



OPEN

A novel optimization method for belief rule base expert system with activation rate

Gang Xiang^{1,2}, Jie Wang³, XiaoXia Han⁴, Shuaiwen Tang⁴ & Guanyu Hu^{3,4}✉

Although the belief rule base (BRB) expert system has many advantages, such as the effective use of semi-quantitative information, objective description of uncertainty, and efficient nonlinear modeling capability, it is always limited by the problem of combinatorial explosion. The main reason is that the optimization of a BRB with many rules will consume many computing resources, which makes it unable to meet the real-time requirements in some complex systems. Another reason is that the optimization process will destroy the interpretability of those parameters that belong to the inadequately activated rules given by experts. To solve these problems, a novel optimization method for BRB is proposed in this paper. Through the activation rate, the rules that have never been activated or inadequately activated are pruned during the optimization process. Furthermore, even if there is a complete data set and all rules are activated, the activation rate can also be used in the parallel optimization process of the BRB expert system, where the training data set is divided into some subprocesses. The proposed method effectively solves the combinatorial explosion problem of BRB and can make full use of quantitative data without destroying the original interpretability provided by experts. Case studies prove the advantages and effectiveness of the proposed method, which greatly expands the application fields of the BRB expert system.

Expert systems are one of the most traditional artificial intelligence methods and have been used in many fields, including finance, industry, medicine, and education¹. It can express extensive knowledge and experience of a complex system and obtain the final results by the inference engine. However, in the era of big data, an expert system cannot effectively utilize multisource data from complex environments and internal systems, which limits its applications. Data-driven approaches can make up for this defect, such as neural networks^{2,3}, dynamic Bayesian networks⁴, and deep learning methods⁵. However, they cannot use expert experience and domain knowledge to guide the setting of initial parameters, which brings much uncertainty and pressure to the model optimization. A semiquantitative model can combine the advantages of the above two types of models, such as the hidden Markov model (HMM)⁶ and fuzzy neural network⁷. Although the above methods have been applied in many cases, they all lack the ability to process various types of uncertainty, including randomness, fuzziness and ignorance, and lack the interpretability and credibility of the results. As an intelligent expert system and interpretable artificial intelligence method, the belief rule base (BRB) expert system^{8–10} can effectively utilize semiquantitative information, including qualitative knowledge and quantitative data, and objectively express uncertain information. Currently, increasing attention is being paid to BRBs, and many variants have been developed^{11–15}.

Although BRB has many advantages, as an expert system, it will also face the problem of combinatorial explosion when the number of attributes is increased. The main reason is that the number of rules in the BRB will increase exponentially with the increase in attributes and reference levels. There are many problems to be solved for BRBs, and the most important problem is model optimization. Because of the increasing parameters in the BRB, the optimization process will expend considerable time and computing resources. For example, there are 8 attributes in distinguishing diabetes. Assuming that each attribute has 3 reference levels, the number of rules in BRB is $3^8 = 6561$. In the following sections, we will see that each rule of BRB has 12 parameters, including rule weights, attribute weights, and belief degrees of consequents. Thus, the number of parameters of the BRB is $6561 \times 12 = 78,732$, which means that searching for optimal parameters will run in a very high-dimensional solution space. In addition, the objective function of BRB optimization is a nonconvex, highly nonlinear, and existing equality constraint problem. Therefore, the optimization process of BRBs with a large number of parameters

¹Beihang University School of Automation Science and Electrical Engineering, Beijing 100191, China. ²Beijing Aerospace Automatic Control Institute, Beijing 100854, China. ³Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China. ⁴High-Tech Institute of Xi'an, Xi'an 710025, China. ✉email: huguanyu0708@163.com

is very difficult to solve. On the other hand, it is also unreasonable to optimize all the parameters of the BRB because some rules may not be activated or only activated a few times by incomplete quantitative data. It is not only meaningless to optimize those rules but will also destroy the original interpretability provided by experts.

Based on the above descriptions, it is necessary to develop an effective method to solve the optimization problem of BRBs. Currently, there are 2 types of methods for this problem: (1) Dimension reduction methods, which can reduce the number of attributes, such as principal component analysis (PCA)¹⁶ and linear discriminant analysis (LDA)¹⁷. For example, Hu used PCA to reduce the characteristics of attacks in network security situation prediction by using BRB¹⁸. (2) Structure reduction methods, which can reduce the number of rules. For example, Zhou proposed an automatic adding and deleting criterion for belief rules in BRBs based on statistical utility¹⁹. Chang proposed a structure learning method for BRBs based on gray targets (GTs) and multidimensional scaling (MDS)¹⁶.

Although the above methods can relieve the pressure of BRB optimization to a certain extent, some shortcomings still exist. Dimension reduction methods cannot keep the original meaning of the attribute, which weakens the advantage that the BRB expert system can effectively utilize domain knowledge. Structure reduction methods are not always efficient and reach practical optimization speeds at the expense of precision. Obviously, from the view of scale reduction, the optimization problem of BRBs with a large number of parameters cannot be effectively solved. Therefore, a novel optimization method for BRBs with activation rates is proposed in this paper, where the activation rate is used to determine which rules should be optimized in a process, and then the whole optimization process of BRBs can be simplified without losing accuracy. Furthermore, in the situation that most of the rules are activated, the activation rate can also be used in the parallel process of BRB optimization, where the training data set is divided into some subprocesses. The proposed method in this paper can reduce the unnecessary optimization of those unactivated rules of the BRB expert system, which ensures that the quantitative data can be utilized as much as possible without destroying the original interpretability provided by experts.

The remainder of this paper is organized as follows. In “The basic description of the BRB expert system” section, the basic description of the BRB expert system is introduced, and the optimization problem of BRBs is analysed. In “Optimization method for BRBs” section, the novel optimization method with activation rate is proposed, and the parallelized process is constructed. In “Case studies” section, two case studies are designed to verify the effectiveness of the proposed method. The paper is concluded in “Conclusion” section.

The basic description of the BRB expert system

The structure of the BRB expert system. BRB is an intelligence expert system that can effectively use qualitative knowledge and quantitative data and can express most uncertainty information. The basic construction of the BRB expert system is as follows⁸.

$$R_k : \text{If } (a_1 \text{ is } A_1^k) \wedge (a_2 \text{ is } A_2^k) \wedge \dots \wedge (a_M \text{ is } A_M^k) \\ \text{Then } \{(D_1, \beta_{1,k}), (D_2, \beta_{2,k}), \dots, (D_N, \beta_{N,k}), (D, \beta_{D,k})\} \\ \text{with rule weight } \theta_k \ (k = 1, 2, \dots, L) \text{ and attribute weight } \delta_i \ (i = 1, 2, \dots, M) \quad (1)$$

where R_k denotes the k th rule in the belief rule base, and a_i ($i = 1, \dots, M$) denotes the i th antecedent attribute, whose referential value is A_i^k , $A_i^k = (A_{i,1}^k, A_{i,2}^k, \dots, A_{i,J_i}^k)$, where J_i denotes the number of referential levels of the i th attribute. M denotes the number of antecedent attributes. To simplify the problem, we assume that the number of attributes in each rule is the same. D_j ($j = 1, \dots, N$) denotes the j th output result, whose belief degree can be expressed by $\beta_{j,k}$. D is the set that includes all the results, so the belief degree $\beta_{D,k}$ assigned to D denotes the remaining belief degree. Because β denotes the probabilities of the results, the sum of those belief degrees must equal 1, which is the constraint condition in training BRB.

The BRB expert system uses the above belief rules to construct the nonlinear relationship between the input and output of a complex system, and as a general probability, the belief degree can express various types of uncertain information in an objective world.

The reasoning process of the BRB expert system. When data are imported into BRB, some rules are activated. The principle of the activation is that when attributes of the data match the corresponding reference levels, the transformation method is used to generate a matching degree of the attribute value relative to the reference value. The transformation method depends on the form of the attributes. If the attributes are quantitative, the matching degrees can be obtained by the equivalence transformation technique^{20,21}. If the attributes are qualitative, the matching degrees can be obtained by the subjective judgment of experts²². All the matching degrees will construct M matching degree vectors, denoted by v_i , where each vector includes J_i matching degrees. Then a matching degree matrix V can be obtained. V includes an element $V_{k,i}$ ($k = 1, \dots, L$; $i = 1, \dots, M$) that is selected from v_i according to the rules arranged by reference level. Thus the activation weight of the k th rule ω_k can be calculated by

$$\omega_k = \frac{\theta_k \prod_{i=1}^M (V_{k,i})^{\bar{\delta}_i}}{\sum_{k=1}^L \theta_k \prod_{i=1}^M (V_{k,i})^{\bar{\delta}_i}}, \text{ where } \bar{\delta}_i = \frac{\delta_i}{\max\{\delta_i\}} \quad (2)$$

If the activation weight is not equal to 0, the corresponding rule is activated. Then, the following evidential reasoning (ER) rule is utilized to fuse the activated rules and finally obtain the distribution of belief degrees $\hat{\beta}_j$ assigned to the results, as shown in Eq.

$$\hat{\beta}_j = \frac{\gamma \times \left[\prod_{k=1}^L \left(\omega_k \beta_{j,k} + 1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) - \prod_{k=1}^L \left(1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) \right]}{1 - \gamma \times \left[\prod_{k=1}^L (1 - \omega_k) \right]} \quad (j = 1, \dots, N) \quad (3)$$

$$\gamma = \left[\sum_{j=1}^N \prod_{k=1}^L \left(\omega_k \beta_{j,k} + 1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) - (N-1) \prod_{k=1}^L \left(1 - \omega_k \sum_{i=1}^N \beta_{i,k} \right) \right]^{-1} \quad (4)$$

Optimization of the BRB expert system. It can be seen from the above descriptions that BRB includes many parameters, of which the initial values are usually given by experts. These initial parameters constitute a rough BRB, which cannot produce accurate results. Therefore, parameter optimization for BRBs is necessary. The first step is to establish an optimization objective, shown as follows.

$$\begin{aligned} & \min \{F(\Omega)\} \\ \text{s.t. } & 0 \leq \theta_k \leq 1, \quad k = 1, \dots, L \\ & 0 \leq \delta_i \leq 1, \quad i = 1, \dots, M \\ & 0 \leq \beta_{j,k} \leq 1, \quad j = 1, \dots, N, k = 1, \dots, L \\ & \beta_{D,k} + \sum_{j=1}^N \beta_{j,k} = 1 \end{aligned} \quad (5)$$

where $F(\Omega)$ denotes the objective function, which can be defined through the mean square deviation between the real values and testing results of the BRB. Ω denotes the parameters to be optimized.

Equation (3) is a highly nonlinear, highly dimensional, strongly constrained optimization problem. Therefore, the second step is to select an appropriate optimization algorithm to solve the optimization objective of the BRB²³ used the sequential quadratic programming (SQP) algorithm to obtain the optimal parameters of BRBs²⁴ proposed the projection covariance matrix adaptation evolution strategy (P-CMA-ES) algorithm, which achieves a good optimization effect.

Although the accuracy of BBR can be improved by the parameter optimization process, when the attributes and reference levels are increased, BRB has to face the problem of combinatorial explosion. As described in “Introduction” section, a large number of parameters will not only reduce the speed of optimization, which will lead to failure in the scene with high real-time requirements but also lead to a decrease in accuracy because of the difficulty of optimal solution search in high-dimensional space. The above limitation greatly restricts the application of BRBs in more complex and wide fields. Therefore, an efficient optimization method for BRBs is proposed in this paper.

Optimization method for BRBs

A novel optimization method for BRBs with activation rates is proposed in this section. The method can be used in two different application scenarios: (1) BRBs with incomplete samples or patterns, which means that a part of the rules will never be activated or only activated a few times. By pruning these rules by using the activation rate and threshold, the optimization dimension will be greatly reduced. (2) Fully activated BRB, where all rules are activated many times. Through parallel operation by using the activation rate and threshold, optimization will be separated into many child processes, which will fundamentally solve the problem of BRB optimization. Next, the basic principles of the proposed optimization method will be introduced in two different scenarios.

BRB optimization method with activation rate. After an in-depth analysis of the combination explosion problem, we found that when samples are incomplete or the actual system does not cover all patterns, a part of the rules will never be activated or only activated very few times in the whole training process, which is called inadequate-activated rules. However, in the traditional optimization process of BRBs, the parameters of all rules are involved in optimization, which is unreasonable. The initial parameters of the BRB are given by experts based on experience and domain knowledge. If the parameters of these nonactivated or inadequate-activated rules are optimized, then we give up expert knowledge without enough quantitative data to provide information for model learning. To solve the above problem, the activation rate for the BRB is first proposed, as follows.

$$ar_k = \frac{an_k}{\sum_{n=1}^L an_n} \quad (6)$$

where ar_k denotes the activation rate of the k th rule and an_k denotes the number of activation times of the k th rule.

To prune the nonactivated and inadequate-activated rules, a threshold h of the activation rate must be given. When the activation rate is greater than the threshold h , the parameters of the corresponding rules can be optimized, and the remaining rules still keep the initial values given by experts.

Remark 1 Note that an_k can be obtained only after all samples are input into the initial BRB, which cannot affect the efficiency of optimization because the initial BRB is without a training process and can quickly obtain output results.

Parallel optimization method of BRBs. The scale of the BRB can only be reduced to a certain extent through the activation rate, but when the data set is relatively complete, the scale reduction will be limited, which cannot solve the optimization problem of BRBs with a large number of parameters in essence. With the development of computer technology, parallel optimization provides a good solution, which can greatly reduce the optimization time. Theoretically, if we have enough computing units, the optimization time will surely meet the requirements of the actual system. Thus, a parallel optimization method for BRBs is proposed in this section.

Inspired by the activation rate and pruning rules, the parallelization of BRB optimization can be achieved by partitioning the data set, which can be denoted as

$$S = \begin{bmatrix} s_{1,1}, & s_{1,2}, & \dots, & s_{1,M} \\ s_{2,1}, & s_{2,2}, & \dots, & s_{2,M} \\ \dots & \dots & \dots & \dots \\ s_{sn,1}, & s_{sn,2}, & \dots, & s_{sn,M} \end{bmatrix} \tag{7}$$

where S denotes the original data set, sn denotes the number of samples, and M denotes the number of attributes of each sample $s_{i,1}, s_{i,2}, \dots, s_{i,M}$.

The parallelization steps of BRB optimization are shown as follows:

Step 1 First, the initial parameter values of the BRB are set according to expert knowledge.

Step 2 Assuming that the number of optimization subprocesses is pn , the training data set can be divided into pn parts, each of which is an average sampled from sn samples of the original data set.

Step 3 Input every sub data set into the initial BRB, and calculate the activation rate $AR_n = (ar_1^n, ar_2^n, \dots, ar_k^n)$; ($n = 1, 2, \dots, pn$) of each sub data set, where AR_i denotes the activation rate set of the i th sub data set. ar_k^n denotes the activation rate of the k th rule activated by the n th sub data set.

Step 4 Set the threshold for the activation rate, which can decide which rules to participate in each optimization subprocess. Then, the BRB can be divided into pn sub-BRB models, denoted as BRB_n .

Step 5 The corresponding sub-BRB is assigned to different computing units and optimized independently according to the corresponding training sub data set. The optimization algorithm is P-CMA-ES, which is used to minimize the objective function shown in Eq. (3). Please refer to algorithm 1 for pseudo code of P-CMA-ES algorithm.

Step 6 After the above steps, we obtain pn groups of belief degree distributions, and each group has sn belief degree distributions for the output results of sn samples in the BRB. The belief degree distribution generated by the n th optimization subprocess of the i th testing sample can be denoted as $B_{n,i} = (\hat{\beta}_{n,i}^1, \hat{\beta}_{n,i}^2, \dots, \hat{\beta}_{n,i}^N)$. To obtain the final belief degree distribution, the weighted average method is utilized, and the weight of the n th distribution can be determined by $p\omega_n$

$$p\omega_n = \frac{\sum_{k=1}^L an_k^n}{\sum_{n=1}^{pn} \sum_{k=1}^L an_k^n} \times (pn - 1) \tag{8}$$

where an_k^n denotes the number of activation times of the k th rule in BRB_n . Then, the final belief degrees distribution of i th sample B_f^i can be obtained by Eq. (8), the final results of the BRB can be obtained by Eq. (9).

$$B_f^i = \left(\frac{\sum_{n=1}^{pn} p\omega_n \times \hat{\beta}_{n,i}^1}{pn}, \frac{\sum_{n=1}^{pn} p\omega_n \times \hat{\beta}_{n,i}^2}{pn}, \dots, \frac{\sum_{n=1}^{pn} p\omega_n \times \hat{\beta}_{n,i}^N}{pn} \right) \tag{9}$$

$$Z_i = \sum_{j=1}^N B_{f,j}^i \times D_j; \quad (i = 1, \dots, sn) \tag{10}$$

Remark 2 The above optimization subprocesses are independent of each other. The weighted average operation for the final belief degree distribution is executed only when the optimization is completed.

Algorithm 1 P-CMA-ES: Projection Covariance Matrix Adaptation Evolution Strategy

Require: $c_\sigma, c_c, c_1, c_\mu$	Learning rates
Require: d_σ	Damping factor
Require: $\mathbf{g} = \mathbf{0}$	Generation counter
Require: $m^{(0)} \in R^N, \sigma^{(0)} \in E_\lambda$	Inputs: initial mean vectors and step size
Require: $C^{(0)} = I, P_\sigma^{(0)} = 0, P_c^{(0)} = 0$	Initialize covariance matrix and evolution paths.

repeat

$$\Omega_i^{g+1} \leftarrow \omega^g + \varepsilon^g N(0, C^g) \quad i = 1, \dots, \lambda$$

$$\Omega_i^{g+1} (1 + n_c \times (j-1) : n_c \times j) \leftarrow \frac{\Omega_i^{g+1} (1 + n_c \times (j-1) : n_c \times j) \times A_j}{\Omega_i^{g+1} (1 + n_c \times (j-1) : n_c \times j) - A_j^T \times (A_j \times A_j^T)}$$

$$MSE(\Omega_{i\lambda}^{g+1}) \leq MSE(\Omega_{i\lambda-1}^{g+1}) \leq \dots \leq MSE(\Omega_{i2}^{g+1}) \leq \dots \leq MSE(\Omega_{i1}^{g+1})$$

$$\omega^{g+1} \leftarrow \sum_{i=1}^{\tau} h_i \Omega_{i\lambda}^{g+1}$$

$$C^{g+1} \leftarrow (1 - c_1 - c_2) C^g + c_1 p_c^{g+1} (p_c^{g+1})^T + c_2 \sum_{i=1}^{\tau} h_i \left(\frac{\Omega_{i\lambda}^{g+1} - \omega^g}{\varepsilon^g} \right) \left(\frac{\Omega_{i\lambda}^{g+1} - \omega^g}{\varepsilon^g} \right)^T$$

$$p_c^{g+1} \leftarrow (1 - c_c) p_c^g + \sqrt{c_c (2 - c_c)} \left(\sum_{i=1}^{\tau} h_i^2 \right)^{-1} \frac{\omega^{g+1} - \omega^g}{\varepsilon^g}$$

$$\varepsilon^{g+1} \leftarrow \varepsilon^g \exp \left(\frac{c_\sigma}{d_\sigma} \left(\frac{\|p_\sigma^{g+1}\|}{E\|N(0, I)\|} - 1 \right) \right)$$

$$p_\sigma^{g+1} \leftarrow (1 - c_\sigma) p_\sigma^g + \sqrt{c_\sigma (2 - c_\sigma)} \left(\sum_{i=1}^{\tau} h_i^2 \right)^{-1} C^{(k)\frac{1}{2}} \frac{\omega^{g+1} - \omega^g}{\varepsilon^g}$$

until hit stopping criteria

return $\mu^{(l)}, \sigma^{(l)}, C^{(l)}$

Case studies

To verify the superiority of the proposed method, two cases, “Health status assessment of laser gyro” and “Leak size estimation of oil pipeline”, were used for verification.

Health status assessment of laser gyro. *Problem formulation.* A laser gyro is a precision instrument in the navigation control system. Its state parameters are zero-order drift coefficient, first-order drift coefficient, X-axis gyroscope light intensity voltage. When these parameters exceed the calibration threshold, it means that the laser gyro has failed, and the navigation control system will fail at this time. However, when these parameters are within the threshold, the laser gyro will also show different states. At this time, evaluating their health is also a necessary means to measure whether the laser gyro meets the navigation accuracy. Therefore, this case studies the health assessment of laser gyro, using the following data sets.

In this case, the data set of the laser gyro is used to prove the advantages of the proposed method. This data set contains a zero-order drift coefficient, first-order drift coefficient, X-axis gyroscope light intensity voltage, and expected utility value. The data set has 2000 samples, as shown in Figs. 1, 2, 3 and 4.

First, we can establish a BRB expert system according to expert experience or domain knowledge. The reference values of the zero-term drift coefficient, first-term drift coefficient, and X-axis gyroscope light intensity voltage are shown in Tables 1, 2 and 3. Thus the BRB expert system of laser gyroscope health status detection can be described

$$R_k : IF \left(\text{zero-term drift coefficient is } A_1^k \right) \wedge \left(\text{first-term drift coefficient is } A_2^k \right) \wedge \left(X\text{-axis is } A_3^k \right) \\ \text{Then laser gyroscope health status is } \{ (H, \beta_1^k), (SH, \beta_2^k), (UH, \beta_3^k) \} \\ \text{with rule weight } \theta_k \text{ and attribute weight } \delta_i \tag{11}$$

where H, SH, UH denote the reference values of the laser gyroscope health status, as shown in Table 4. The other parameters of the BRB expert system are shown in Table 5.

As described in the above sections, the data set may not cover all patterns about the zero-order drift coefficient, first-order drift coefficient and X-axis gyroscope light intensity voltage, which means that perhaps only a part of the rules shown in Table 6 can be activated. To prove this point, the laser gyro data samples are input into the above BRB expert system. The samples are divided into a training data set and a testing data set. The training data set includes 500 samples, which are regularly selected from the 2000 samples, and the remaining

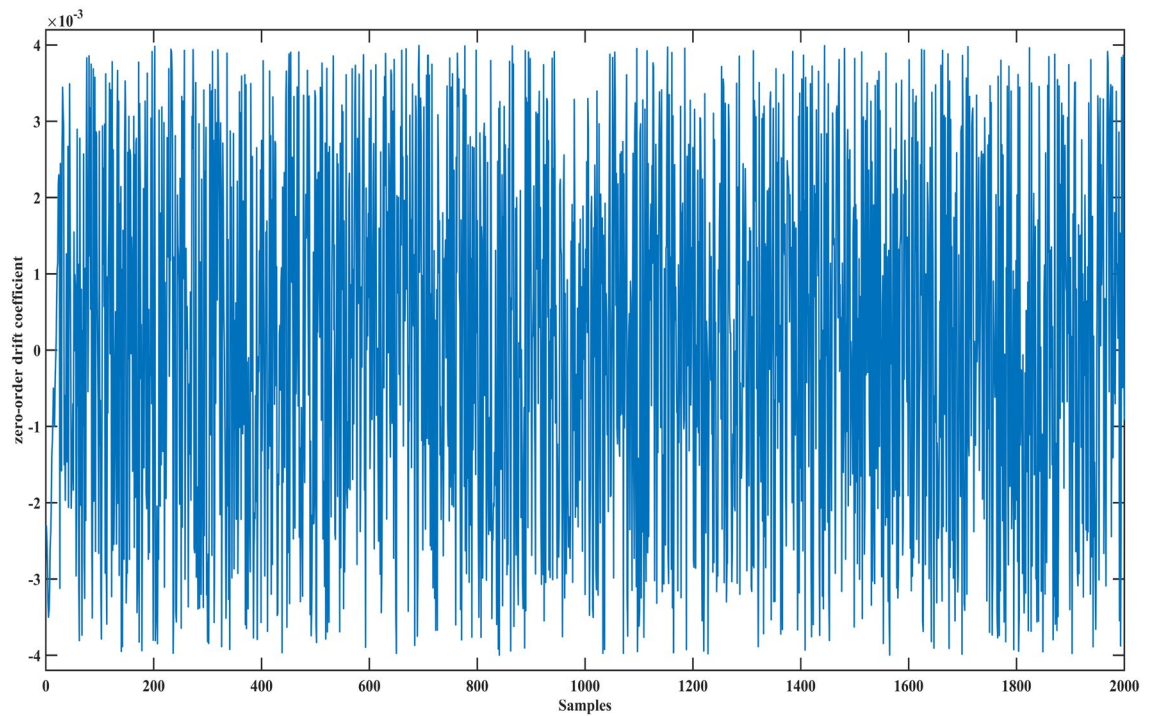


Figure 1. The zero-order drift coefficient of the laser gyro.

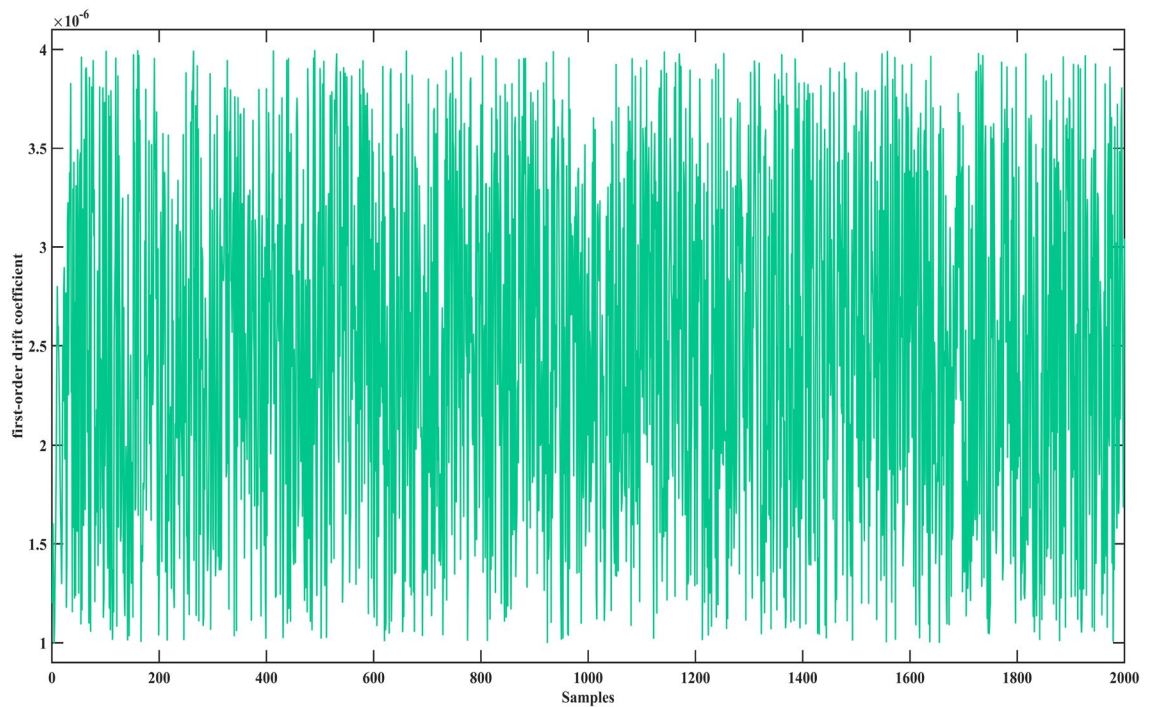


Figure 2. The first-order drift coefficient of the laser gyro.

1500 samples are used as a testing data set. The P-CMA-ES algorithm is utilized as an optimization tool of the BRB expert system, and to ensure the fairness of the optimization process, the parameters and iterations are the same for different conditions.

Case 1-optimization process of a BRB using activation rates. Figure 5 shows the activation rates of 64 rules in the BRB expert system when the training data set is entered. It can be seen that the activation rates of different rules are not the same, and the activation rates of some rules are even very small and will hardly be activated.

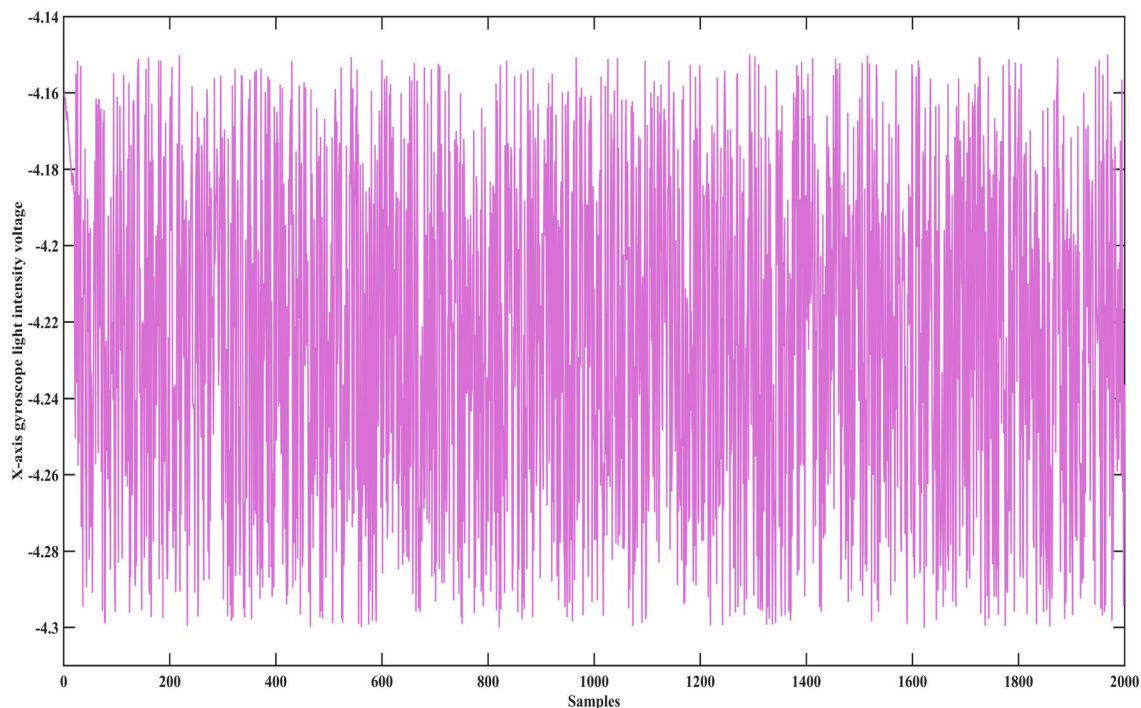


Figure 3. The X-axis gyroscope light intensity voltage of the laser gyro.

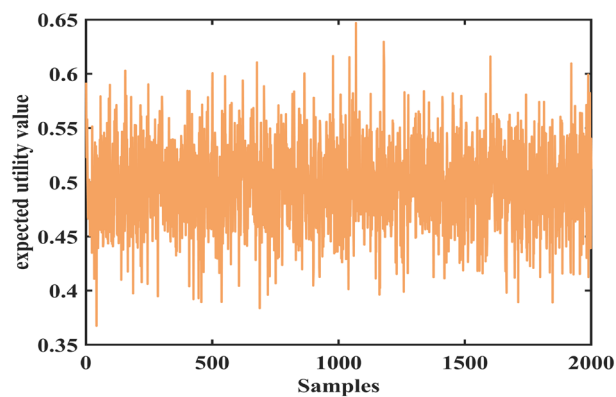


Figure 4. The expected utility value of the laser gyro.

Linguistic terms	<i>NL</i>	<i>NS</i>	<i>PS</i>	<i>PL</i>
Values	-4×10^{-3}	-1.3×10^{-3}	1.3×10^{-3}	4×10^{-3}

Table 1. The reference values and points of the zero-order term drift coefficient.

Linguistic terms	<i>PS</i>	<i>PM</i>	<i>PL</i>	<i>PVL</i>
Values	1×10^{-6}	2×10^{-6}	3×10^{-6}	4×10^{-6}

Table 2. The reference values and points of the primary term drift coefficient.

Linguistic terms	NL	NM	NS	NVS
Values	-4.3	-4.25	-4.2	-4.15

Table 3. The reference values and points of the X-axis gyroscope light intensity voltage.

Linguistic terms	H	SH	UH
Values	0	2	4

Table 4. The reference values and points of laser gyroscope health status.

Four BRBs with different activation rate thresholds are established for comparison with the traditional BRB. In these BRBs, the rule whose activation rate is larger than the threshold value will be selected to participate in the optimization process. Thus, the different threshold values will retain different numbers of rules. The optimization algorithm is P-CMA-ES, and the iteration time is 100. Table 6 shows the details and testing results generated by different BRBs. Where BRB stands for traditional BRB, BRB_1 to BRB_4 represents four BRBs with different activation thresholds, BRB_SQP stands for BRB using Sequential Quadratic Programming optimization algorithm, BRB_PSO stands for BRB using Particle Swarm Optimization algorithm, BRB_S stands for BRB using only the selected rules. Table 7 shows the health status distribution of different laser gyros under different BRBs.

It can be seen that BRB_1 with fewer rules generates the best result. The mean square error (MSE) of BRB_1 with 243 optimization parameters is even smaller than that of the traditional BRB with 259 optimization parameters.

Figures 6, 7 and 8 can better demonstrate the results of comparative experiments, where 1500 samples are divided into three parts.

Remark 3 Note that the number of selected rules shown in Table 6 does not mean that the BRB expert system only has these rules; the remaining rules are also included. They do not participate in the optimization process.

Remark 4 As seen from the above experimental results, although BRB_1 achieves the best accuracy with minimum rules in the optimization process, the accuracy is not sufficient because of the few iterations of the P-CMA-ES algorithm. It is obvious that, in theory, further optimization will not have an obvious effect.

Case 2-parallel processing for BRBs using activation rates. In this section, all samples are treated as training data to activate more rules in the BRB to start up the parallel optimization process of the BRB. Figure 9 shows the activation rates generated by training data. The activation rates of 64 rules are more balanced than the activation rates shown in Fig. 5 of case 4.1.2. In Fig. 9, most rules have been activated, although some rules are activated less frequently. When more rules participate in the optimization process, the number of parameters to be optimized will increase dramatically. Therefore, the method of decomposing the data set is proposed to generate independent optimization subprocesses.

Remark 5 Note that the training data set of optimization subprocesses is divided into some independent parts, and each part is average sampled from all samples in the original data set. The number of parts of the training data set is equal to the number of optimization subprocesses whose parameters are the same as those in case 4.1.2.

By using the parallel optimization method described in “Parallel optimization method of BRBs” section, the training data sets are divided into 6 parts, each of which belongs to the corresponding 6 sub-BBR models, denoted as sub.1–6. Thus, the parallel optimization processes can also be divided into 6 parts. The operating environment uses the MATLAB parallel toolkit with an Intel Core i7-8750H 2.2 GHz CPU and 16 GB memory. The optimization algorithm is P-CMA-ES, and the iteration time is 100. The activation rates of 64 rules generated by subs.1–6 are shown in Fig. 10.

In this case, the activation rate threshold of the parallel optimization is 0.001; then, the number of rules involved in the optimization of each sub-BRB are shown in Table 8.

Thus, the 6 sub-BRBs are assigned to 6 optimization subprocesses and optimized independently. The comparison results between the original BRB without using the activation rate and optimization parallelization, the BRB with the activation rate but without optimization parallelization (nonparallelization BRB_a), and the BRB with the activation rate and optimization parallelization (parallelization BRB_a) are shown in Table 9.

In Table 9, parallelization BRB_a uses the least time to obtain the best results. Its running time is approximately one-third that of other models. It can be predicted that with the increase in the number of processors, the running time will decrease without affecting the accuracy.

The experimental results show that the BRB expert system with activation rate has higher accuracy than the initial BRB or other BRB optimization algorithms in the process of laser gyro health state assessment, and the time spent is also very little. It effectively solves the problem of BRB expert system combination explosion.

Rule	Weight	Attributes			Belief degrees distribution
		$a_1(t)$	$a_2(t)$	$a_3(t)$	$\{H, SH, UH\}$
1	1	NL	PS	NL	{1, 0, 0}
2	1	NL	PS	NM	{0.9, 0.1, 0}
3	1	NL	PS	NS	{0.8, 0.2, 0}
4	1	NL	PS	NVS	{0.7, 0.3, 0}
5	1	NL	PM	NL	{0.9, 0.1, 0}
6	1	NL	PM	NM	{0.8, 0.1, 0.1}
7	1	NL	PM	NS	{0.7, 0.2, 0.1}
8	1	NL	PM	NVS	{0.7, 0.3, 0}
9	1	NL	PL	NL	{0.8, 0.2, 0}
10	1	NL	PL	NM	{0.7, 0.2, 0.1}
11	1	NL	PL	NS	{0.6, 0.3, 0.1}
12	1	NL	PL	NVS	{0.6, 0.4, 0}
13	1	NL	PVL	NL	{0.8, 0.2, 0}
14	1	NL	PVL	NM	{0.7, 0.3, 0}
15	1	NL	PVL	NS	{0.6, 0.4, 0}
16	1	NL	PVL	NVS	{0.5, 0.4, 0.1}
17	1	NS	PS	NL	{0.8, 0.2, 0}
18	1	NS	PS	NM	{0.7, 0.2, 0.1}
19	1	NS	PS	NS	{0.6, 0.3, 0.1}
20	1	NS	PS	NVS	{0.5, 0.4, 0.1}
21	1	NS	PM	NL	{0.5, 0.3, 0.2}
22	1	NS	PM	NM	{0.4, 0.5, 0.1}
23	1	NS	PM	NS	{0.3, 0.7, 0}
24	1	NS	PM	NVS	{0.3, 0.6, 0.1}
25	1	NS	PL	NL	{0.1, 0.6, 0.3}
26	1	NS	PL	NM	{0.1, 0.6, 0.3}
27	1	NS	PL	NS	{0.1, 0.6, 0.3}
28	1	NS	PL	NVS	{0.1, 0.6, 0.3}
29	1	NS	PVL	NL	{0.1, 0.3, 0.6}
30	1	NS	PVL	NM	{0.1, 0.5, 0.4}
31	1	NS	PVL	NS	{0.1, 0.3, 0.6}
32	1	NS	PVL	NVS	{0.1, 0.2, 0.7}
33	1	PS	PS	NL	{0.3, 0.5, 0.2}
34	1	PS	PS	NM	{0.1, 0.6, 0.3}
35	1	PS	PS	NS	{0.1, 0.5, 0.4}
36	1	PS	PS	NVS	{0.1, 0.4, 0.5}
37	1	PS	PM	NL	{0.1, 0.6, 0.3}
38	1	PS	PM	NM	{0.1, 0.7, 0.2}
39	1	PS	PM	NS	{0, 0.6, 0.4}
40	1	PS	PM	NVS	{0, 0.5, 0.5}
41	1	PS	PL	NL	{0, 0.5, 0.5}
42	1	PS	PL	NM	{0, 0.6, 0.4}
43	1	PS	PL	NS	{0, 0.5, 0.5}
44	1	PS	PL	NVS	{0, 0.4, 0.6}
45	4	PS	PVL	NL	{0.1, 0.4, 0.5}
46	1	PS	PVL	NM	{0, 0.4, 0.6}
47	1	PS	PVL	NS	{0, 0.5, 0.5}
48	1	PS	PVL	NVS	{0, 0.4, 0.6}
49	1	PL	PS	NL	{0.2, 0.3, 0.5}
50	1	PL	PS	NM	{0.3, 0.2, 0.5}
51	1	PL	PS	NS	{0.1, 0.4, 0.5}
52	1	PL	PS	NVS	{0, 0.5, 0.5}
53	1	PL	PM	NL	{0.3, 0.2, 0.5}
54	1	PL	PM	NM	{0.2, 0.3, 0.5}
55	1	PL	PM	NS	{0.1, 0.4, 0.5}

Continued

Rule	Weight	Attributes			Belief degrees distribution
		$a_1(t)$	$a_2(t)$	$a_3(t)$	$\{H, SH, UH\}$
56	1	PL	PM	NVS	{0, 0.4, 0.6}
57	1	PL	PL	NL	{0.1, 0.4, 0.5}
58	1	PL	PL	NM	{0.1, 0.4, 0.5}
59	1	PL	PL	NS	{0.1, 0.3, 0.6}
60	1	PL	PL	NVS	{0, 0.3, 0.7}
61	1	PL	PVL	NL	{0.1, 0.1, 0.8}
62	1	PL	PVL	NM	{0, 0.2, 0.8}
63	1	PL	PVL	NS	{0, 0.1, 0.9}
64	1	PL	PVL	NVS	{0, 0, 1}

Table 5. Initial parameters of the BRB expert system.

Details Models	Threshold	Number of selected rules	Number of parameters	MSE
BRB	--	64	259	0.0224
BRB_1	0.005	60	243	0.0118
BRB_2	0.007	54	219	0.0558
BRB_3	0.0156	31	127	0.4974
BRB_4	0.02	15	63	1.1067
BRB_SQP	--	64	259	0.0218
BRB_PSO	--	64	259	0.0221
BRB_S	--	60	243	0.3362

Table 6. The details and final testing results generated by different BRBs.

Leak size estimation of oil pipeline. *Problem formulation.* In this case, an actual data set of an oil pipeline leak in Britain is used to prove the advantage of the proposed method¹. This oil pipeline is a hundred kilometers long, and when the pipeline leaks, the flow rate and pressure of the oil in the pipeline will change according to a certain mode. The data set has 2007 samples which include flow difference, average pressure difference, and leak size, as shown in Figs. 11, 12 and 13.

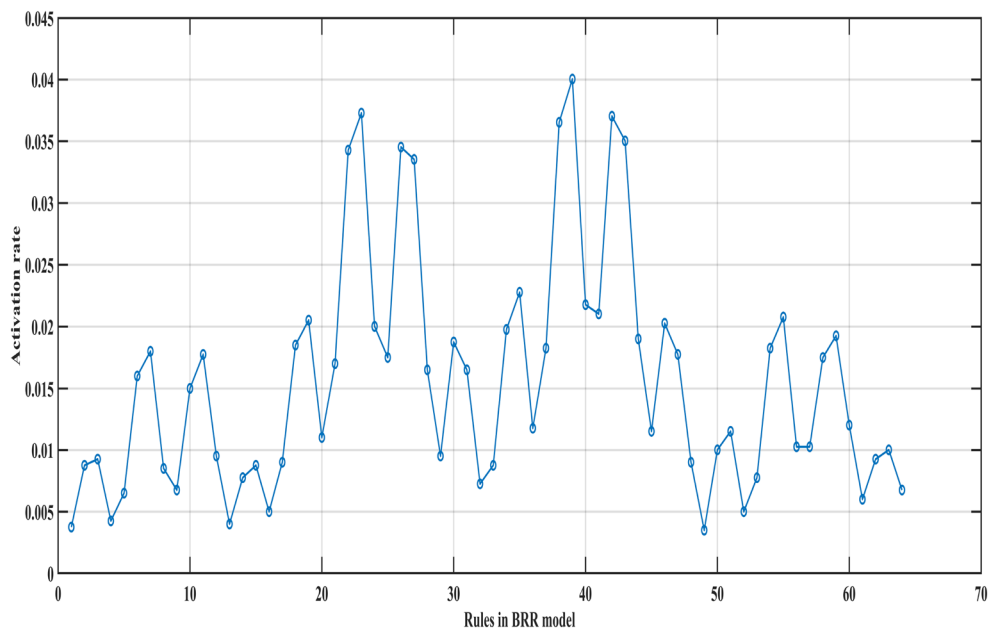


Figure 5. The activation rates of 64 rules in the BRB expert system.

Details Models	Threshold	health	sub-health	unhealth
BRB	--	77.94%	13.93%	8.13%
BRB_1	0.005	79.64%	13.27%	7.09%
BRB_2	0.007	73.76%	15.61%	10.63%
BRB_3	0.0156	61.48%	20.49%	18.03%
BRB_4	0.02	50.21%	25.82%	23.97%
BRB_SQP	--	79.15%	13.01%	7.84%
BRB_PSO	--	78.63%	13.68%	7.69%
BRB_S	--	68.96%	17.69%	13.35%

Table 7. Average belief degrees distribution of different BRBs.

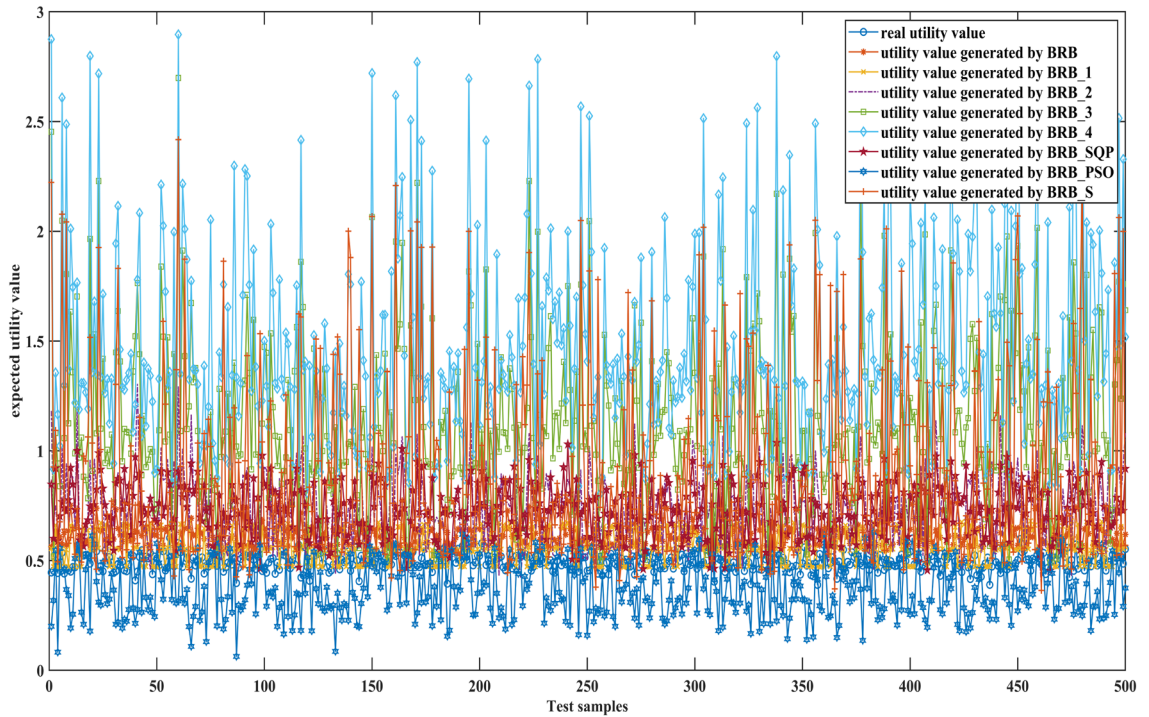


Figure 6. The 1–500 comparative results generated by different BRBs.

The reference values of flow difference and pressure difference are shown in Tables 10, 11. Thus, the BRB expert system of pipeline leak detection can be described as

$$R_k : \text{If flow_difference is } A_1^k \wedge \text{pressure_difference is } A_2^k \\ \text{Then Leak_size is } \{ (Z, \beta_1^k), (VS, \beta_2^k), (M, \beta_3^k), (H, \beta_4^k), (VH, \beta_5^k) \} \\ \text{with rule weight } \theta_k \text{ and attribute weight } \delta_i$$

where Z, VS, M, H, VH denote the reference values of leak size, as shown in Table 12. The other parameters of the BRB expert system are shown in Table 13.

As described in the above sections, the data set may not cover all patterns about flow difference and pressure difference, which means that perhaps only a part of the rules shown in Table 13 can be activated. To prove this point, the samples of oil pipeline leaks are entered into the above BRB expert system. The samples are divided into training data set and testing data set. The training data set includes 510 samples, which are selected from the 2007 samples regularly, and the remaining 1497 samples are used as a testing data set. P-CMA-ES algorithm is utilized as an optimization tool of the BRB expert system, and to ensure the fairness of the optimization process, the parameters and iterations are the same for different conditions.

Case 1-optimization process of BRB using activation rates. Figure 14 shows the activation rates of 56 rules in the BRB expert system when the training data set are entered. It can be seen that only a few rules are activated and their activation times of them are different.

Four BRBs with different activation thresholds were established to compare with traditional BRBs. In these BRBs, rules with activation rates greater than the threshold are selected to participate in the optimization process. Therefore, different thresholds preserve different numbers of rules. The optimization algorithm is P-CMA-ES with an iteration time of 100. Table 14 shows the details and test results for different BRBs. Where BRB stands for traditional BRB, BRB_1 to BRB_4 represents four BRBs with different activation thresholds, BRB_SQP stands for

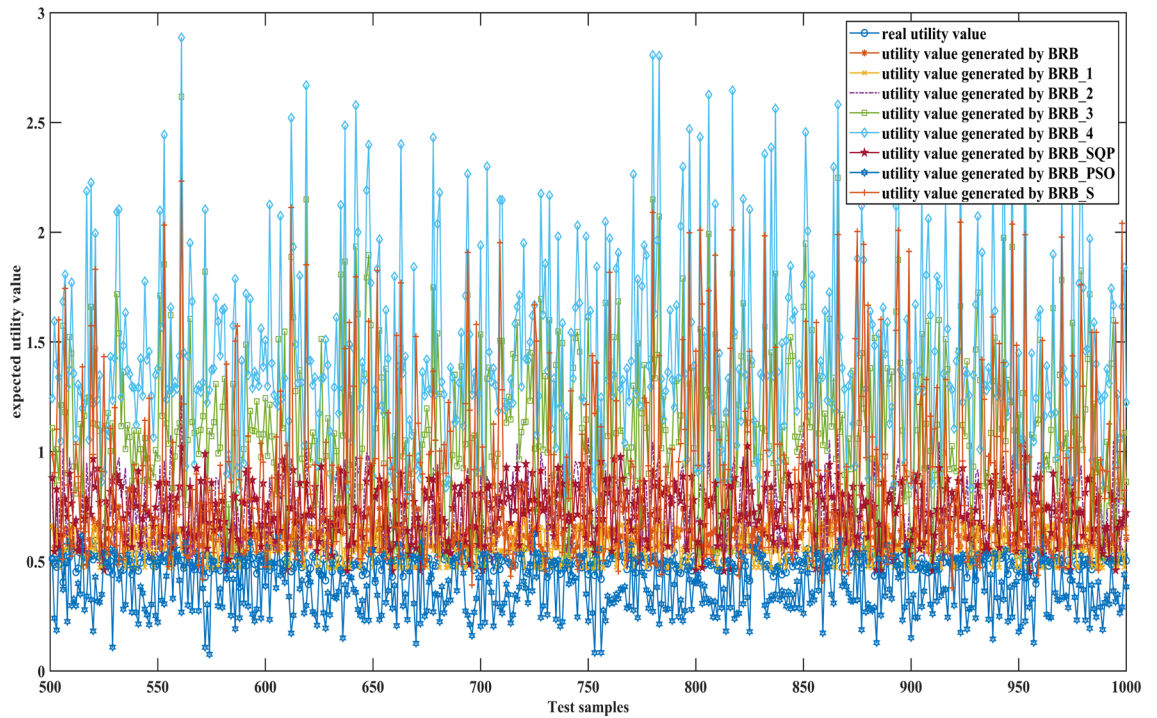


Figure 7. The 501–1000 comparative results generated by different BRBs.

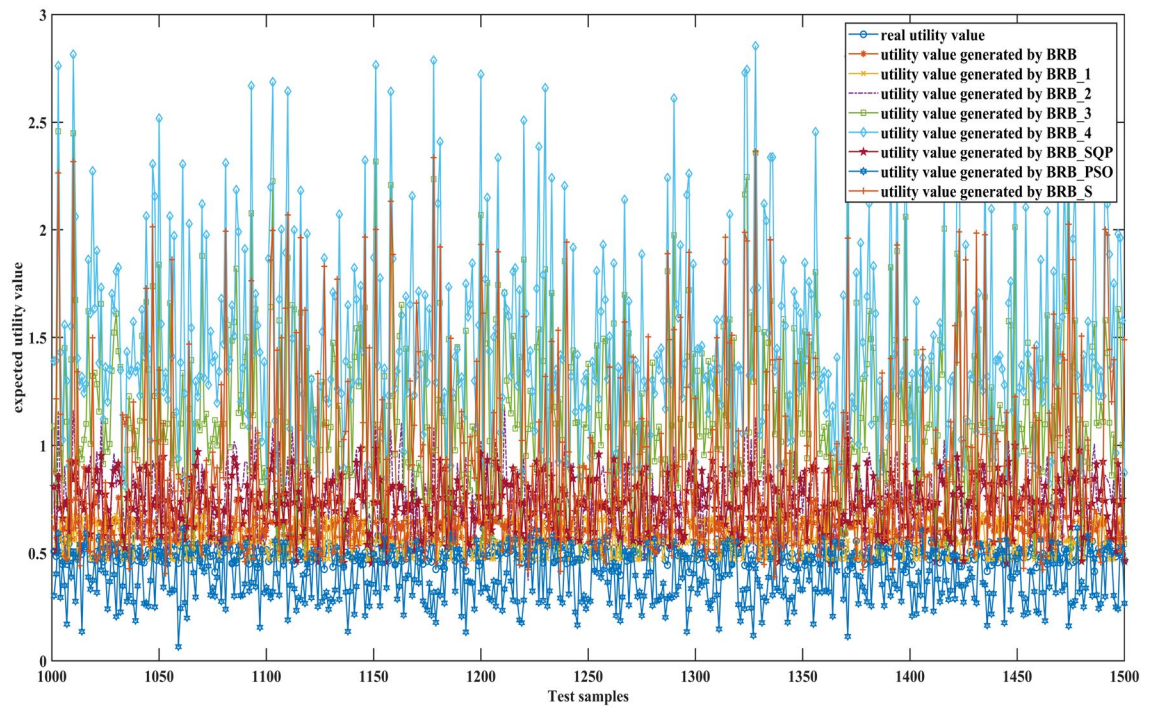


Figure 8. The 1001–1500 comparative results generated by different BRBs.

BRB using Sequential Quadratic Programming optimization algorithm, BRB_PSO stands for BRB using Particle Swarm Optimization algorithm, BRB_S stands for BRB using only the selected rules.

It can be seen that BRB_4 with minimum rules generates the best result. The mean square error (MSE) of BRB_4 with 26 optimization parameters is even smaller than traditional BRB with 338 optimization parameters.

Figures 15, 16 and 17 can better demonstrate the results of comparative experiments, where 1497 samples are divided into three parts.

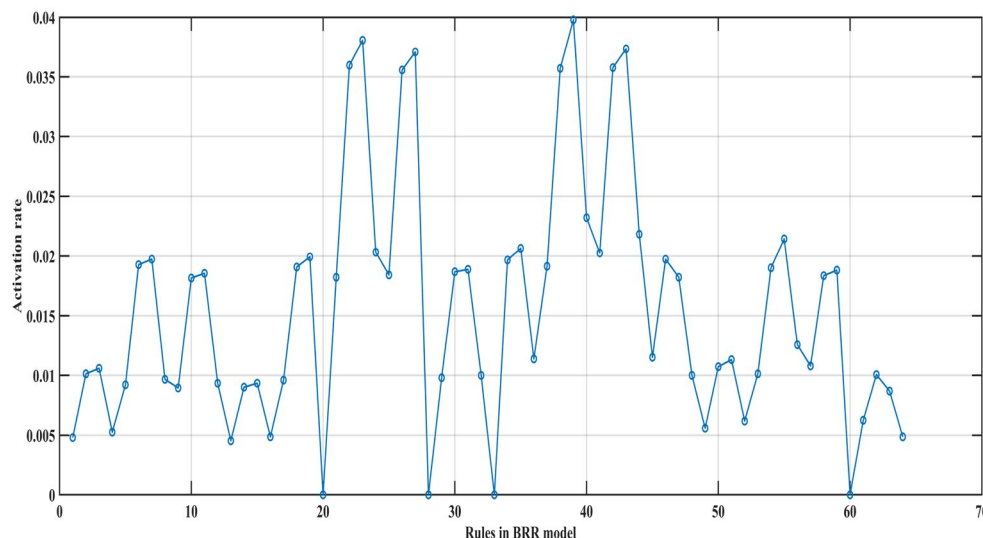


Figure 9. The activation rates of 64 rules generated by all samples.

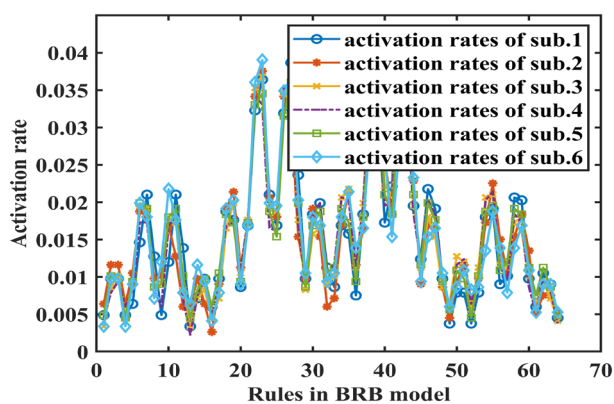


Figure 10. The activation rates of 64 rules generated by sub.1–6.

Sub-BRBs	Sub1	Sub2	Sub3	Sub4	Sub5	Sub6
The number of rules	38	42	43	43	44	41

Table 8. The number of rules involved in optimization in each sub-BRB.

Details		Threshold	Running time(s)	MSE
Models				
Original BRB	--	--	1733.0071	0.0224
nonparallelization BRB_a	0.001	--	1743.0693	0.5102
nonparallelization BRB_a	0.015	--	1762.9366	0.8611
parallelization BRB_a	0.001	--	635.1335	0.0217
BRB_SQP	--	--	2742.3305	0.0218
BRB_PSO	--	--	1668.2579	0.0221
BRB_S	--	--	1496.4573	0.3362

Table 9. The comparison results between different BRBs.

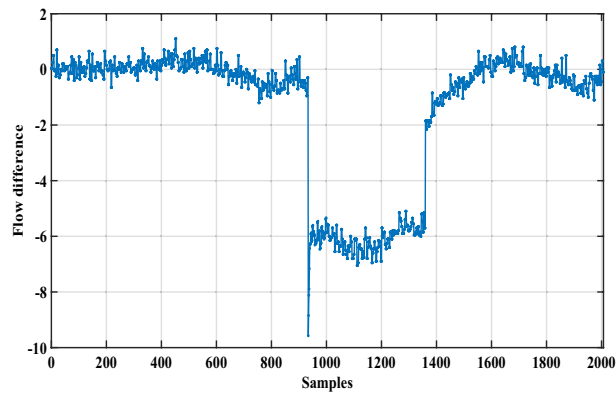


Figure 11. The flow difference of the oil pipeline.

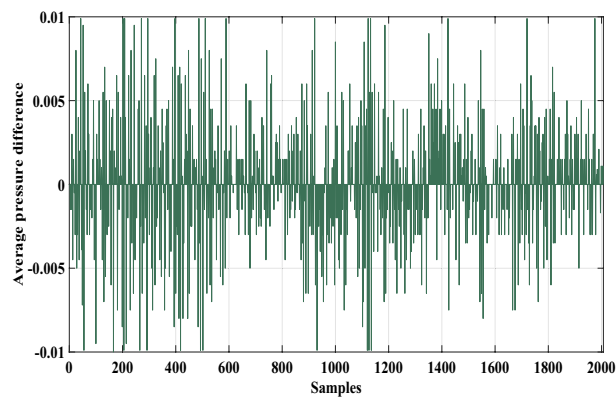


Figure 12. The average pressure difference of the oil pipeline.

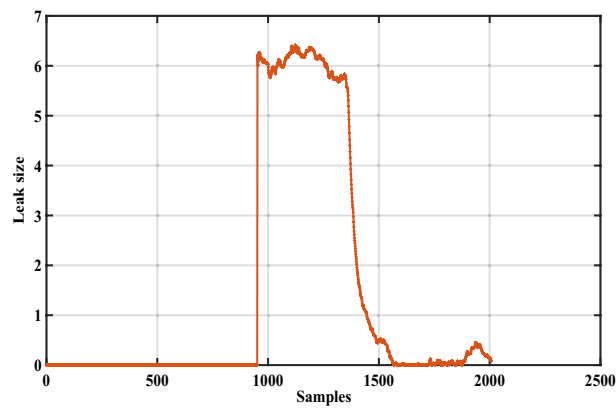


Figure 13. The leak size of the oil pipeline.

Linguistic terms	Negative large	Negative medium	Negative small	Negative very small
Values	- 10	- 5	- 3	- 1
Linguistic terms	Zero	Positive small	Positive medium	Positive large
Values	0	1	2	3

Table 10. Referential values of flow difference.

Linguistic terms	Negative large	Negative medium	Negative small	Negative very small
Values	- 0.042	- 0.025	- 0.01	void
Linguistic terms	Zero	Positive small	Positive medium	Positive large
Values	0	0.01	0.025	0.042

Table 11. Referential values of average pressure difference.

Linguistic terms	Zero	Very small	Medium	High	Very high
Values	0	2	4	6	8

Table 12. Referential values of leak size.

Case 2-parallel processing for BRB using activation rates. In this section, all samples are treated as training data to activate more rules in BRB, which is to start up the parallel optimization process of BRB. Figure 18 shows the activation rates generated by training data. It can be seen that the activation rates of 56 rules are more balanced than activation rates shown in Fig. 14 of case 4.2.1. In Fig. 18, most rules have been activated, although some rules are activated less frequently. When more rules participate in the optimization process, the number of parameters to be optimized will increase dramatically. Therefore, the method of the decomposing data set is proposed to generate independent optimization sub-processes.

By using the parallel optimization method described in “Parallel optimization method of BRBs” section, the training data sets are divided into 6 parts, each of which belongs to the corresponding 6 sub-BBR models, denoted as sub.1–6. Thus, the parallel optimization processes can be also divided into 6 parts. The operating environment uses MATLAB parallel toolkit with Intel Core i7-8750H 2.2 GHz CPU, 16 GB memory. The optimization algorithm is P-CMA-ES, and the iteration time is 100. The activation rates of 56 rules generated by sub.1–6 are shown in Fig. 19.

In this case, the activation rate threshold of the parallel optimization is 0.002, then the number of the rules involved in optimization of each sub-BRB are shown in Table 15.

Thus, the 6 sub-BRBs are assigned to 6 optimization sub-processes and optimized independently. The comparison results between original BRB without using activation rate and optimization parallelization, BRB with activation rate but without optimization parallelization (non-parallelization BRB_a), and BRB with activation rate and optimization parallelization (parallelization BRB_a) are shown in Table 16.

In Table 16, parallelization BRB_a uses the least time to get the best results. Its running time is about one-fourth of other models. It can be predicted that with the increase of the number of processors, the running time will decrease without affecting the accuracy.

The experimental results show that the BRB expert system with activation rate has higher accuracy than the initial BRB or other BRB optimization algorithms in the process of oil pipeline leakage size assessment, and it takes less time. It effectively solves the problem of BRB expert system combination explosion.

Conclusion

To solve the combinatorial explosion problem of the BRB expert system, a novel optimization method is proposed in this paper. The proposed method can be applied in two situations: (1) Only a few rules in BRB are activated. In this case, activation rate is used to prune the rules that have never been activated or are inadequately activated. The initial parameters of these removed rules are given by experts, which will not only keep the role of expert experience but also enhance the credibility of the results. (2) Most rules in BRB are activated. In this case, parallel optimization is proposed by decomposing training data, and activation rate is utilized in each parallel optimization sub-process. The final results of BRB can be obtained by the weighted average method. The proposed optimization method for the BRB expert system can be applied to all kinds of BRB models and can increase the accuracy while reducing the calculation pressure by reducing the parameters and using parallel operation.

Rule	Weight	Flow and pressure difference with weight = 1	Belief degrees distribution for {Z, VS, M, H, VH}
1	1	Negative large, Negative large	0 0 0 1
2	1	Negative large, Negative medium	0 0 0 0.3 0.7
3	1	Negative large, Negative small	0 0 0.2 0.8 0
4	1	Negative large, Zero	0 0 0.8 0.2 0
5	1	Negative large, Positive small	0.65 0.35 0 0 0
6	1	Negative large, Positive medium	0.85 0.15 0 0 0
7	1	Negative large, Positive large	0.95 0.05 0 0 0
8	1	Negative medium, Negative large	0 0 0.1 0.9 0
9	1	Negative medium, Negative medium	0 0 0.7 0.3 0
10	1	Negative medium, Negative small	0 0 0.7 0.3 0 0
11	1	Negative medium, Zero	0 0 0.9 0.1 0 0
12	1	Negative medium, Positive small	0.8 0.2 0 0 0
13	1	Negative medium, Positive medium	0.9 0.1 0 0 0
14	1	Negative medium, Positive large	0.99 0.01 0 0 0
15	1	Negative small, Negative large	0 0 0.4 0.6 0
16	1	Negative small, Negative medium	0 0 0.8 0.2 0
17	1	Negative small, Negative small	0 0.3 0.6 0.1 0
18	1	Negative small, Zero	0.1 0.7 0.2 0 0
19	1	Negative small, Positive small	0.7 0.3 0 0 0
20	1	Negative small, Positive medium	0.9 0.1 0 0 0
21	1	Negative small, Positive large	1 0 0 0 0
22	1	Negative very small, Negative large	0.02 0.11 0.39 0.48 0
23	1	Negative very small, Negative medium	0.1 0.78 0.12 0 0
24	1	Negative very small, Negative small	0.36 0.64 0 0 0
25	1	Negative very small, Zero	1 0 0 0 0
26	1	Negative very small, Positive small	1 0 0 0 0
27	1	Negative very small, Positive medium	1 0 0 0 0
28	1	Negative very small, Positive large	1 0 0 0 0
29	1	Zero, Negative large	1 0 0 0 0
30	1	Zero, Negative medium	1 0 0 0 0
31	1	Zero, Negative small	1 0 0 0 0
32	1	Zero, Zero	1 0 0 0 0
33	1	Zero, Positive small	1 0 0 0 0
34	1	Zero, Positive medium	1 0 0 0 0
35	1	Zero, Positive large	1 0 0 0 0
36	1	Positive small, Negative large	0.39 0.61 0 0 0
37	1	Positive small, Negative medium	0.9 0.1 0 0 0
38	1	Positive small, Negative small	1 0 0 0 0
39	1	Positive small, Zero	1 0 0 0 0
40	1	Positive small, Positive small	1 0 0 0 0
41	1	Positive small, Positive medium	1 0 0 0 0
42	1	Positive small, Positive large	1 0 0 0 0
43	1	Positive medium, Negative large	0.1 0.9 0 0 0
44	1	Positive medium, Negative medium	0.3 0.7 0 0 0
45	1	Positive medium, Negative small	0.85 0.15 0 0 0
46	1	Positive medium, Zero	0.98 0.02 0 0 0
47	1	Positive medium, Positive small	1 0 0 0 0
48	1	Positive medium, Positive medium	1 0 0 0 0
49			1 0 0 0 0
50	1	Positive large, Negative large	0.9 0.1 0 0 0
51	1	Positive large, Negative medium	0.99 0.01 0 0 0
52	1	Positive large, Negative small	1 0 0 0 0
53	1	Positive large, Zero	1 0 0 0 0
54	1	Positive large, Positive small	1 0 0 0 0
55	1	Positive large, Positive medium	1 0 0 0 0
56	1	Positive large, Positive large	1 0 0 0 0

Table 13. Initial parameters of BRB expert system.

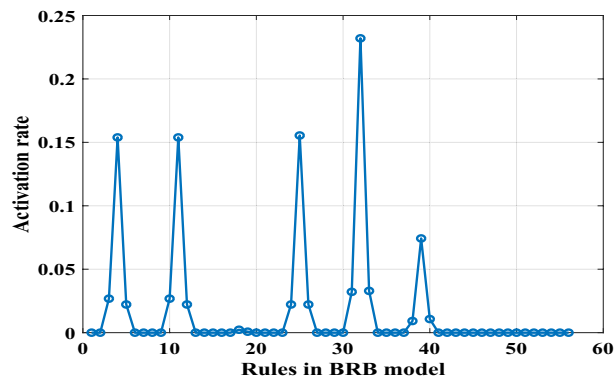


Figure 14. The activation rates of 56 rules in the BRB expert system.

Models \ Details	Threshold	Number of selected rules	Number of parameters	MSE
BRB	--	56	338	0.8451
BRB_1	0	17	104	1.4809
BRB_2	0.01	14	86	1.4663
BRB_3	0.05	5	32	0.8267
BRB_4	0.1	4	26	0.7811
BRB_SQP	--	56	338	0.8027
BRB_PSO	--	56	338	0.8349
BRB_S	--	4	26	0.9154

Table 14. The details and final testing results generated by different BRBs.

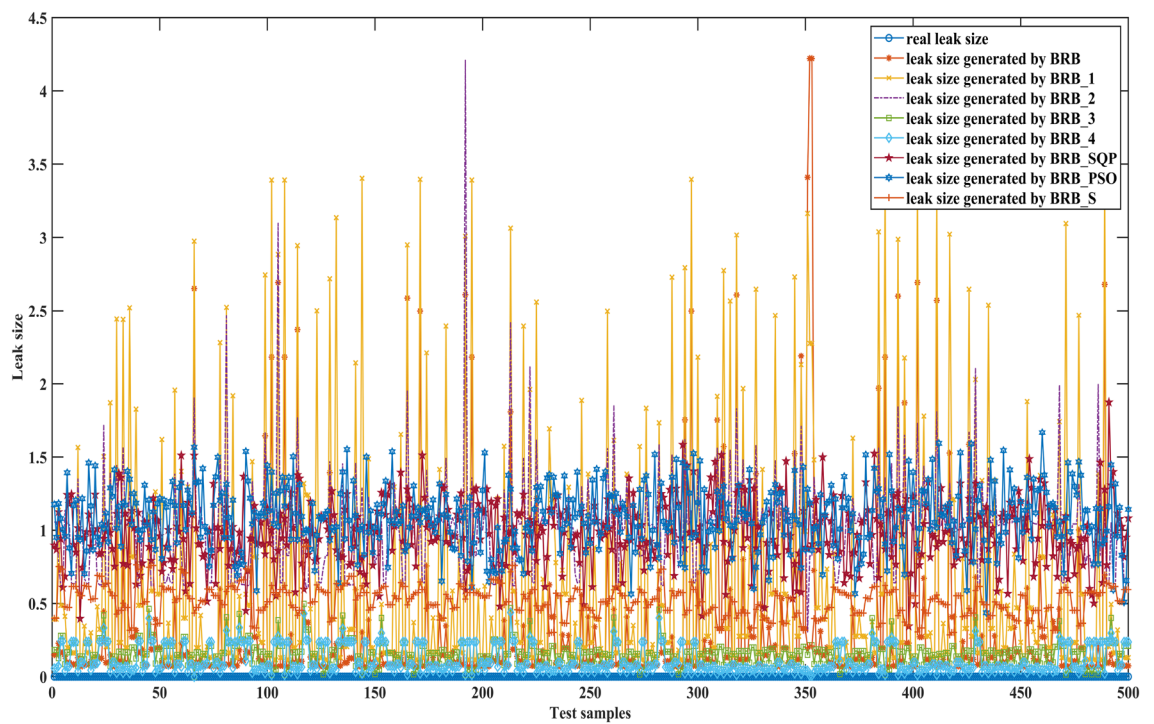


Figure 15. The 1–500 comparative results generated by different BRBs.

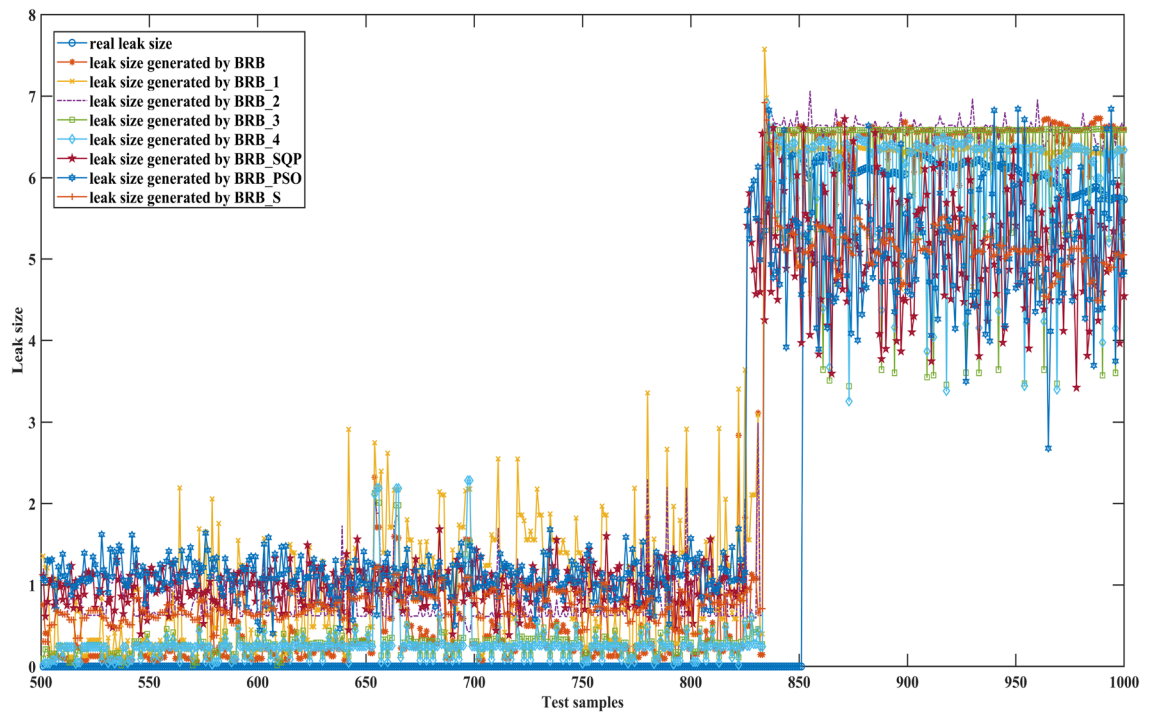


Figure 16. The 501–1000 comparative results generated by different BRBs.

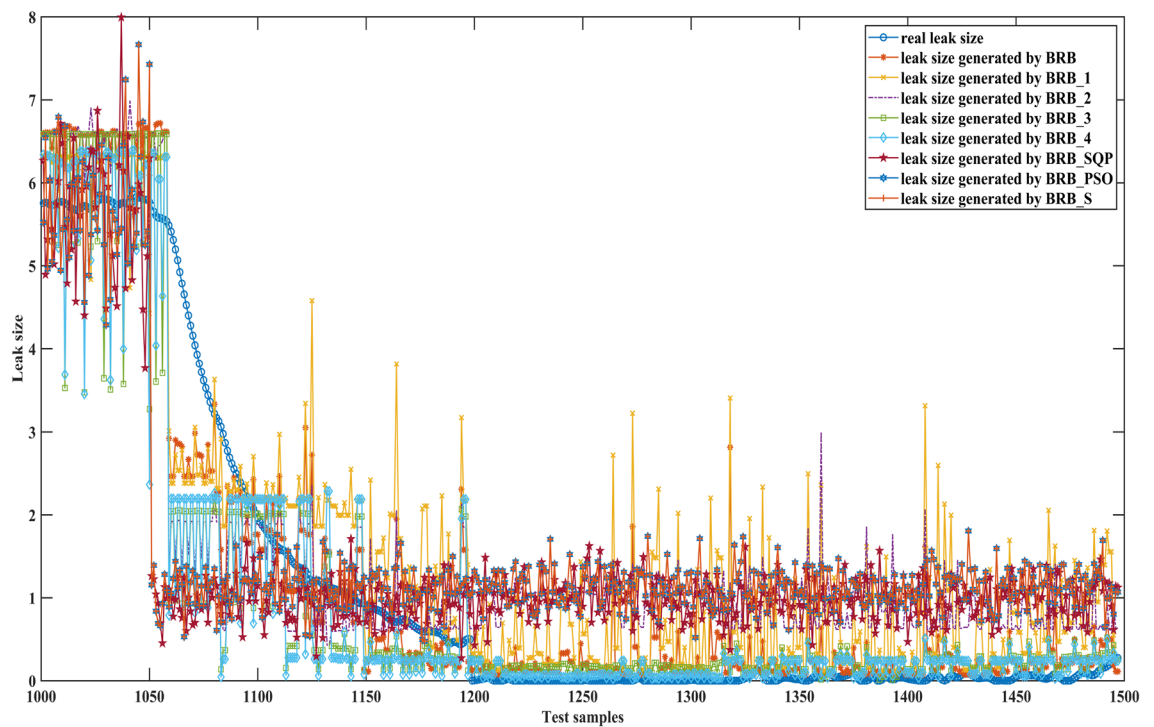


Figure 17. The 1001–1497 comparative results generated by different BRBs.

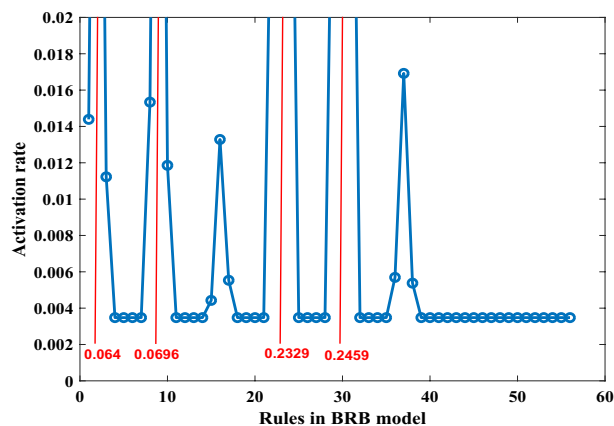


Figure 18. The activation rates of 56 rules generated by all samples.

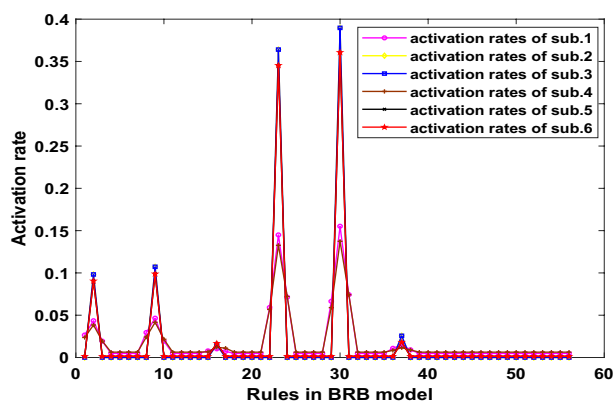


Figure 19. The activation rates of 56 rules generated by sub.1–6.

Sub-BRBs	sub1	sub2	sub3	sub4	sub5	sub6
The number of rules	56	6	6	56	6	6

Table 15. The number of rules involved in optimization in each sub-BRB.

Details Models	Threshold	Running time(s)	MSE
Original BRB	--	1758.0274	0.8451
nonparallelization BRB_a	0.005	1634.7291	1.4628
nonparallelization BRB_a	0.002	1595.6152	0.6873
parallelization BRB_a	0.002	496.6395	0.6133
BRB_SQP	--	2377.3517	0.8027
BRB_PSO	--	1125.1589	0.8349
BRB_S	--	311.9046s	0.9154

Table 16. The comparison results between different BRBs.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due Laboratory requirements but are available from the corresponding author on reasonable request.

Received: 1 August 2022; Accepted: 3 January 2023

Published online: 11 January 2023

References

- Duda, R. O. & Shortliffe, E. H. Expert systems research. *Science* **220**(4594), 261 (1983).
- Ebadzadeh, M. M. & Salimi-Badr, A. IC-FNN: A novel fuzzy neural network with interpretable intuitive and correlated-contours fuzzy rules for function approximation. *IEEE Trans. Fuzzy Syst.* **26**(3), 1288–1302 (2018).
- He, W., Chen, Y. & Yin, Z. Adaptive neural network control of an uncertain robot with full-state constraints. *IEEE Trans. Cybern.* **46**(3), 620–629 (2017).
- Ghahramani, Z. An introduction to hidden Markov models and Bayesian networks. *Int. J. Pattern Recognit. Artif. Intell.* **15**(1), 9–42 (2001).
- Greenspan, H., Ginneken, B. & Summers, R. M. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* **35**(5), 1153–1159 (2016).
- Zhou, Z. J., Hu, C. H. & Xu, D. L. A model for real-time failure prognosis based on hidden Markov model and belief rule base. *Eur. J. Oper. Res.* **207**(1), 269–283 (2010).
- Juang, C. F. & Chen, C. Y. Data-driven interval type-2 neural fuzzy system with high learning accuracy and improved model interpretability. *IEEE Trans. Cybern.* **43**(6), 1781–1795 (2013).
- Zhou, Z. J., Hu, G. Y., Hu, C. H., Wen, C. L. & Chang, L. L. A survey of belief rule base expert system. *IEEE Trans. Syst. Man Cybern. Syst.* <https://doi.org/10.1109/TSMC.2019.2944893> (2019).
- Yang, J. B., Liu, J. & Wang, J. Belief rule-base inference methodology using the evidential reasoning approach-RIMER. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **36**(2), 266–285 (2006).
- Zhou, Z. J., Hu, C. H., Yang, J. B., Xu, D. L. & Zhou, D. H. Online updating belief-rule-based systems using the RIMER approach. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **41**(6), 1225–1243 (2011).
- Hu, G. Y. *et al.* A method for predicting the network security situation based on hidden BRB model and revised CMA-ES algorithm. *Appl. Soft Comput.* **48**, 404–418 (2016).
- Hu, G. Y. & Qiao, P. L. Cloud belief rule base model for network security situation prediction. *IEEE Commun. Lett.* **20**(5), 914–917 (2016).
- Yang, J. B., Wang, Y. M., Xu, D. L., Chin, K. S. & Chatton, L. Belief rule-based methodology for mapping consumer preferences and setting product targets. *Expert Syst. Appl.* **39**, 4749–4759 (2012).
- Feng, Z. C. *et al.* Fault diagnosis based on belief rule base with considering attribute correlation. *IEEE Access* **6**, 1–1 (2017).
- Zhou, Z. G. *et al.* A bilevel belief rule-based decision support system for diagnosis of lymph node metastasis in gastric cancer. *Knowl.-Based Syst.* **54**, 128–136 (2013).
- Chang, L. L., Zhou, Y., Jiang, J., Li, M. J. & Zhang, X. H. Structure learning for belief rule base expert system: A comparative study. *Knowl.-Based Syst.* **39**, 159–172 (2013).
- Kim, T. K. & Kittler, J. Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(3), 318–327 (2005).
- Zhang, B. C. *et al.* Network intrusion detection based on directed acyclic graph and belief rule base. *ETRI J.* **39**(4), 592–604 (2017).
- Zhou, Z. J. *et al.* Online updating belief rule based system for pipeline leak detection under expert intervention. *Expert Syst. Appl.* **36**(4), 7700–7709 (2009).
- Bustince, H. & Burillo, P. Mathematical analysis of interval-valued fuzzy relations: Application to approximate reasoning. *Fuzzy Sets Syst.* **113**(2), 205–219 (2000).
- Chen, S. M. & Hsiao, W. H. Bidirectional approximate reasoning for rule-based systems using interval-valued fuzzy sets. *Fuzzy Sets Syst.* **113**(2), 185–203 (2000).
- Zimmermann, H. J. An application-oriented view of modeling uncertainty. *Eur. J. Oper. Res.* **122**(2), 190–199 (2000).
- Jian, J. B. A superlinearly and quadratically convergent SQP type feasible method for constrained optimization. *Appl. Math.-A J. Chin. Univ. (B)* **15**(3), 319–332 (2000).
- Zhou, Z. J. *et al.* A model for hidden behavior prediction of complex systems based on belief rule base and power set. *IEEE Trans. Syst. Man Cybern. Syst.* **88**(PP), 1–7 (2017).

Author contributions

Conceptualization, G.X., J.W. and G.H.; methodology, G.X., G.H.; software, J.W.; validation, G.X., J.W. and G.H.; formal analysis, G.X.; investigation, X.H.; data curation S.T.; writing—original draft preparation, J.W.; writing—review and editing, G.X. and G.H.; visualization, J.W. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 62273113, China Postdoctoral Science Foundation under Grant 2020M673668, and Guangxi Key Laboratory of Trusted Software under Grant KX202050, and National Defense Science and Technology Key Laboratory fund under grant 6142101210309, and basic scientific research items of equipment under Grant 514010203-214.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to G.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023