.

(2) Secondly, from the analysis of model interpretability: The IBRB-Di model not only improves the accuracy and

TABLE 7. Comparing the accuracy of the IBRB-Di model to the original IBRB-r.

Method	θ_1		θ_2		θ_3		θ_4	
	TPR(%)	PPV(%)	TPR(%)	PPV(%)	TPR(%)	PPV(%)	TPR(%)	PPV(%)
IBRB-D	100	100	100	100	100	94.34	95.16	100
IBRB-Di	100	100	98.04	100	100	91.74	94.35	100
IBRB-r	91.67	98.51	86.27	87.13	80.00	72.73	84.68	87.50

TABLE 8. Comparison of the performances of various methods.

	Model	ACR (%)	TPR_{avg} (%)	PPV_{avg} (%)
Part I	IBRB-Di	97.74	98.09	97.94
	IBRB-D	98.49	98.79	98.58
	IBRB-r	85.18	85.66	86.46
Part II	WOA-IBRB	93.47	93.90	94.16
	DE-IBRB	91.21	91.88	92.03
	PSO-IBRB	92.71	92.42	93.06
Part III	BPNN	87.94	95.62	95.38
	RF	89.44	89.94	91.58
	ELM	79.45	80.99	80.81

FIGURE 12. The belief distributions of each rule.

adaptability of the model, but also ensures interpretability in the parameter tuning process by introducing the P-CMA-ES optimisation algorithm with expert constraints. Compared with the IBRB-D model, it can be seen from Figure 12 that the IBRB-Di model has the same trend in the belief distribution with the belief distribution of the initial expert knowledge setting. However the belief distribution results of the optimization algorithm for rules 5, 7, 8, and 9 of IBRB-D in Figure 12 defeat the original purpose of interpretability. A detailed comparison between IBRB-D and IBRB-Di in terms of rule weights and reliability is shown in Figure 13. IBRB-Di shows a higher reliability and a more stable distribution of weights for most of the rules, which reflects the IBRB-Di model's transparency and interpretability.interpretability of the IBRB-Di model in terms of rule application.

(3) Finally, In terms of model accuracy: From Table 7, it can be seen that the PPV on θ_3 and the TPR on θ_2 and θ_4 of IBRB-D are slightly higher than that of IBRB-Di. This is due to the fact that in the absence of an expert set optimisation range IBRB-D ignores interpretability in order to just go for accuracy. IBRB-Di, on the other hand, makes the accuracy too high while maintaining the interpretability of the model at the same time. The overall accuracy of IBRB-Di is higher than that of IBRB-r.

Since IBRB-Di can dynamically adjust the rules according to the fluctuation and non-linear trend of the actual data. This flexibility allows the model to maintain high accuracy when dealing with complex data without being affected by uneven data distribution. Finally, as can be seen from the results of the 20 repeated experiments in Figure 14(a), the accuracy of the IBRB-Di model is almost always maintained above 95%, showing its excellent robustness in complex fault situations. In contrast, the accuracies of the IBRB-D and IBRB-r models fluctuate considerably, especially the IBRB-r model performs poorly in the detection of all fault types.

In Part II, different optimization algorithms such as WOA-IBRB, DE-IBRB, and PSO-IBRB are compared with IBRB-Di. The different optimization algorithms use the same training samples and the same model. All the IBRB-Di algorithms have higher accuracy than the other algorithms, as shown in Table 8, Part II. The IBRB-Di optimization algorithm adds interpretability constraints, which makes the optimized parameters interpretable.

In Part III, IBRB-Di is compared with Backpropagation Neural Network (BPNN), Random Forest (RF), and Extreme

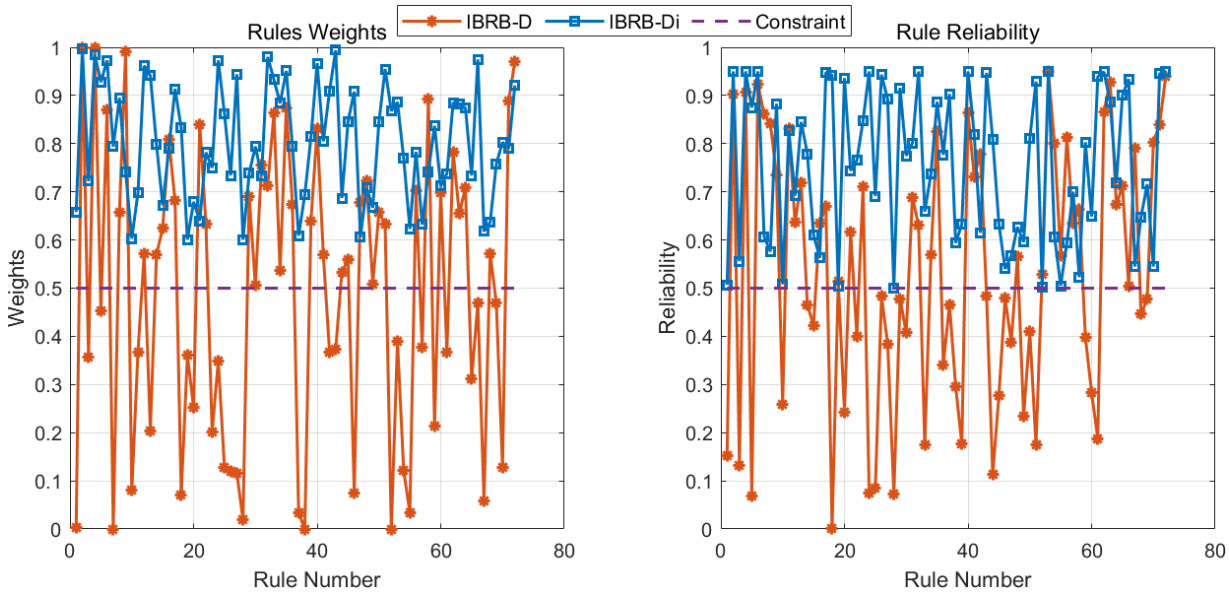


FIGURE 13. Comparison of the rules reliabilities and rules weights of IBRB-D and IBRB-Di.

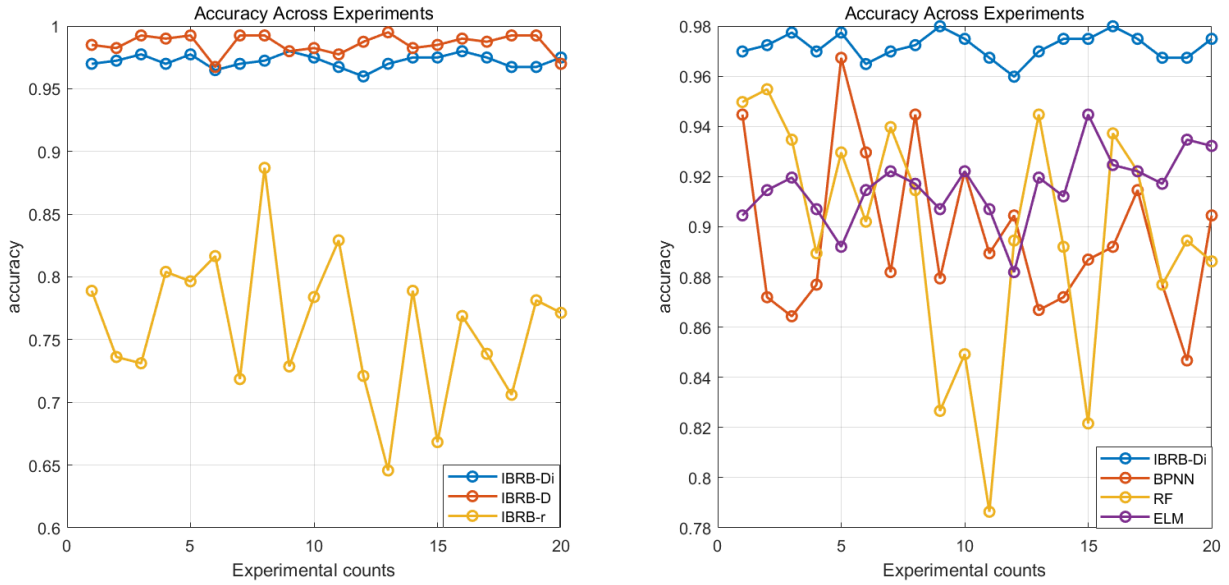


FIGURE 14. Comparison of the accuracy of different models.

Learning Machine (ELM). (1) In terms of accuracy: as shown in Part III of Table 8, the IBRB-Di model outperforms the other models in predicting bearing fault diagnosis. (2) In terms of model interpretability: all three models compared are black-box models and cannot be interpreted. (3) Most of the data-driven models require a large amount of data; however, IBRB-Di can maintain high accuracy even under small sample conditions. Figure 14 (b) shows that after 20 repetitions of the experiment, the accuracy of IBRB-Di is higher than other models.

E. EXPERIMENTAL SUMMARY

Through the above analysis, several significant advantages of IBRB-Di model in bearing fault diagnosis can be summarized:

(1) Improve accuracy and reduce the number of rules: As shown in Table 6, the IBRB-Di model optimizes the model structure by effectively reducing the number of rules. It realizes the simplification of rules by dynamically calculating the matching degree and weight method between samples and rules in each interval. This method not only

TABLE 9. Definition of numeric symbols.

Notation	Meaning
c	The interpretable criterion
ϕ_{best}	The parameter set after model optimization
ϕ_0	The initial parameters set by the expert
r	The set of parameters in the optimization process
$optimize(.)$	The optimization function
z	The output of the model
$h(.)$	The nonlinear functional relationship of the model
X	The input model fault characterization sample data
$Rule_k$	The k_{th} belief rule.
$x_1, \ldots, x_T$	There are a total of T fault features
(a_i, b_i)	The reference interval
$\theta_1, \theta_2, \ldots, \theta_N$	The rule k_{th} has N levels of evaluation that correspond to the results.
$\beta_{N,k}$	Belief distributions in rule k_{th}
r_k	Rule reliability in the k_{th}
w_k	Rule weights in the k_{th}
S	There is a total of S rules
δ_k	Activation rule matching in the k_{th}
med	The value of the midpoint of the reference interval
ℓ_k	Rules in the reasoning process
$\tilde{m}_{n,k}$	Weighted Confidence Distribution with Evidential Reliability
$b_{rw,k}$	Normalizing factor
$\hat{m}_{n,\ell(k)}$	The joint confidence level of the combination of k pieces of evidence for the evaluation level θ.
$m_{n,\ell(k)}$	The quality of the underlying probability that the first k pieces of evidence are fused into
Θ	A discursive framework that contains all possible propositions.
N	Number of types of health states
$MSE(*)$	The difference between the predicted and actual values of model
G	Optimize the number of iterations of the algorithm
ι^0	Initialize the parameters to be optimized
ϕ_j^{s+1}	The j_{th} solution in the first $(s+1)_{th}$ generation optimization.
EK	Expert knowledge
d	The distance determined by the expert
$B^{(s+1)}$	The overall covariance matrix
υ^{s+1}	The value of the step size for generation $s+1_{th}$
d_σ	The damping parameter

significantly reduces the complexity of the model, but also enhances the accuracy of the diagnosis.

(2) Enhanced model interpretability: IBRB-Di model effectively retains expert knowledge in the process of model fault diagnosis. Thus, the transparency and interpretability of the model are improved. In addition, it also improves the understanding and trust of decision-makers in the decision-making process of the model. Is a transparent and explainable decision support system. A balance has been struck between interpretability and accuracy. The accuracy of the model is improved, and the interpretability of the model is ensured.

VI. CONCLUSION

Based on IBRB-Di model, a new bearing fault diagnosis model is developed in this paper. The goal of the model is to enable decision makers to effectively predict bearing failures in an easy-to-understand manner.

Firstly, the IBRB-Di model introduces a new method to dynamically calculate the rule matching degree and rule weights of data samples within a reference interval. This method allows the model to dynamically adjust the matching degree and weight according to the actual data distribution, which maintains high accuracy and reliability even in the case of extremely heterogeneous data. Secondly, this method also effectively simplifies the number of rules and ensures the effectiveness of fault diagnosis. Finally, the optimisation algorithm using expert knowledge constraints to adjust the model parameters not only further improves the accuracy of the model, but also enhances the interpretability of the model.

At present, although the reference interval is a key parameter in the model, it is not directly involved in the optimization process, and its influence on the model interpretation has not been fully played. Future research needs to explore how to integrate the reference interval values into the optimization process of the model while maintaining their contribution to the interpretation of the results to further improve the adaptability and accuracy of the model.

APPENDIX

See Table 9.

REFERENCES

[1] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism Mach. Theory*, vol. 70, pp. 441–453, Dec. 2013.

[2] H.-B. Shi, D. Huang, L. Wang, M.-Y. Wu, Y.-C. Xu, B.-E. Zeng, and C. Pang, "An information integration approach to spacecraft fault diagnosis," *Enterprise Inf. Syst.*, vol. 15, no. 8, pp. 1128–1161, Sep. 2021.

[3] Z. Wei, Y. Wang, S. He, and J. Bao, "A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection," *Knowl.-Based Syst.*, vol. 116, pp. 1–12, Jan. 2017.

[4] Z. Wang, X. Gao, S. Huang, Q. Sun, Z. Chen, R. Tang, and Z. Di, "Measuring systemic risk contribution of global stock markets: A dynamic tail risk network approach," *Int. Rev. Financial Anal.*, vol. 84, Nov. 2022, Art. no. 102361.

[5] M. He and D. He, "Deep learning based approach for bearing fault diagnosis," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 3057–3065, May 2017.

[6] Y. Sun and S. Li, "Bearing fault diagnosis based on optimal convolution neural network," *Measurement*, vol. 190, Feb. 2022, Art. no. 110702.

[7] X. Jin, Z. Que, Y. Sun, Y. Guo, and W. Qiao, "A data-driven approach for bearing fault prognostics," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 3394–3401, Jul. 2019.

[8] J. Li, C. Hai, Z. Feng, and G. Li, "A transformer fault diagnosis method based on parameters optimization of hybrid kernel extreme learning machine," *IEEE Access*, vol. 9, pp. 126891–126902, 2021.

[9] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, "Feature extraction for data-driven remaining useful life prediction of rolling bearings," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[10] B. Cui, Y. Weng, and N. Zhang, "A feature extraction and machine learning framework for bearing fault diagnosis," *Renew. Energy*, vol. 191, pp. 987–997, May 2022.

[11] C. Liu and K. Gryllias, "Simulation-driven domain adaptation for rolling element bearing fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 5760–5770, Sep. 2022.

[12] L. Ma, B. Jiang, L. Xiao, and N. Lu, "Digital twin-assisted enhanced meta-transfer learning for rolling bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 200, Oct. 2023, Art. no. 110490.

[13] L. Hou, J. Zhao, S. Dun, Y. Cai, Y. Yang, J. Xu, and C. Sun, "Feature extraction of weak-bearing faults based on Laplace wavelet and orthogonal matching pursuit," *Shock Vibrat.*, vol. 2022, pp. 1–12, Mar. 2022.

[14] S. Zhang, B. Wang, M. Kanemaru, C. Lin, D. Liu, M. Miyoshi, K. H. Teo, and T. G. Habetler, "Model-based analysis and quantification of bearing faults in induction machines," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2158–2170, May 2020.

[15] V. T. Tran, H. Thom Pham, B.-S. Yang, and T. Tien Nguyen, "Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine," *Mech. Syst. Signal Process.*, vol. 32, pp. 320–330, Oct. 2012.

[16] M. Chen, N.-C. Xiao, M. J. Zuo, and Y. Ding, "An efficient algorithm for finding modules in fault trees," *IEEE Trans. Rel.*, vol. 70, no. 3, pp. 862–874, Sep. 2021.

[17] T. Yin, N. Lu, G. Guo, Y. Lei, S. Wang, and X. Guan, "Knowledge and data dual-driven transfer network for industrial robot fault diagnosis," *Mech. Syst. Signal Process.*, vol. 182, Jan. 2023, Art. no. 109597.

[18] Y. Li, S. Xie, J. Wang, J. Zhang, and H. Yan, "Sparse sample train axle bearing fault diagnosis: A semi-supervised model based on prior knowledge embedding," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

[19] S. Eke, G. Clerc, T. Aka-Ngnui, and I. Fofana, "Transformer condition assessment using fuzzy C-means clustering techniques," *IEEE Elect. Insul. Mag.*, vol. 35, no. 2, pp. 47–55, Mar. 2019.

[20] R. Soni and B. Mehta, "Diagnosis and prognosis of incipient faults and insulation status for asset management of power transformer using fuzzy logic controller & fuzzy clustering means," *Electr. Power Syst. Res.*, vol. 220, Jul. 2023, Art. no. 109256.

[21] X. Cheng, P. Han, W. He, and G. Zhou, "A new interval constructed belief rule base with rule reliability," *J. Supercomput.*, vol. 79, no. 14, pp. 15835–15867, Sep. 2023.

[22] J.-B. Yang, J. Liu, J. Wang, H.-S. Sii, and H.-W. Wang, "Belief rule-base inference methodology using the evidential reasoning approach-RIMER," *IEEE Trans. Syst. Man, Cybern. Part A, Syst. Humans*, vol. 36, no. 2, pp. 266–285, Mar. 2006.

[23] Y. Gao, X. Liu, and J. Xiang, "FEM simulation-based generative adversarial networks to detect bearing faults," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4961–4971, Jul. 2020.

[24] G.-Y. Hu, Z.-J. Zhou, B.-C. Zhang, X.-J. Yin, Z. Gao, and Z.-G. Zhou, "A method for predicting the network security situation based on hidden BRB model and revised CMA-ES algorithm," *Appl. Soft Comput.*, vol. 48, pp. 404–418, Nov. 2016.

[25] C.-C. Zhang, Z.-J. Zhou, S.-W. Tang, L.-Y. Chen, and P. Zhang, "BR-FRL: A belief rule-based fault recognition and location model for bus network systems," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[26] J. Wu, Q. Wang, Z. Wang, and Z. Zhou, "AutoBRB: An automated belief rule base model for pathologic complete response prediction in gastric cancer," *Comput. Biol. Med.*, vol. 140, Jan. 2022, Art. no. 105104.

[27] W. He, X. Cheng, X. Zhao, G. Zhou, H. Zhu, E. Zhao, and G. Qian, "An interval construction belief rule base with interpretability for complex systems," *Expert Syst. Appl.*, vol. 229, Nov. 2023, Art. no. 120485.

[28] Y. Zhang, G. Zhou, W. Zhang, W. He, Y. Wang, Y. Zhang, and P. Han, "A new performance analysis method for rolling bearing based on the evidential reasoning rule considering perturbation," *Sci. Rep.*, vol. 12, no. 1, p. 17842, Oct. 2022.

[29] B. Zhao, Q. Zhang, W. He, P. Han, Y. Cao, and G. Zhou, "A deep belief rule base-based fault diagnosis method for complex systems," *ISA Trans.*, vol. 150, pp. 77–91, Jul. 2024.

[30] Y. Cao, Z. Zhou, C. Hu, W. He, and S. Tang, "On the interpretability of belief rule-based expert systems," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 11, pp. 3489–3503, Nov. 2021.

[31] P. Han, W. He, Y. Cao, Y. Li, Q. Mu, and Y. Wang, "Lithium-ion battery health assessment method based on belief rule base with interpretability," *Appl. Soft Comput.*, vol. 138, May 2023, Art. no. 110160.

[32] A. Ramachandran, S. Gupta, S. Rana, C. Li, and S. Venkatesh, "Incorporating expert prior in Bayesian optimisation via space warping," *Knowl.-Based Syst.*, vol. 195, May 2020, Art. no. 105663.

[33] J. Wang, Z.-J. Zhou, C.-H. Hu, X.-X. Han, S.-W. Tang, and P.-Y. Ning, "An evidential reasoning rule considering parameter uncertainty," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 2, pp. 1391–1404, Apr. 2022.

[34] X. Yin, W. He, Y. Cao, N. Ma, G. Zhou, and H. Li, "A new health state assessment method based on interpretable belief rule base with bimetric balance," *Rel. Eng. Syst. Saf.*, vol. 242, Feb. 2024, Art. no. 109744.

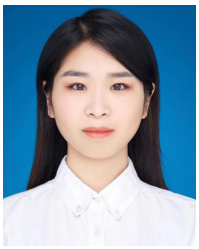
[35] Z. Ming, Z. Zhou, Y. Cao, S. Tang, Y. Chen, X. Han, and W. He, "A new interpretable fault diagnosis method based on belief rule base and probability table," *Chin. J. Aeronaut.*, vol. 36, no. 3, pp. 184–201, Mar. 2023.

[36] Q. Zhang, K. Li, G. Zhang, H. Zhu, and W. He, "A complex system health state assessment method with reference value optimization for interpretable BRB," *Sci. Rep.*, vol. 14, no. 1, p. 2334, Jan. 2024.

[37] H. Li, T. Liu, X. Wu, and Q. Chen, "An optimized VMD method and its applications in bearing fault diagnosis," *Measurement*, vol. 166, Dec. 2020, Art. no. 108185.

[38] G. Yu, T. Lin, Z. Wang, and Y. Li, "Time-reassigned multisynchrosqueezing transform for bearing fault diagnosis of rotating machinery," *IEEE Trans. Ind. Electron.*, vol. 68, no. 2, pp. 1486–1496, Feb. 2021.

[39] M. Chen, Z. Zhou, B. Zhang, G. Hu, and Y. Cao, "A novel combination belief rule base model for mechanical equipment fault diagnosis," *Chin. J. Aeronaut.*, vol. 35, no. 5, pp. 158–178, May 2022.



HAIFENG WAN received the B.Eng. degree from Heilongjiang Institute of Technology, China, in 2022. She is currently pursuing the master's degree with Harbin Normal University, Harbin, China. Her research interests include evidential reasoning, safety assessment, fault prognosis, and belief rule base.

MINGYUAN LIU received the B.Eng. degree from Shijiazhuang University, Shijiazhuang, China, in 2022. He is currently pursuing the master's degree with Harbin Normal University, Harbin, China. His research interests include evidential reasoning, deep learning, safety assessment, fault prognosis, and belief rule base.

HAILONG ZHU received the B.Eng. degree from Harbin Engineering University, Harbin, China, in 2001, the M.Eng. degree from Harbin University of Technology, Harbin, in 2008, and the Ph.D. degree from Harbin Institute of Technology, Harbin, in 2012. He is currently a Professor with the School of Computer Science and Information Engineering, Harbin Normal University, Harbin. He has published more than ten articles and monographs. His research interests include pattern recognition and digital image processing.

NING MA received the Ph.D. degree from Harbin Engineering University, in 2018. He is currently with the School of Computer Science and Information Engineering, Harbin Normal University. His research interests include evidential reasoning, belief rule base, and intelligent optimization.

WEI HE received the Ph.D. degree from Harbin University of Science and Technology, in 2018. He is currently with the School of Computer Science and Information Engineering, Harbin Normal University. His research interests include evidential reasoning, deep learning, and belief rule base.