A Belief Rule Extraction Method Based on Rough Sets

YU Meng, HUANG Jian and KONG Jiangtao
National University of Defense Technology, Changsha, China
1546737193@qq.com

Abstract. Most of the existing research on belief rule-base focused on parameter optimization problem. However, there are few studies on the problem of establishing the initial belief rule-base due to the difficulty in obtaining expert experience and the unconformity of standards. Based on this, a method of extracting belief rule based on rough sets is proposed. First, the initial attribute index is used to divide the data set into equivalent classes. Then, the belief rule is extracted with the knowledge of rough sets. Finally, the extracted belief rule is used for reasoning and the index is optimized according to the result of reasoning. An example of pipeline leak detection is introduced in the experiment section. Experimental results show that the method is effective and feasible.

1. Introduction
In order to effectively use quantitative information and incomplete or uncertain information provided by experts to model complex problems, Yang et al.[1] proposed a belief rule-base inference methodology using the evidential reasoning approach (RIMER). RIMER combines D-S evidence theory, decision theory, fuzzy theory, and traditional IF-THEN rule base knowledge to model data with ambiguous or fuzzy uncertainty, incomplete or probabilistic uncertainties, and nonlinear features. At present, RIMER has been widely used in pipeline leak detection[2-3], fault diagnosis[4-7], intent identification[8-10], and capability evaluation[11-13].

RIMER mainly includes the expression of knowledge and the reasoning of knowledge. Among them, the expression of knowledge is realized by the belief rule-base (BRB) system, and the reasoning of knowledge is carried out by using the evidence reasoning (ER) algorithm. Due to the uncertainty of expert experience, the initial BRB systems is often inaccurate. To this end, the obtained data set can be used to optimize the inference parameters of the BRB system in combination with the parameter training method. At present, the research on RIMER is mainly concentrated in this field. Including Yang et al. proposed to use the FMINCON function in Matlab[14] for parameter training; B Qian et al. proposed a particle swarm optimization algorithm with improved velocity update way and repair methods (PSO_VR) to deal with the NNOP of SPO-BRB[15]; J Liu et al. proposed a new learning method based on the given sample data for optimally generating a consistent BRB[16]; R Chang proposed a new simple optimization algorithm to realize the self-learning capability of the rule base parameters[17]. However, in the process of establishing the initial BRB system, due to the difficulties in obtaining expert experience and the inconsistency of standards, it is often the case that the initial BRB system has a long establishment period and the initial BRB system does not match the actual system. How to quickly and effectively establish the initial BRB system has become an urgent problem to be solved, and at present there is relatively little research in this area.

Based on the above reasons, this paper proposes a method based on rough set[18-19] for the
extraction of belief rule. Firstly, according to the attribute index, the data is divided into equivalence classes; then, the belief degree is calculated for the class of the same precondition attribute, and the belief rule is extracted; finally, the extracted belief rule is used for reasoning, and the initial attribute index is optimized according to the inference result, which improves the inference precision of the belief rule-base. An example of pipeline leak detection was introduced in the experimental part to verify the effectiveness of the method.

2. Related basic concepts

2.1. belief rule representation

The confidence rule is developed according to the IF-THEN rule, introducing a distributed confidence framework and weight parameters. The $k$th rule in the BRB system is as follows:

$$R_k: \text{if } x_1 \text{ is } A^k_1 \land x_2 \text{ is } A^k_2 \land \ldots \land x_M \text{ is } A^k_M$$

Then\{$(D_1, \beta_{1,k}), (D_2, \beta_{2,k}), \ldots, (D_N, \beta_{N,k})$\}

With a rule weight $\theta_k$ and attribute weight $\delta_i$.

2.2. Belief rule-base inference methodology

2.2.1 Calculation of the activation weight of the belief rule. The activation weight of the input information $x$ for the $k$th rule can be calculated by the following formula, ie

$$\omega_k = \frac{\theta_k \prod_{i=1}^{M} (\alpha^k_i)^{\delta_i}}{\sum_{l=1}^{N} \theta_l \prod_{i=1}^{M} (\alpha^l_i)^{\delta_i}}$$

(2)

Where $\omega_k \in [0, 1], k = 1, 2, \ldots, L; \alpha^k_i (i = 1, 2, \ldots, M)$ indicates the confidence of the $i$th input $x_i$ in the $k$th rule with respect to the reference value $A^k_i$.

2.2.2 ER algorithm fusion. After calculating the activation degree of the rule, the ER algorithm can be used to fuse the rules in the BRB. First, convert the belief degree in the result attribute to the basic probability mass.

$$m_{j,k} = \omega_k \beta_{j,k}$$

(3)

$$m_{D,k} = 1 - \omega_k \sum_{j=1}^{N} \beta_{j,k}$$

(4)

$$\bar{m}_{D,k} = 1 - \omega_k$$

(5)
\[ \tilde{m}_{D,k} = \omega_k (1 - \sum_{j=1}^{N} \beta_{j,k}) \]  

Then, using the ER parsing algorithm to fuse the \( L \) rules, the formula is as follows:

\[ \hat{\beta}_j = \frac{\mu \times \left[ \prod_{k=1}^{l} (\omega_i \beta_{j,k} + 1 - \omega_k \sum_{i=1}^{n} \beta_{i,k}) \right]}{1 - \mu \times \left[ \prod_{k=1}^{l} (1 - \omega_k) \right]} \]

\[ \mu = \frac{\sum_{j=1}^{N} \prod_{k=1}^{l} (\omega_j \beta_{j,k} + 1 - \omega_k \sum_{j=1}^{N} \beta_{j,k})}{-(N-1) \prod_{k=1}^{l} (1 - \omega_k \sum_{j=1}^{N} \beta_{j,k})} \]  

Among them, \( \hat{\beta}_j \) indicates the belief level with respect to the evaluation result \( D_j \).

2.3. Rough set theory  
Rough set theory is a data analysis theory proposed by Polish mathematician Z. Pawlak in 1982. Its basic idea is to divide the instances in the database according to the attribute values of different attributes, and then establish rules for the relationship between the divided subsets [20-22].

The knowledge expression system is an ordered pair \( S = (U, C, D, V, f) \), in which \( U \) is a non-empty finite set, called a universe, \( C \) is a set of conditional attributes, \( D \) is a set of decision attributes, \( A = C \cup D \) called a set of attributes. \( V \) is a collection of attribute values, the elements of the universe \( U \) are called objects or instances, \( f \) specifies the mapping between objects and attribute values.

In the knowledge expression system, for any attribute \( a \) in the attribute set \( A \), if the instance \( x_i \) and \( x_j \) have the same value for the attribute \( a \), we call \( x_i \) and \( x_j \) equal based on the attribute set and are indistinguishable. A collection of all equivalent records based on a set of attributes \( A \) is defined as an equivalence class. Records belonging to the same equivalence class are grouped into one class, called \( R \) a division based on an attribute set \( A \), expressed as \( U/\text{ind}(R) = \{R_1, R_2, \ldots, R_n\} \), \( [x]_R \) is an equivalence class contained \( x \) in the \( U/\text{ind}(R) \). When \( X \subseteq U \), and \( X \) is the sum of \( U/\text{ind}(R) = \{R_1, R_2, \ldots, R_n\} \), \( X \) is said to be \( R \)-definable, called the exact set of \( R \), otherwise \( X \) is \( R \) undefined, also known as the rough set of \( R \).

Rough sets can be characterized by upper and lower approximations. The lower approximation of \( X \) with respect to \( R \) is \( R_+(X) = \bigcup \{Y \in U/R; Y \subseteq X\} \), and the upper approximation of \( X \) with respect to \( R \) is \( R^+(X) = \bigcup \{Y \in U/R; Y \cap X \neq \emptyset\} \). The positive domain for the decision attribute \( D \) with respect to \( R \) is \( \text{pos}_R(X) = R_+(X) \), \( X \in U/\text{ind}(D) \). When \( R_+(X) \neq R^+(X) \), \( X \) is rough for \( R \), the roughness is defined as \( \alpha_R(X) = \text{card}(R_+(X))/\text{card}(R^+(X)) \).

3. Belief rule extraction method

3.1 Equivalence class division
First, establish the knowledge expression system $S = \{U, C, D, V, f\}$, where $U = \{x_1, x_2, x_3, \cdots, x_n\}$ is a set containing all instances, called the universe. $C = \{c_1, c_2, c_3, \cdots, c_m\}$ is the set of conditional attributes, $D = \{d_1, d_2, d_3, \cdots, d_k\}$ is the set of decision attributes, $A = C \cup D$ is called the set of attributes, $V$ is the set of attribute values, $f$ specifies the instance mapping relationship between the attribute value and the attribute value.

In order to divide the instance into equivalence classes, you need to first determine the attribute indicators. For attributes $a_i$, if you want to set the $z_i$ level index, you can determine the index width

$$\text{wide}_i = \frac{\max(a_i) - \min(a_i)}{z - 1}$$

(9)

It is thus determined that the indexes at each level are \{min($a_i$), min($a_i$) + wide$_i$, $\cdots$, min($a_i$) + ($z - 1$)wide$_i$\}. For the attribute value $v_i$, it can be determined its index level is

$$j_i = \text{round} \left( \frac{v_i - \min(a_i)}{\text{wide}_i} \right) + 1$$

(10)

For each attribute, the index is determined, and the index can be divided into $\sum_{i=1}^{z_x} z_i$ x equivalent classes. For the instance $x_i = \{v_1, v_2, \cdots, v_{m+1}\}$, the index level corresponding to each attribute is calculated separately, and the instance can be divided into its corresponding equivalence class.

### 3.2 Extracting belief rules

For a decision rule "if X then Y" (or denoted as $X \rightarrow Y$), where $X$ is the set of objects with the same conditional attributes; $Y$ is the set of objects with the same decision attributes. The confidence factor of the rule, or belief level, can be expressed as:

$$cf = \frac{\text{card}(X \cap Y)}{\text{card}(X)} (X \neq \phi)$$

(11)

$\text{card}(X)$ represents the number of elements of the set $X$.

The belief level of the decision rules with the same conditional attributes and different decision attributes is obtained separately, and then these decision rules are combined to form a belief rule in the following form:

If $C$ then \{(d$_1$, cf$_1$), (d$_2$, cf$_2$), $\cdots$, (d$_k$, cf$_k$)\}

(12)

Where $C = \{c_1, c_2, c_3, \cdots, c_m\}$ is the set of conditional attributes.

### 3.3 Index optimization

Index optimization can be achieved by optimizing the index levels. The selection of the attribute index series can be determined by using the control variables and the principle of stepwise determination. Specific steps are as follows:

1. Use the index level of the first attribute as the variable, and select an empirical value as the fixed value for the other attribute index series to determine the indexes at all levels;
2. then use the data set for rule extraction;
3. The cumulative error is obtained by inference of the test set;
4. determining the optimal index level of the first attribute according to the obtained cumulative
error of reasoning:

⑤ The index level of the second attribute is taken as the variable. The optimal index level of the first attribute and the empirical value of the other attribute are used as fixed values to determine the indexes at all levels, and repeat ② to ④ to obtain optimal index level of the second attribute;

⑥ Similar to ⑤, determine the optimal index level of the third attribute;

......

And so on, determine the index levels for all attributes.

4. Experimental verification

4.1. Experimental conditions

In order to verify the effectiveness of the method, an example of oil pipeline leak detection is introduced. The pipeline leak detection problem is to be installed in a 100-kilometer-long oil pipeline in the UK. When a leak occurs in the pipeline, the (FD) and (PD) of the pipeline will change according to a certain pattern, which in turn affects the (LS). A belief rule base system is constructed by taking the FlowDiff and PressureDiff as inputs and the LeakSize as an output.

In the literature, the initial belief rule base system FD has 8 indicators, respectively {-10,-5,-3,-1,0,1,2,3}; PD has 7 indicators, respectively {-0.042,-0.025,-0.01,0,0.01,0.025,0.042}; LS has 5 indicator levels, respectively {0, 2, 4, 6, 8}. The experiment collected real-time data of the 2008 group from normal to 25% leakage every 10 seconds as a cycle. The initial belief rule-base system output is shown in Figure 1.

![Figure 1. Initial confidence rule base output and test set data.](image)

4.2. Experimental process

The framework of the experiment in this paper is shown in Figure 2. First determine the index level, calculate the indexes at each level, and then select 500 sets of data from 7:00 am to 7:33 am, 9:46 am to 10:20 am and 10:50 am to 11:08 am as the data set for rule extraction, and then the 2008 data is used as the test set. The extracted rules are used to perform the belief rule-base reasoning, and the inference cumulative error is obtained. Finally, the index is optimized until the optimal index is obtained, which makes the inference cumulative error minimum.
In the experiment, through the control variables, the index levels of the three attributes are determined step by step. The index of each index is extracted from 2 to 50, and then the test set is used to perform the belief rule-base reasoning. By comparing the final inference cumulative error, the optimal index level of the three attributes is determined to be 10, 3, 10. The calculated index of FD is \{-7.0500, -6.1888, -5.3277, -4.4666, -3.6055, -2.7444, -1.8833, -1.0222, -0.1611, 0.7000\}, and the index of PD is \{-0.0140, 0.0012, 0.0165\}, the indicators of LS are \{0, 0.7131, 1.4262, 2.1393, 2.8525, 3.5656, 4.2787, 4.9918, 5.7050, 6.4181\}. Finally, the extracted rule is used to perform the belief rule-base reasoning. The result is shown in Figure 3.

**Figure 2.** Experimental framework.

**Figure 3.** System output after optimization of indexes and test set data.

In order to reflect the superiority of the method, other methods were used in the experiment to compare with the method. In the previous research, most of the optimization algorithms are used to optimize the rule weight and precondition attribute weight of the belief rule-base based on the initial belief rule-base. In this paper, we use the fmincon in MATLAB to optimize the rule weight and precondition attribute weight of the initial belief rule-base. The optimization result is shown in Figure 4.

Comparing Fig. 1 and Fig. 3, it can be found that the reasoning error of the belief rule-base
extracted by the index optimization is significantly smaller than the initial belief rule-base in Fig. 1, which shows that the method is effective and greatly reduces the inference error. Comparing Fig. 4, the inference error of the proposed method is also smaller than the belief rule-base after rule weight and the preconditioner attribute weight optimization, which indicates that the method has certain advantages over other methods.

Figure 4. System output after rule weight and attribute weight optimization and test set data.

5. Conclusion
In this paper, the difficulty of obtaining expert experience and the inconsistency of standards in the process of establishing the belief rule-base are proposed. A method based on rough set is proposed. Firstly, the data is divided into equivalence classes by the initially determined attribute indexes; then the rules are extracted according to the relevant knowledge of the rough sets; finally, the extracted rules are used to perform the belief rule-base reasoning, and the attribute indexes are optimized according to the inference results. Through experiments, it is found that the method can extract better rules based on the data set, so that the belief rule no longer depends solely on expert experience. However, in the actual process, similar to other optimization methods, the quality of the rule extraction depends on the quality of the selected data set. How to avoid the selection of the "poor" data set requires further research.

References
[1] Yang J B, Liu J, Wang J, et al. Belief rule-base inference methodology using the evidential reasoning Approach-RIMER[J]. IEEE Transactions on Systems Man & Cybernetics Part A Systems & Humans, 2006, 36(2):266-285.
[2] Xu D L, Liu J, Yang J B, et al. Inference and learning methodology of belief-rule-based expert system for pipeline leak detection[J]. Expert Systems with Applications, 2007, 32(1):103-113.
[3] Zhou Z J, Hu C H, Yang J B, et al. Online updating belief rule based system for pipeline leak detection under expert intervention[J]. Expert Systems with Applications, 2009, 36(4):7700-7709.
[4] Zhao, Wang, Shicheng, et al. Real-time fault detection method based on belief rule base for aircraft navigation system[J]. Chinese Journal of Aeronautics (English), 2013, 26(3):717-729.
[5] Feng Z, Zhou Z, Hu C, et al. Fault diagnosis based on belief rule base with considering attribute correlation[J]. IEEE Access, 2017, PP(99):1-1.
[6] Zhou Z G, Liu F, Li L L, et al. A cooperative belief rule based decision support system for lymph node metastasis diagnosis in gastric cancer[J]. Knowledge-Based Systems, 2015, 85(C):62-70.

[7] Solic K, Ocevcic H, Golub M. The information systems' security level assessment model based on an ontology and evidential reasoning approach[M]. Elsevier Advanced Technology Publications, 2015.

[8] Yan X, Xu X, Sheng C, et al. Intelligent wear mode identification system for marine diesel engines based on multi-level belief rule base methodology[J]. Measurement Science & Technology, 2018, 29(1).

[9] Yu H E, Chang L L, Jiang J, et al. Intension Identification in Air Defense Based on Belief Rule Base Expert System under Expert Guidance[J]. Fire Control & Command Control, 2017.

[10] Zhao F J, Zhou Z J, Chang-Hua H U, et al. Aerial Target Intention Recognition Approach Based on Belief-Rule-Base and Evidential Reasoning[J]. Electronics Optics & Control, 2017.

[11] Zhang J, Jiang J, Chen Y, et al. A Belief Rule-Based Safety Evaluation Approach for Complex Systems[J]. Mathematical Problems in Engineering, 2015, 2015:1-10.

[12] Liu J J, Chang-Hua H U, Zhou Z J, et al. Life assessment approach of equipment based on belief-rule-base and evidential reasoning[J]. Control Theory & Applications, 2015.

[13] Long Q, Li Y, Sun Y, et al. An Approach of Converter Transformer Condition Evaluation Based on The Belief Rule Base Inference Methodology and Evidence Reasoning[C]// International Conference on Civil, Transportation and Environment. 2016.

[14] Yang J B, Liu J, Xu D L, et al. Optimization Models for Training Belief-Rule-Based Systems[J]. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 2007, 37(4):569-585.

[15] Qian B, Wang Q Q, Hu R, et al. An effective soft computing technology based on belief-rule-base and particle swarm optimization for tipping paper permeability measurement[J]. Journal of Ambient Intelligence & Humanized Computing, 2017(15):1-10.

[16] Liu J, Martinez L, Da R, et al. Optimization algorithm for learning consistent belief rule-base from examples[J]. Journal of Global Optimization, 2011, 51(2):255-270.

[17] Chang R, Zhang S. An Algorithm for Training Parameters in Belief Rule-bases Based on Gradient Methods with Optimization Step Size[J]. Journal of North China Institute of Water Conservancy & Hydroelectric Power, 2011.

[18] Pawlak Z. Rough set[J]. International Journal of Computer & Information Sciences, 1982, 11(5).

[19] Swiniarzki R W, Skowron A. Rough set methods in feature selection and recognition[J]. Pattern Recognition Letters, 2003, 24(6):833-849.

[20] Saedudin R R, Mahdin H, Kasim S, et al. A Relative Tolerance Relation of Rough Set for Incomplete Information Systems[J]. 2018.

[21] Qian Y, Liang X, Wang Q, et al. Local rough set: A solution to rough data analysis in big data[J]. International Journal of Approximate Reasoning, 2018, 97:38-63.

[22] Yang G J, Yang Y U. Selection of Rough Set and Attribute Reduction Based on BIC[J]. Statistics & Information Forum, 2018.