Belief rule-base expert system with multilayer tree structure for complex problems modeling

Yang, L., Ye, F., Liu, J., & Wang, Y. (2023). Belief rule-base expert system with multilayer tree structure for complex problems modeling. *Expert Systems with Applications*, *217*, Article 119567. https://doi.org/10.1016/j.eswa.2023.119567

Published in:
Expert Systems with Applications

Publication Status:
Published (in print/issue): 01/05/2023

DOI:
10.1016/j.eswa.2023.119567

Document Version
Author Accepted version

Belief Rule-Base Expert System with Multilayer Tree Structure for Complex Problems Modeling

Long-Hao Yang

Decision Sciences Institute, Fuzhou University, Fuzhou, 350108, China

Email: more026@hotmail.com

Fei-Fei Ye*

School of Cultural Tourism and Public Administration, Fujian Normal University, Fuzhou, 350117, China

Email: 13075810934@163.com

Jun Liu

School of Computing, Ulster University at Belfast Campus, York Street, Belfast BT15 1ED, Northern Ireland, UK

Email: j.liu@ulster.ac.uk

Ying-Ming Wang

Decision Sciences Institute, Fuzhou University, Fuzhou, 350108, China

Email: msymwang@hotmail.com

*The corresponding author

Belief Rule-Base Expert System with Multilayer Tree Structure for Complex Problems Modeling

Long-Hao Yang [a], Fei-Fei Ye [b, *], Jun Liu [c], Ying-Ming Wang [a]

[a]Decision Sciences Institute, Fuzhou University, Fuzhou, 350108, China

[b]School of Cultural Tourism and Public Administration, Fujian Normal University, Fuzhou, 350117, China

[c]School of Computing, Ulster University at Belfast Campus, York Street, Belfast BT15 1ED, Northern Ireland, UK

*The corresponding author

Abstract: Belief rule-base (BRB) expert system is one of recognized and fast-growing approaches in the areas of complex problems modeling. However, the conventional BRB has to suffer from the combinatorial explosion problem since the number of rules in BRB expands exponentially with the number of attributes in complex problems, although many alternative techniques have been looked at with the purpose of downsizing BRB. Motivated by this challenge, in this paper, multilayer tree structure (MTS) is introduced for the first time to define hierarchical BRB, also known as MTS-BRB. MTS-BRB is able to overcome the combinatorial explosion problem of the conventional BRB. Thereafter, the additional modeling, inferencing, and learning procedures are proposed to create a self-organized MTS-BRB expert system. To demonstrate the development process and benefits of the MTS-BRB expert system, case studies including benchmark classification datasets and research and development (R&D) project risk assessment have been done. The comparative results showed that, in terms of modelling effectiveness and/or prediction accuracy, MTS-BRB expert system surpasses various existing, as well as traditional fuzzy system-related and machine learning-related methodologies.

Keywords: Belief rule base, multilayer tree structure, expert system, complex problems, combinatorial explosion problem

1. Introduction

Rule-based system is one kind of explainable artificial intelligence (XAI) techniques (Mendel and Bonissone, 2021). It applies rules to store, manage, and manipulate quantitative data and qualitative knowledge for complex problems modeling, where the rules usually take a form of "*IF statements THEN consequents*" and constitute the kernel component of the rule-based system, namely rule-base. Unsurprisingly, a rule-base directly affects the performance of rule-based systems, because the systems response is generated on the basis of the information fusion of IF-THEN rules activated from the rule-base (Sun 1995). For this reason, the modeling of rule-base has attracted lots of attention and researches from the field of information sciences, decision support systems, and computer sciences (Ouyang *et al*., 2019).

Fuzzy rule-base (FRB) is the first typical and popular rule-base, which makes full use of semantic language and fuzzy logic with fuzzy predicates (Sugeno and Yasukawa, 1993). On the basis of the FRB, various kinds of Mamdani-type and Takagi-Sugeno-Kang (TSK)-type fuzzy systems were proposed for complex problems modeling in the past decades (Wang *et al*., 2022; Elkano *et al*., 2016; Fernandez *et al*., 2010; Ishibuchi *et al*., 2005; Cordon *et al*., 1999) and have also been proven the capacity of modeling systems from data, permitting the incorporation of human knowledge, and integrating numerical and semantic reasoning into an explainable scheme (Pedrycz, 1996), where the well-known fuzzy systems include, steady-state generic algorithm for extracting fuzzy classification rule from data (SGREED) (Mansoori *et al*., 2008), fuzzy association rule-based classification method for high-dimensional problem (FARC-HD) (Alcala-Fdez et al., 2011), and deep convolutional fuzzy system (DCFS) (Wang, 2020).

The FRB has an excellent performance in representing vague information. However, a real problem usually coexists various kinds of uncertain information (Cao *et al*., 2021; Aminravan *et al*., 2015). For example, "Project risk is high with a

certainty of 0.9", where the "high" is linguistic expression of project risk, indicating vague information; the "0.9" is numeric expression of project risk, indicating probabilistic information. In order to improve the ability of FRB, the belief rule-base (BRB) was proposed by embedding belief structure into traditional fuzzy IF-THEN rules (Yang *et al*., 2006), and its corresponding model is often referred to as BRB expert system in the researches afterwards (Ahmed *et al*., 2022; Chang *et al*., 2021; Sachan *et al*., 2020; Zhou *et al*., 2019; Zhang *et al*., 2017; Chen *et al*., 2015). Because the BRB inherits many benefits from FRB, the BRB expert system has been successfully applied in handling many complex problems, such as consumer preference prediction (Yang *et al*., 2012), network security prediction (Hu *et al*., 2016), Datacentor PUE prediction (Hossain *et al*., 2017), hidden fault prediction (Zhou *et al.*, 2019), thyroid nodules diagnosis (Chang *et al*., 2022).

However, the conventional BRB has to suffer from the combinatorial explosion problem (AbuDahab *et al*., 2016; Chang *et al*., 2013;). This is because the modeling of a conventional BRB is required to cover all possible combinations of each assessment rating for all antecedent attributes. For example, the problem of R&D project risk assessment involves M risk factors with J assessment ratings, the resulting conventional BRB would have J^M rules, *e.g.*, 3^{20}=3.4 billion rules for 20 risk factors and 3 assessment ratings. Moreover, it is evident that the size of the conventional BRB grows exponentially along with that of risk factors and assessment ratings (Zhou *et al*., 2020; Diao *et al*., 2022). Although many studies have been done (The detailed literatures can be found in ***Section* 2.2**) to improve the modeling of conventional BRB in the past decade, the combinatorial explosion problem of conventional BRB is still a great challenge when BRB expert system is applied for modeling complex problems.

From a general survey of the existing studies detailed in ***Section* 2.2**, it seems that hierarchical BRB may provide an effective and feasible solution to solve the combinatorial explosion problem of BRB when modeling complex problems. Here, it worth noting that, apart from hierarchical BRB, the application and modeling of hierarchical FRB attract ongoing development in the literature, recent studies include proposing hierarchical deep rule-based classifier (Gu and Angelov, 2020), the extraction of hierarchical TSK fuzzy systems from data (Kerr-Wilson and Pedrycz, 2020), designing granular fuzzy models using hierarchical approach (Pedrycz *et al*., 2015), and the application of hierarchical fuzzy system to traffic congestion prediction (Zhang *et al*., 2014). However, none of existing studies on hierarchical BRB and FRB clearly define a generic representation scheme of hierarchical rule-base, which is the basic foundation of constructing hierarchical rule-base from data for modeling complex problems. Moreover, owing to the visible differences between FRB and BRB in terms of rule representation and reasoning, these limit the use of existing hierarchical FRB modeling procedures directly to construct hierarchical BRB. Hence, it is necessary and valuable to study hierarchical modeling procedure for BRB.

In order to fill the gap mentioned above, this study aims to provide a generic representation scheme of hierarchical BRB using multilayer tree structure (MTS), where the new BRB is referred to as MTS-BRB. On the basis of MTS, it is convenient to illustrate the size reduction of MTS-BRB and its ability to overcome the combinatorial explosion problem of conventional BRB. Furthermore, the relevant MTS-BRB modeling, inferencing and learning procedures are also proposed to provide the necessity elements of an expert system, including how to construct a MTS-BRB from complex problems and reply a given input data through the MTS-BRB? What kinds of parameters need to be trained for optimizing MTS-BRB? Finally, a novel expert system, called MTS-BRB expert system, is developed for complex problems modeling.

To verify the effectiveness of the proposed MTS-BRB expert system, the research and development (R&D) projects risk assessment from Chinese industries and the four classification datasets from the well-known University of California, Irvine (UCI) database are introduced to conduct case studies. The details of how to apply the proposed MTS-BRB modeling,

inferencing, and learning procedures are illustrated based on the R&D projects risk assessment. Several comparative studies are carried out to justify excellent computing efficiency and prediction accuracy of the MTS-BRB expert system better than some existing BRB expert systems, typical fuzzy systems, and classical machine learning methods.

The original contributions of this study can be summarized as follows:

(1) MTS is used to represent the hierarchical BRB for the first time, which can be used to handle complex problems modeling and overcome the combinatorial explosion problem of conventional BRB.

(2) An effective MTS-BRB modeling procedure is proposed and it allows that the complex problems can be handled by the self-organized hierarchical BRB, instead of human-made hierarchical BRB.

(3) An effective MTS-BRB inferencing procedure is proposed and it provides a generic process of how to produce an inferential output from given input data in a hierarchical BRB for the first time.

(4) An effective MTS-BRB learning procedure is proposed and it is able to provide optimal parameters for MTS-BRB expert system without restricting by limited expert knowledge in complex problems.

The remainder of this paper is organized by: ***Section* 2** briefly reviews the background of BRB expert system. ***Section* 3** gives the definition and analysis of hierarchical BRB using MTS. ***Section* 4** proposes a MTS-BRB expert system. The case studies are carried out for verification in ***Section* 5**, and finally ***Section* 6** concludes the study.

2. Background of the BRB Expert System

In this section, the relevant basics of this study, including the conventional BRB and its combinatorial explosion problem in ***Section* 2.1** and the related works to overcome the problem in ***Section* 2.2**, are presented.

2.1. Conventional BRB and its combinatorial explosion problem

BRB is a rule base of BRB expert system and is comprised of a series of belief rules (Yang *et al*., 2006). Normally, the *k*th belief rule in a conventional BRB can be written as:

$$\begin{aligned} &R_k : IF\ U_1\ is\ A_1^k \wedge \cdots \wedge U_M\ is\ A_M^k, THEN\ D\ is\ \{(D_n, \beta_{n,k}); n = 1,..., N\}, \\ &\quad with\ rule\ weight\ \theta_k\ and\ attribute\ \ weights\ \{\delta_1,..., \delta_M\} \end{aligned} \tag{1}$$

where $\{U_m; m=1,\ldots, M\}$ denotes a set of M antecedent attributes used in IF part; D denotes one consequent attribute used in THEN part; $\{A_m^k; m=1,\ldots, M\}$ denotes a set of assessment ratings used to describe the kth (k=1,…, L) belief rule, L is a total number of belief rules in the BRB. Noting that A_m^k belongs to $\{A_{m,j}; j=1,\ldots, J_m\}$ that is a complete set of J_m assessment ratings used to describe the mth antecedent attribute; $\{(D_n, \beta_{n,k}); n=1,\ldots, N\}$ denotes a belief distribution in consequent attribute D, where $\beta_{n,k}$ is a belief degree to which the consequent D_n is believed to be true.

To illustrate conventional BRB in details, suppose that a decision problem with M+1 attributes, denoted by $U_1,\ldots, U_M$, D, is provided to construct a conventional BRB, where attribute D is associated with M attributes $U_1,\ldots, U_M$ simultaneously. Here, it needs to note that the data of $U_1,\ldots, U_M$, D can be continuous and discrete data. More detailed, taking R&D project risk assessment (Yang *et al*., 2019) as the decision problem for example, the risk of a project (namely D) is determined by business cooperation (namely U_1) and company scale (namely U_2) and their data, denoted by $x(U_1)$, $x(U_2)$, and $y(D)$, are all continuous data within ranges $x(U_1)\in[1.33, 5.0]$, $x(U_2)\in[2.0, 5.0]$, and $y(D)\in[2.0, 5.0]$. When there are three assessment ratings *e.g.*, $\{A_{1,1}, A_{1,2}, A_{1,3}\}=\{A_{2,1}, A_{2,2}, A_{2,3}\}=\{Low, Middle, High\}$, provided for U_1 and U_2, and three consequents, *e.g.*, $\{D_1, D_2, D_3\}=\{Small, Middle, High\}$, provided for D by domain expert, a conventional BRB can be therefore constructed by covering all possible combinations of each assessment rating for all antecedent attributes, as shown in Table 1.

Table 1. Example of a conventional BRB regarding R&D project risk assessment

Rule No.	Rule weight	Business cooperation U_1 (δ_1=0.8)	And	Company scale U_2 (δ_2=0.9)	Project risk D		
					D_1=*Small*	D_2=*Middle*	D_3=*High*
R_1	θ_1=0.9	A_1^1=*Low*	∧	A_2^1=*Low*	$\beta_{1,1}$=0.6	$\beta_{2,1}$=0.3	$\beta_{3,1}$=0.0
R_2	θ_2=0.5	A_1^2=*Low*	∧	A_2^2=*Middle*	$\beta_{1,2}$=0.5	$\beta_{2,2}$=0.5	$\beta_{3,2}$=0.0
R_3	θ_3=0.6	A_1^3=*Low*	∧	A_2^3=*High*	$\beta_{1,3}$=0.7	$\beta_{2,3}$=0.3	$\beta_{3,3}$=0.0
R_4	θ_4=0.7	A_1^4=*Middle*	∧	A_2^4=*Low*	$\beta_{1,4}$=0.3	$\beta_{2,4}$=0.5	$\beta_{3,4}$=0.2
R_5	θ_5=0.8	A_1^5=*Middle*	∧	A_2^5=*Middle*	$\beta_{1,5}$=0.2	$\beta_{2,5}$=0.5	$\beta_{3,5}$=0.3
R_6	θ_6=0.4	A_1^6=*Middle*	∧	A_2^6=*High*	$\beta_{1,6}$=0.1	$\beta_{2,6}$=0.9	$\beta_{3,6}$=0.0
R_7	θ_7=0.3	A_1^7=*High*	∧	A_2^7=*Low*	$\beta_{1,7}$=0.0	$\beta_{2,7}$=0.3	$\beta_{3,7}$=0.7
R_8	θ_8=0.9	A_1^8=*High*	∧	A_2^8=*Middle*	$\beta_{1,8}$=0.0	$\beta_{2,8}$=0.2	$\beta_{3,8}$=0.8
R_9	θ_9=1.0	A_1^9=*High*	∧	A_2^9=*High*	$\beta_{1,9}$=0.0	$\beta_{2,9}$=0.1	$\beta_{3,9}$=0.9

From Table 1, it can be found that a belief rule such as R_1 contains that when business cooperation is *Low* and company scale is *Low*, then 60% sure that project risk is Small and 30% is *Middle*. Meanwhile, some other information should be highlighted: 1) the weight of R_1 is 0.9, indicating the importance of R_1 over other rules; 2) the weight of business cooperation and company scale is 0.8 and 0.9, respectively, indicating the different importance of these two attributes; 3) the total belief degree of project risk is 60% + 30% = 90% < 100%, indicating that R_1 contains 10% incomplete uncertainty.

Here, it is worth noting that a conventional BRB has to enumerate all combinations of each assessment rating. Hence, when there are M antecedent attributes in a decision problem and J_i assessment ratings used for each attribute, the size of conventional BRB is $\prod_{i=1}^{M} J_i$. From this situation, the following problem can be observed:

***Problem* 1 (*Combinatorial explosion problem of BRB*)**: The size of conventional BRB has exponential-relation with the number of antecedent attributes and/or their assessment ratings. Moreover, the size of the conventional BRB would grow exponentially when the number of antecedent qualities and/or assessment ratings increased.

2.2. Related works to overcome the problem of conventional BRB

The following previous investigations, which can be grouped into four categories, have been conducted in the past decade in an effort to solve the combinatorial explosion problem of conventional BRB:

(1) BRB modeling with feature extraction. Feature extraction is one of the well-known techniques in machine learning for dimensionality reduction and its fundamental is the transformation of original data to a data set with a reduced number of features, which contains the most discriminatory information. In this respect, Wang *et al*. (2009) utilized a BRB expert system to predict the consumer preference of orange juices that are usually involved with many sensory attributes, while they extracted the first two or three principal components (PCs) from sensory attributes to avoid constructing an overlarge BRB. Later, Yang *et al*. (2012) also proposed extracting the first two or three PCs of relevant sensory attributes to construct BRB with an acceptable size when a BRB expert system is used for consumer preference prediction in retro-fit design of food and drink product. Afterwards, Yang *et al*. (2016) incorporated factor analysis into BRB expert system for predicting consumer preference of new products. They extracted a few key factors from product attributes based on both exploratory and confirmatory factor analysis. Recently, Cheng *et al*. (2020) proposed a data-driven health monitoring method for running gears of a high-speed train by using principal component analysis (PCA) method to extract PCs and BRB expert

system to perform health monitoring tasks. Zhang *et al*. (2021) ranked the original attributes based on eigenvalue decomposition and contribution rate and transformed them into attribute vectors, so the resulting BRB expert system could overcome the issue caused by overnumbered attributes.

(2) BRB modeling with feature selection. Feature selection is another well-known technique in machine learning for dimensionality reduction. The difference between feature selection and feature extraction is that the former one is to search for an optimal subset of original features. In this respect, Chang *et al*. (2013) introduced a concept "structure learning" for solving combinatorial explosion challenge. They proposed four kinds of structure learning approaches using grey target (GT), multidimensional scaling (MDS), Isomap, and PCA. Later, some similar studies based on conditional generalized minimum variance (Li *et al*., 2017) were carried out for selecting key attributes from numerous attributes. Yang *et al*. (2019) applied information gain (IG) to calculate the weight of all attributes, which were regarded as probability to randomly select subset of original attributes, so that the BRB expert system can be used to evaluate project performance of R&D. Recently, the idea of feature selection and ensemble learning were used together to construct BRB expert system under the situation of lots of attributes (Yang *et al*., 2020). Another similar study can be found in (You *et al*., 2021), they proposed an ensemble -BRB model with the use of bagging framework to reduce the size of BRB.

(3) Extended BRB modeling. For completely solving the combinatorial explosion problem, several extensions were proposed in the term of rule generation to improve the conventional BRB modeling. In this respect, Liu *et al*. (2013) proposed an extended BRB (EBRB) modeling method, which is based on the transformation of data into rules. A similar study was proposed in (Jiao *et al*., 2016). Afterwards, the EBRB modeling was extended by Yang *et al*. (2021) to cluster all rules into division domains, so that an EBRB did not have an ever-increasing size due to the increase of available data. Chang *et al*. (2019) developed a BRB modeling method under disjunctive assumption, namely DBRB modeling, whose essential modifies the logical relationship between attributes using disjunctive assumption to replace conjunctive assumption. Cao *et al*. (2021) proposed a new approximate belief rule with single attributes to solve the rule explosion problem and week extendability of BRB expert system and its effectiveness was confirmed by using a case study of the Lithium-ion power battery. In the viewpoint of BRB size, the above extensions break the precondition of the combinatorial explosion problem, so the BRB expert system has been successfully applied in the classification problems with high dimension (Fang *et al*., 2020; Gao *et al*., 2021; Zhuang *et al*., 2021). However, these new modeling methods lose the completeness of rule representation, which is a core characteristic of using BRB to represent expert knowledge and historical data.

(4) Human-made hierarchical BRB. This kind of solution is usually based on expert knowledge to build a top-down structure of BRB so that all attributes can be considered in a BRB construction. In this respect, Yang *et al*. (2006) first provided a rough description of hierarchical BRB construction and its example using prior information when the BRB expert system is used to address complex systems with 9 attributes. Later, Yang *et al*. (2012) adopted the hierarchical BRB for quality evaluation of lemonade products, which has a three-level top-down structure of BRB according to expert knowledge. Zhou *et al*. (2015) proposed a bi-level BRB for the clinical decision support system. By considering both clinical data and specific domain knowledge, the second layer BRB was constructed to distinguish the overlapped and fuzzy information. Afterwards, He *et al*. (2018) proposed a reliability evaluation method based on hierarchical BRB expert system for the field of wireless sensor network, in which the hierarchical BRB expert system only focused on the hierarchical inferencing and optimization procedures in a given hierarchical BRB. Recently, Sachan *et al*. (2020) presented a BRB-based

decision-support-system to automate loan underwriting process. The hierarchical structure of BRB was given as prior knowledge to avoid the situation that the number of rules increases exponentially with the numbers of antecedent attributes and their referential values.

From the above-mentioned four kinds of previous studies, the combinatorial explosion problem is one of the biggest obstacles in the BRB modeling when a BRB expert system is applied to dealing with complex system modeling problems. However, it seems that the hierarchical BRB can be a potential solution to address the combinatorial explosion problem, because a hierarchical BRB is required to cover the assessment ratings of a part of antecedent attributes, instead of all attributes. Hence, the present study focuses on how to develop a BRB expert system using the hierarchical BRB.

3. Using MTS to Represent the Hierarchical BRB

In this section, the definition of hierarchical BRB using MTS, called MTS-BRB, is firstly provided in ***Section* 3.1**, followed by the comparison of different MTS-BRBs in ***Section* 3.2** to show the necessity of this study.

3.1. Definition of hierarchical BRB using MTS

On the basis of the previous studies shown in ***Section* 2.2**, the hierarchical BRB can be a potential solution to address the combinatorial explosion problem. However, the representation of a hierarchical BRB is still lack of a definitive scheme. Hence, in this subsection, the multi-layer tree structure is introduced to provide a new definition of hierarchical BRB.

Firstly, based on the usual representation of a tree structure, there are three kinds of nodes, namely *root node*, *internal node*, and *leaf node*. Hence, the definitions of these three nodes are provided as follows:

***Definition* 1**: In a tree structure, the node that has no parent node is called *root node*.

***Definition* 2**: In a tree structure, the node that has no child node is called *leaf node*.

***Definition* 3**: In a tree structure, the node that has both parent node and child node is called *internal node*.

Furthermore, the definition regarding the number of layers in a tree structure is provided as follows:

***Definition* 4**: *The number of layers* in a tree structure is the length of the longest path from the top internal node to the bottom internal nodes

Based on ***Definitions* 1** to **4**, the conventional BRB shown in ***Section* 2.1** can be represented as a tree structure, as shown in Fig. 1, where the BRB has M antecedent attributes U_m (m=1,…, M) and one consequent attribute D.

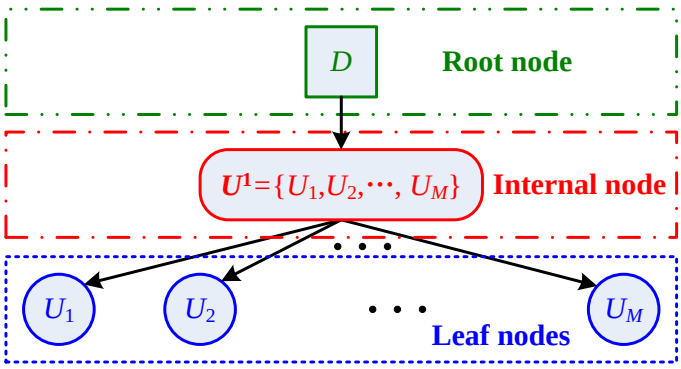


Fig. 1. Tree structure of conventional BRB

From Fig. 1, it can be found that the tree structure of conventional BRB is a single-layer tree structure and has one root node, one internal node, and M leaf nodes, where the all M leaf nodes are the child node of the internal node. This corresponds to the situation that the all M antecedent attributes should be used together to construct a conventional BRB.

Secondly, in order to avoid the above situation, the all M antecedent attributes can be packaged into multiple subsets

e.g., $\{U_1,\ldots, U_{m1}\}$, $\{U_{m1+1},\ldots, U_{m2}\}$, and $\{U_{m2+1},\ldots, U_M\}$ ($1<m_1<m_2<M$), for constructing multiple downsized sub-BRBs. Moreover, in the any case of the sub-BRB being still oversize, the subset of attributes can be recursively packaged into smaller subsets until the new sub-BRB has a desired size. As a result, all these sub-BRBs forms a hierarchical BRB and the relationship between subsets constructs a multi-layer tree structure. Without of loss generality, assume that the multi-layer tree structure is the tree structure shown in Fig. 2.

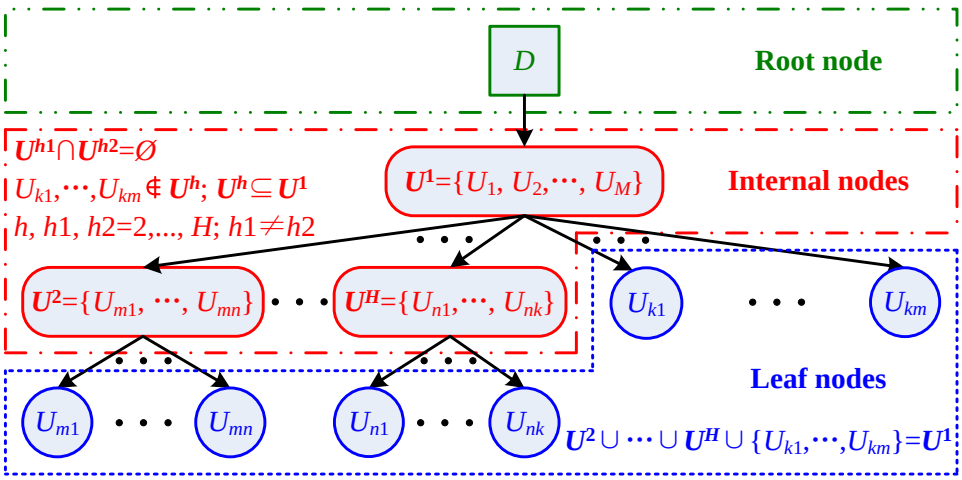


Fig. 2. Tree structure of hierarchical BRB

From Fig. 2, it is clear that the tree structure of hierarchical BRB is a two-layer tree structure and, apart from M leaf nodes and one root node, it has H internal nodes, which consist of a subset of the M antecedent attributes. In this case, a new definition based on ***Definitions* 1** to **4** is offered to define MTS-BRB as follows:

***Definition* 5 (MTS-BRB)**: Assume that a tree structure of hierarchical BRB contains one root node, H internal nodes, and M leaf nodes, where all these nodes should meet the following requirements:

(1) The root node, internal node, and leaf node are corresponding to consequent attribute, subset of antecedent attributes, and antecedent attribute, respectively, *i.e.*, consequent attribute D is root node, subsets of antecedent attributes $\boldsymbol{U^h}$ (h=1,…, H) are internal nodes, and antecedent attributes U_m (m=1,…, M) are leaf nodes in Fig. 2.

(2) The internal nodes are located on two layers at least and they satisfy: 1) the top internal node includes all antecedent attributes, *e.g.*, $\boldsymbol{U^1}=\{U_1,\ldots, U_M\}$ in Fig. 2; 2) the internal node in a lower layer is a subset of the internal node in an upper layer, *e.g.*, $\boldsymbol{U^h}\subseteq\boldsymbol{U^1}$ (h=2,…, H) in Fig. 2.

(3) The internal node is corresponding to sub-BRB. Taking the hth (h=1,…, H) internal node $\boldsymbol{U^h}$ as an example, assume that it has N_h consequents $\{D_{h,n}; n=1,\ldots, N_s\}$ and M_h child nodes $\boldsymbol{U_{h,m}}$ (m=1,…, M_h) with $J_{h,m}$ assessment ratings $\{A_{h,m,j}; j=1,\ldots, J_{h,m}\}$. Thus, the kth (k=1,…, L_h) belief rule in the sub-BRB for the hth internal node can be written as:

$$\begin{aligned} &R_{h,k}: IF\ \boldsymbol{U_{h,1}}\ is\ A_1^{h,k}\wedge\boldsymbol{U_{h,2}}\ is\ A_2^{h,k}\wedge\cdots\wedge\boldsymbol{U_{h,M_h}}\ is\ A_{M_h}^{h,k}, THEN\ \boldsymbol{U^h}\ is\ \{(D_{h,n},\beta_{h,n,k}); n=1,\ldots, N_h\}, \\ &\quad with\ rule\ weight\ \theta_{h,k}\ and\ attribute\ \ weights\ \{\delta_{h,1},\ldots,\delta_{h,M_h}\} \end{aligned} \tag{2}$$

where $A_m^{h,k}\in\{A_{h,m,j}; j=1,\ldots, J_{h,m}\}$; $\delta_{h,m}$ denotes the weight of the mth child node of the hth internal node; $\theta_{h,k}$ denotes the weight of the kth rule in the sub-BRB for the hth internal node. The above defined tree structure of hierarchical BRB is called a MTS-BRB.

Here, it is worth noting from ***Definition* 5** that the child node is an important element in each sub-BRB and it can be embodied into internal nodes and/or leaf nodes. As shown in Fig. 2, the child nodes of $\boldsymbol{U^1}$ are H-1 internal nodes $\boldsymbol{U^h}$ (h=2,…, H) and a series of leaf nodes $U_{k1},\cdots, U_{km}$; the child nodes of $\boldsymbol{U^2}$ are a series of leaf nodes $U_{m1},\cdots, U_{mn}$.

3.2. Size of BRBs under different numbers of attributes

To facilitate the application of MTS-BRB, the size of different BRBs needs to be investigated. For this purpose, the following three examples are provided to carry out comparative studies.

***Example* 1**: In the case of the conventional BRB detailed in ***Section* 2.1**, when there are M antecedent attributes and 3 assessment ratings for each attribute, the resulting BRBs has 3^M rules. Furthermore, by using a certain feature selection or extraction technique to retain 30%, 50%, and 80% attributes, the resulting three BRBs has $3^{0.3\times M}$, $3^{0.5\times M}$, and $3^{0.8\times M}$ rules, respectively, as shown in Fig. 3.

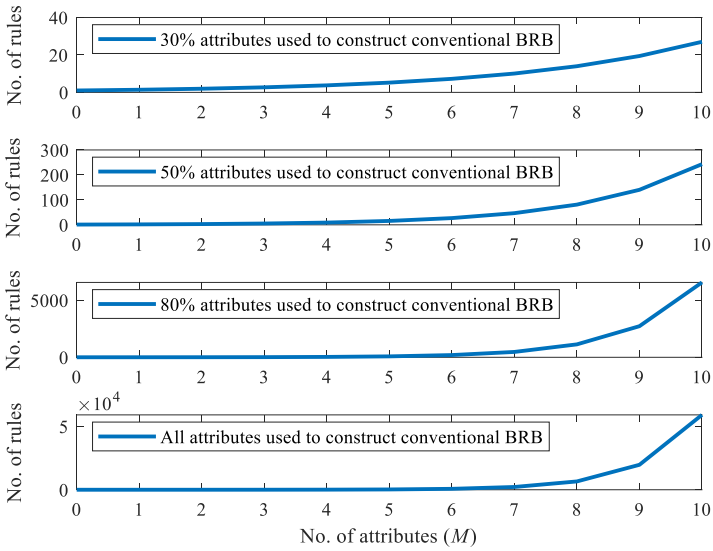


Fig. 3. Comparison of conventional BRBs with feature selection or extraction

From Fig. 3, it is clear that feature selection or extraction techniques can downsize BRB, *e.g.*, 3^{10}=59049 rules for 10 antecedent attributes, and only $3^{0.8\times10}$=6561, $3^{0.5\times10}$=243, and $3^{0.3\times10}$=27 rules while retaining 80%, 50%, and 30% attributes, respectively. However, the size of downsized BRB still grows exponentially along with the increase of attributes. Hence, it can be concluded that feature selection or extraction techniques are able to reduce the size of a BRB, but they are unable to be an effective approach to solve the combinatorial explosion problem.

***Example* 2**: From (Wang *et al*., 2009), the problem of consumer preferences for orange juices can be related with 9 features, including color, bitty, citrus, marmalade, body, sweet, sour, mouthcoating, and astringent. For this problem, two kinds of BRBs can be constructed using the tree structure shown in Fig. 4 and Fig. 5, where Fig. 4 is corresponding to the conventional BRB detailed in ***Section* 2.1**, and Fig. 5 is corresponding to a MTS-BRB.

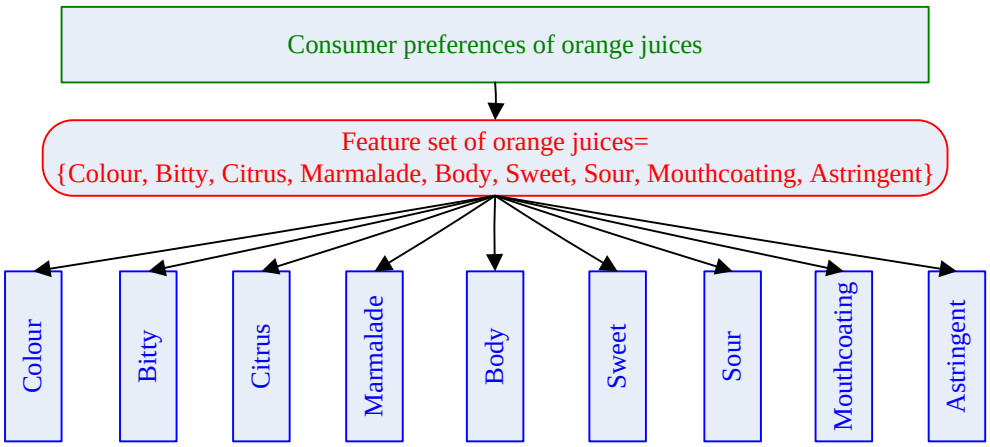


Fig. 4. Tree structure of conventional BRB for consumer preferences of orange juices

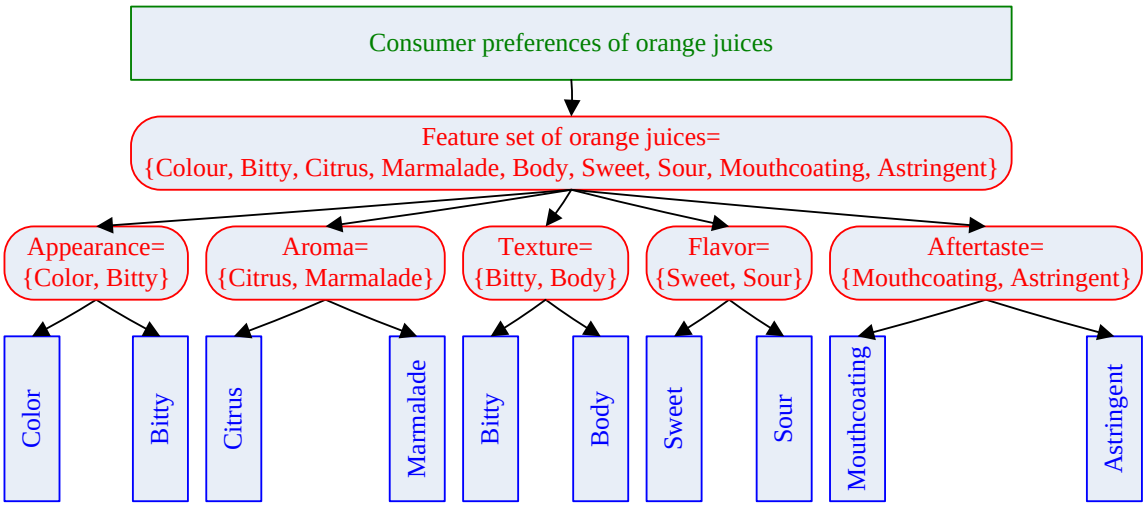
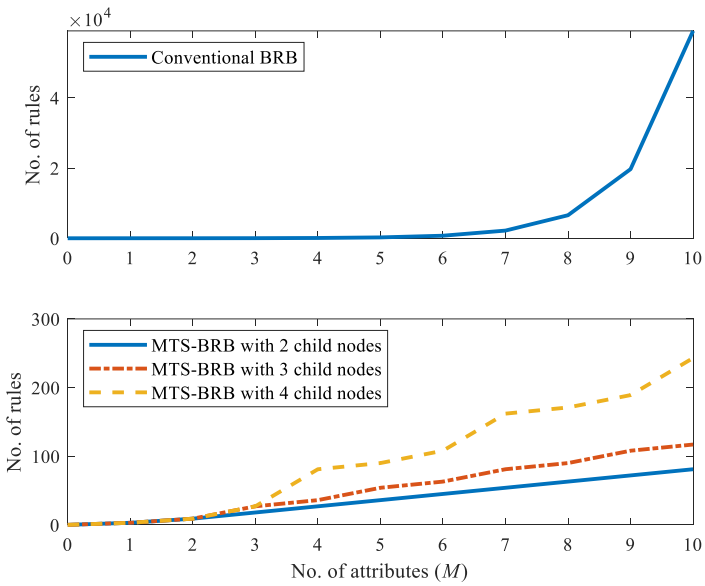


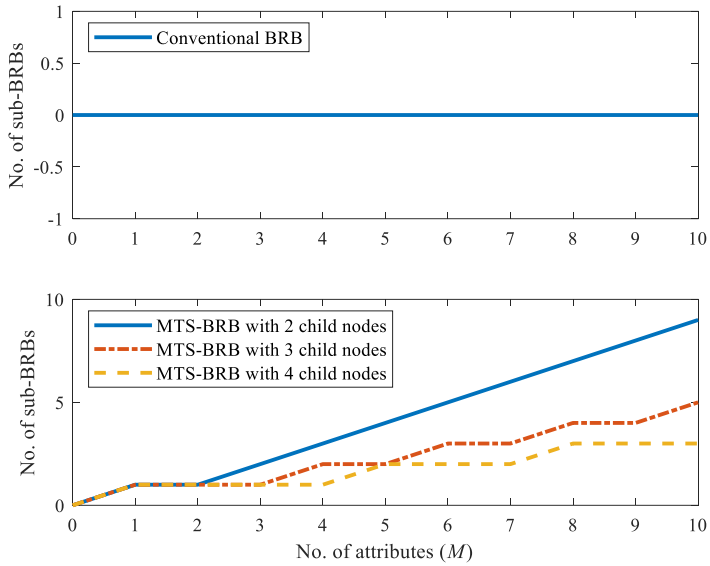
Fig. 5. Tree structure of hierarchical BRB for consumer preferences of orange juices

From Fig. 4 and Fig. 5, it is clear that the size of conventional BRB is 3^9=19683 rules when each feature is evaluated by three ratings. But there are only $3^5+3^2+3^2+3^2+3^2+3^2$=288 rules in the MTS-BRB when the 9 features are packaged into 6 subsets, *i.e.*, color and bitty are packaged into a subset, indicating appearance, and the resulting sub-BRB has 3^2=9 rules, and the other 5 subsets regarding aroma, texture, flavor, aftertaste, and consumer preference can be also used to construct 5 sub-BRBs with 3^2=9, 3^2=9, 3^2=9, 3^2=9, and 3^5=243 rules.

***Example* 3**: In order to give a more intuitive comparison, assume that there are M antecedent attributes and 3 assessment ratings for each attribute to construct the four kinds of BRBs, including the conventional BRB and the MTS-BRB with 2/3/4 child nodes. As a result, Fig. 6 shows the number of rules and sub-BRBs in these BRBs, in which the number of child nodes means the maximum number of child nodes for each internal node.



(a) Comparison of number of rules in the conventional BRB and MTS-BRBs



(b) Comparison of number of sub-BRBs in the conventional BRB and MTS-BRBs

Fig. 6. Comparison of different BRBs with different numbers of attributes

From Fig. 6, it is clear that the size of the MTS-BRBs with 2 (or 3 or 4) child nodes almost grows linearly along with the increase of attributes and the number of rules in the MTS-BRBs is far less than that in the conventional BRB, *e.g.*, in the case of 10 attributes, the conventional BRB has 3^{10}=59049 rules, but MTS-BRB with 2 (or 3 or 4) child nodes has only 81, 117, and 243 rules. This is because the all 10 attributes are used to construct a single BRB for conventional BRB, but 9 (or 5 or 3) sub-BRBs for MTS-BRBs with 2 (or 3 or 4) child nodes, respectively. As a result, in terms of BRB size, it can be concluded that MTS-BRB is a preferred method for resolving the combinatorial explosion problem of BRB. This method is superior than feature selection and extraction methods.

4. Proposed MTS-BRB Modeling, Inferencing, and Learning Procedures

In this section, the novel procedures of MTS-BRB modeling, inferencing, and learning are proposed firstly in ***Sections* 4.1** to **4.3**, respectively. Next, on the basis of the three procedures, the MTS-BRB expert system is developed in ***Section* 4.4** for modeling complex systems problems and solving the combinatorial explosion problem.

4.1. MTS-BRB modeling procedure

According to definition of MTS-BRB in ***Section* 3.1**, multiple subsets of antecedent attributes should be separated from upper internal nodes to downsize the size of sub-BRB. For this separation operation, one of the most important things is how to put each antecedent attribute in a right subset. Hence, in this subsection, hierarchical attribute clustering is applied to achieve the separation operation. However, despite that various types of clustering analyses exist in (Aggarwa and Reddy, 2013), this study is particularly interested in hierarchical clustering analysis, because it creates clusters in a hierarchical tree structure similar to the three structure of MTS-BRB.

First of all, before performing hierarchical attribute clustering, it is necessary to investigate the way of quantifying correlations among antecedent attributes. From the existing study (Li *et al.*, 2020), the entropy and mutual information are commonly used indicators to measure interdependence information, which is one of the most significant characteristics of inter-relationship among antecedent attributes. To define the relationship of antecedent attributes in light of this argument, the entropy and mutual information are introduced as follows:

***Definition* 6**: Assume that a MTS-BRB has M antecedent attributes U_m (m=1,…, M) and J_m assessment ratings {$A_{m,j}$;

j=1,…, J_m} for each attribute. Based on the entropy and mutual information (MacKay, 2003), *the attribute relation* $AR(U_m, U_n)$ between two attributes U_m and U_n (n=1,…, M) can be quantified by:

$$AR(U_m,U_n)=\frac{I(U_m,U_n)}{H(U_m,U_n)} \tag{3}$$

where $H(U_m, U_n)$ and $I(U_m, U_n)$ denote the entropy and the mutual information of U_m and U_n, and they can be calculated by:

$$H(U_m,U_n)=-\sum_{i=1}^{J_m}\sum_{j=1}^{J_n}P_{m,n}(U_m \text{ is } A_{m,i}\wedge U_n \text{ is } A_{n,j})\times\log P_{m,n}(U_m \text{ is } A_{m,i}\wedge U_n \text{ is } A_{n,j}) \tag{4}$$

$$I(U_m,U_n)=\sum_{i=1}^{J_m}\sum_{j=1}^{J_n}P_{m,n}(U_m \text{ is } A_{m,i}\wedge U_n \text{ is } A_{n,j})\times\log\frac{P_{m,n}(U_m \text{ is } A_{m,i}\wedge U_n \text{ is } A_{n,j})}{P_m(U_m \text{ is } A_{m,i})\times P_n(U_n \text{ is } A_{n,j})} \tag{5}$$

where $P_{m,n}(U_m \text{ is } A_{m,i}\wedge U_n \text{ is } A_{m,j})$ denotes the probability of U_m and U_n being equal to assessment ratings $A_{m,i}$ and $A_{n,j}$ simultaneously; $P_m(U_m \text{ is } A_{m,i})$ denotes the probability of U_m being equal to assessment rating $A_{m,i}$.

Based on the MTS-BRB shown in ***Definition* 5** and the relation defined in ***Definition* 6**, the MTS-BRB modeling procedure is presented in Fig. 7. Here, it should be noted that each part of tree structure can be used to construct a sub-BRB by using the upper-layer node as consequent attribute and the lower-layer nodes as antecedent attributes. Hence, the number of lower-layer nodes is directly related to the size of sub-BRB.

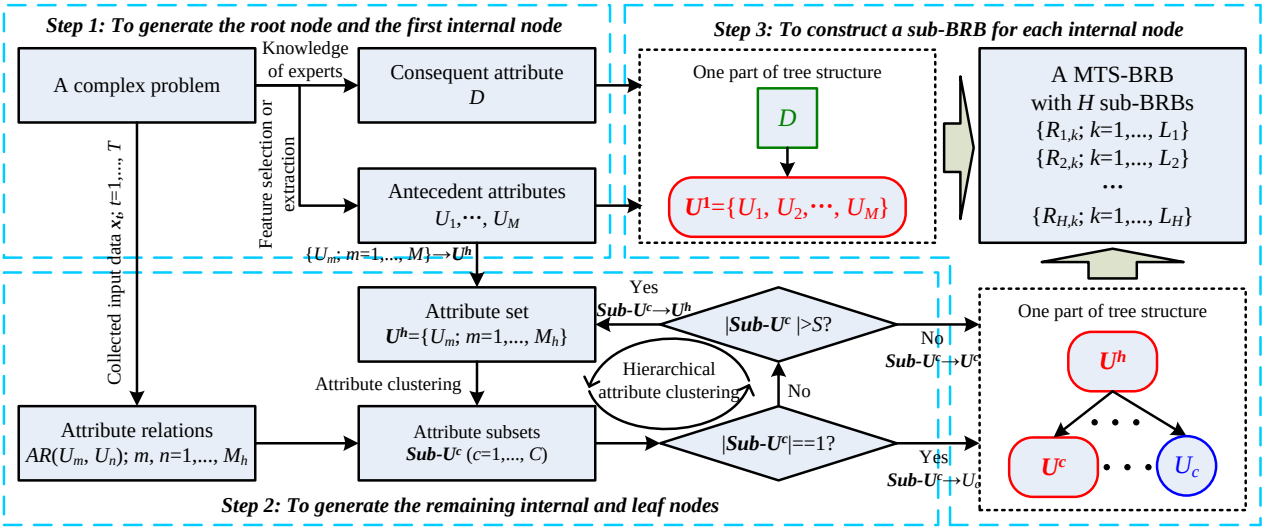


Fig. 7. Illustration of MTS-BRB modeling procedure

From Fig. 7, three steps of the MTS-BRB modeling procedure are provided as follows:

***Step* 1**: To generate the root node and the first internal node of the MTS-BRB.

Based on ***Definition* 5**, the root node and the first internal node of MTS-BRB are corresponding to consequent attribute and all antecedent attributes. However, due to the fact that the initial antecedent attributes sometimes contain redundant and irrelevant information, feature selection or extraction techniques should be used to filter initial attributes. To facilitate description, the consequent attribute is denoted as D and the filtered antecedent attributes are denoted as $\boldsymbol{U}^1=\{U_1,\ldots, U_M\}$.

***Step* 2**: To generate the remaining internal and leaf nodes of the MTS-BRB.

In order to determine a top-down relationship among internal nodes and leaf nodes, a process of hierarchical attribute clustering is developed into the following two sub-steps:

***Step* 2.1**: To calculate attribute relations for antecedent attributes.

Suppose that each antecedent attribute U_m (m=1,…, M) has assessment ratings $A_{m,j}$ (j=1,…, J_m) and their utility values $u(A_{m,j})$ (j=1,…, J_m). When a set of T input data $\boldsymbol{x}_t=(x_{t,1},\ldots, x_{t,M})$ (t=1,…, T) is given, the following belief distributions can be

obtained using the utility-based equivalence transformation technique:

$$S(\boldsymbol{x}_t, U_m) = \{(A_{m,j}, \alpha_{m,j}^t); j = 1,..., J_m\} \tag{6}$$

where $\alpha_{m,j}^t$ denotes the belief degree of assessment rating $A_{m,j}$ at antecedent attribute U_m and it can be calculated by:

$$\alpha_{m,j}^t = \frac{u(A_{m,j+1}) - x_{t,m}}{u(A_{m,j+1}) - u(A_{m,j})} \text{ and } \alpha_{m,j+1}^t = 1 - \alpha_{m,j}^t, \text{ if } u(A_{m,j}) \le x_{t,m} \le u(A_{m,j+1}) \tag{7}$$

$$\alpha_{m,k}^t = 0, \text{ for } k = 1,..., J_m \text{ and } k \ne j, j+1 \tag{8}$$

Based on $T \times M$ belief distributions $S(\boldsymbol{x}_t, U_m)$ (t=1,…, T; m=1,…, M), attribute relation $AR(U_m, U_n)$ between antecedent attributes U_m and U_n (n=1,…, M) can be obtained according to ***Definition* 6**, in which the probabilities $P_{m,n}(U_m$ is $A_{m,i} \wedge U_n$ is $A_{m,j})$ and $P_m(U_m$ is $A_{m,i})$ are calculated by:

$$P_{m,n}(U_m \text{ is } A_{m,i} \wedge U_n \text{ is } A_{n,j}) = \frac{\sum_{t=1}^{T} \alpha_{m,i}^t \times \alpha_{n,j}^t}{T} \tag{9}$$

$$P_m(U_m \text{ is } A_{m,i}) = \frac{\sum_{t=1}^{T} \alpha_{m,i}^t}{T} \tag{10}$$

***Step* 2.2**: To hierarchically cluster antecedent attributes based on attribute relations.

Suppose that the set of antecedent attributes is denoted as $\boldsymbol{U^h}$. According to the given number of clusters C ($M \ge C \ge 2$), C pivot factors P_c (c=2,…, C) should be selected from $\boldsymbol{U^h}$ by

$$P_c = \begin{cases} \arg\min_{U_m \in \boldsymbol{U^h}} \{\sum_{U_n}^{\boldsymbol{U^h}} AR(U_m, U_n)\}, \textit{if } c = 1 \\ \arg\min_{U_m \in \boldsymbol{U^h} - \{P_1,...,P_{c-1}\}} \{\sum_{U_n}^{\{P_1,...,P_{c-1}\}} AR(U_m, U_n)\}, \textit{otherwise} \end{cases} \tag{11}$$

where Eq. (11) indicates that pivot attribute P_c has the minimum correlation to other attributes in $\boldsymbol{U^h}$-{P_1,…, P_{c-1}}.

Next, each pivot attribute P_c is used to generate an attribute subset $\boldsymbol{Sub\text{-}U^c}$, namely $\boldsymbol{Sub\text{-}U_c}$={$P_c$}, and each antecedent attribute U_m ($U_m \in \boldsymbol{U^h}$) is thereafter assigned to these C subsets as follows:

$$\boldsymbol{Sub-U^c} = \{U_m\} \cup \boldsymbol{Sub-U^c}, c = \arg\max_{n=1,...,C} \{AR(P_n, U_m)\} \tag{12}$$

where Eq. (12) indicates that antecedent attribute U_m has the maximum correlation to P_c comparing to other pivot attributes.

For each subset $\boldsymbol{Sub\text{-}U^c}$, a new pivot attribute P_c' should be determined by using Eq. (13). If { P_c'; c=1,…, C} is completely equal to {P_c; c=1,…, C}, the procedure of attribute clustering is finished; Otherwise, it needs to reassign each antecedent attribute U_m ($U_m \in \boldsymbol{U^h}$) to the C subsets $\boldsymbol{Sub\text{-}U^c}$ which are initialized by $\boldsymbol{Sub\text{-}U_c}$={ P_c' }.

$$P_c' = \arg\max_{U_m \in \boldsymbol{Sub-U^c}} \{\sum_{U_n}^{\boldsymbol{Sub-U^c}} AR(U_m, U_n)\} \tag{13}$$

where Eq. (13) indicates that the new pivot attribute has the maximum correlation to other attributes in $\boldsymbol{Sub\text{-}U^c}$.

Finally, when the number of attributes in subset $\boldsymbol{Sub\text{-}U^c}$ is $|\boldsymbol{Sub\text{-}U^c}|>S$, the set $\boldsymbol{U^h}$ is replaced by the subset $\boldsymbol{Sub\text{-}U^c}$ to perform ***Step* 2** for a new round of attribute clustering; Otherwise, the subset $\boldsymbol{Sub\text{-}U^c}$ is regarded as an internal node when $|\boldsymbol{Sub\text{-}U^c}|>1$ and as a leaf node when $|\boldsymbol{Sub\text{-}U^c}|=1$. Here, S denotes the number of attributes to construct a downsized BRB.

***Step* 3**: To construct a sub-BRB for each internal node in the MTS-BRB.

Suppose that there are H internal nodes obtained from ***Step 2***. H sub-BRBs with belief rules {$R_{h,k}$; k=1,…, L_h} (h=1,…, H) can be constructed based on ***Definition 5*** and the three categories below:

(1) For the internal node whose child nodes only include leaf nodes, its sub-BRB can be constructed by covering all combinations of all assessment ratings of all child leaf nodes.

(2) For the internal node whose child nodes include internal and leaf nodes, its sub-BRB is constructed by covering all combinations of all consequents of all child internal nodes and all assessment ratings of all child leaf nodes.

(3) For the internal node whose child nodes only include internal nodes, its sub-BRB is constructed by covering all combinations of all consequents of all child internal nodes.

For the above-mentioned MTS-BRB modeling procedure, C and S are vital thresholds. Obviously, if C and S are big, the resulting MTS-BRB is approximate to a conventional BRB; otherwise, MTS-BRB is composed of lots of sub-BRBs with a smaller size.

4.2. MTS-BRB inferencing procedure

When a MTS-BRB is constructed, the relevant MTS-BRB inference procedure should be proposed to guarantee that the BRB expert system can produce an inference result. Considering that belief structure is a powerful scheme of information representation, it is regarded as a carrier to transfer information from child nodes to their parent node in the MTS-BRB. Fig. 8 shows a process of producing an inference result $f(\boldsymbol{x})$ for a given input data vector $\boldsymbol{x}=\{x_1,\ldots, x_M\}$ from the leaf nodes to the root node in a MTS-BRB.

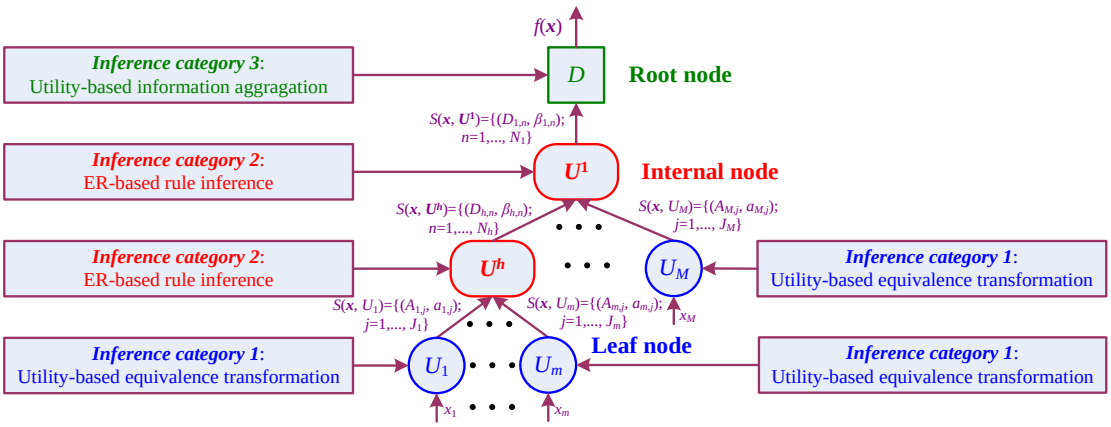


Fig. 8. Illustration of MTS-BRB inferencing procedure

From Fig. 8, it is clear that the calculation of belief distribution used to transfer information can be divided into three inference categories. They are provided as follows:

***Inference category* 1**: Calculation of belief distribution in leaf nodes.

For the leaf node in MTS-BRB, its input is data vectors, denoted as $\boldsymbol{x}=\{x_m; m=1,\ldots, M\}$. Taking the mth leaf node as an example, the mth input data x_m can be used to calculate the belief distribution of the mth leaf node using Eqs. (7) - (8):

$$S(\boldsymbol{x},U_m)=\{(A_{m,j},a_{m,j}); j=1,\ldots, J_m\} \tag{14}$$

where $A_{m,j}$ denotes the jth assessment rating of the mth antecedent attribute; $a_{m,j}$ denotes the belief degree of assessment rating $A_{m,j}$; J_m is the total number of assessment ratings for the mth antecedent attribute.

***Inference category* 2**: Calculation of belief distribution in internal nodes.

For the internal node in MTS-BRB, its input is the belief distribution provided by child nodes. Taking the hth internal node $\boldsymbol{U}^h$ as an example, assume that there are M_h child nodes $U_{h,m}$ (m=1,…, M_h) and $J_{h,m}$ assessment ratings $A_{h,m,j}$ (j=1,…, $J_{h,m}$) for the mth child node, as well as N_h consequents $\{D_{h,n}; n=1,\ldots, N_h\}$, the hth internal node $\boldsymbol{U}^h$ corresponds to the hth sub-BRB shown in ***Definition* 5**. When the input belief distribution for the mth child node is $\{(A_{h,m,j}, a_{h,m,j}); j=1,\ldots, J_{h,m}\}$,

the output of the sub- BRB can be obtained by the following two steps:

***Step* 1**: To calculate activation weights in the *h*th sub-BRB. The activation weight of the *k*th (k=1,…, L_h) rule in the *h*th sub-BRB is calculated by

$$w_{h,k} = \frac{\theta_{h,k}\prod_{m=1}^{M_h}(a_m^{h,k})^{\bar{\delta}_m}}{\sum_{l=1}^{L_h}(\theta_{h,l}\prod_{m=1}^{M_h}(a_m^{h,l})^{\bar{\delta}_m})}, \bar{\delta}_m = \frac{\delta_{h,m}}{\max_{i=1,\ldots,M_h}\{\delta_{h,i}\}} \tag{15}$$

where $a_m^{h,k}$ =$a_{h,m,j}$ if assessment rating $A_m^{h,k}$ used in rule $R_{h,k}$ is equal to assessment rating $A_{h,m,j}$ of the *m*th child node.

***Step* 2**: To integrate activation rules in the *h*th sub-BRB. After calculating activation weights, the rule whose activation weight is greater than 0 should be used to generate a combined belief distribution using the ER algorithm as follows:

$$S(\boldsymbol{x},\boldsymbol{U}^h) = \{(D_{h,n},\beta_{h,n}); n = 1,\ldots,N_h\} \tag{16}$$

where

$$\beta_{h,n} = \frac{\prod_{k=1}^{L_h}\left(w_{h,k}\beta_{h,n,k} + 1 - w_{h,k}\sum_{n=1}^{N_h}\beta_{h,n,k}\right) - \prod_{k=1}^{L_h}\left(1 - w_{h,k}\sum_{n=1}^{N_h}\beta_{h,n,k}\right)}{\sum_{n=1}^{N_h}\prod_{k=1}^{L_h}\left(w_{h,k}\beta_{h,n,k} + 1 - w_{h,k}\sum_{j=1}^{N_h}\beta_{h,j,k}\right) - (N_h - 1)\prod_{k=1}^{L_h}\left(1 - w_{h,k}\sum_{j=1}^{N_h}\beta_{h,j,k}\right) - \prod_{k=1}^{L_h}\left(1 - w_{h,k}\right)}. \tag{17}$$

***Inference category* 3**: Calculation of output in root node.

For the root node in MTS-BRB, its input is the belief distribution provided by the 1st internal node $\boldsymbol{U}^1$. Assume that the input belief distribution of root node is $S(\boldsymbol{x}, \boldsymbol{U}^1)=\{(D_n, \beta_n); n=1,\ldots, N\}$. Hence, when the modeling problem is a regression problem and $u(D_n)$ is the utility value of the *n*th consequent, the output of root node can be calculated as follows:

$$f(\boldsymbol{x}) = \sum_{n=1}^{N}\left(u(D_n)\beta_n\right) + \frac{u(D_1)+u(D_N)}{2}\left(1 - \sum_{n=1}^{N}\beta_n\right). \tag{18}$$

When the problem is a classification problem and D_n is the *n*th class, the output of root node can be calculated as follows:

$$f(\boldsymbol{x}) = D_t, t = \arg\max_{n=1,\ldots,N}\{\beta_n\}. \tag{19}$$

4.3. MTS-BRB learning procedure

The parameter values of a MTS-BRB may initially be given by domain experts based on personal experiences, but usually fail to constitute a BRB expert system with sufficiently high performance, because it difficult or even impossible for experts to provide right parameter values in a complex problem. Hence, historical data need to be collected to optimally train these parameters. For this purpose, this subsection aims to propose a MTS-BRB learning procedure. Fig. 9 shows the possible parameters needed to be optimized in a MTS-BRB from the leaf nodes to the root node.

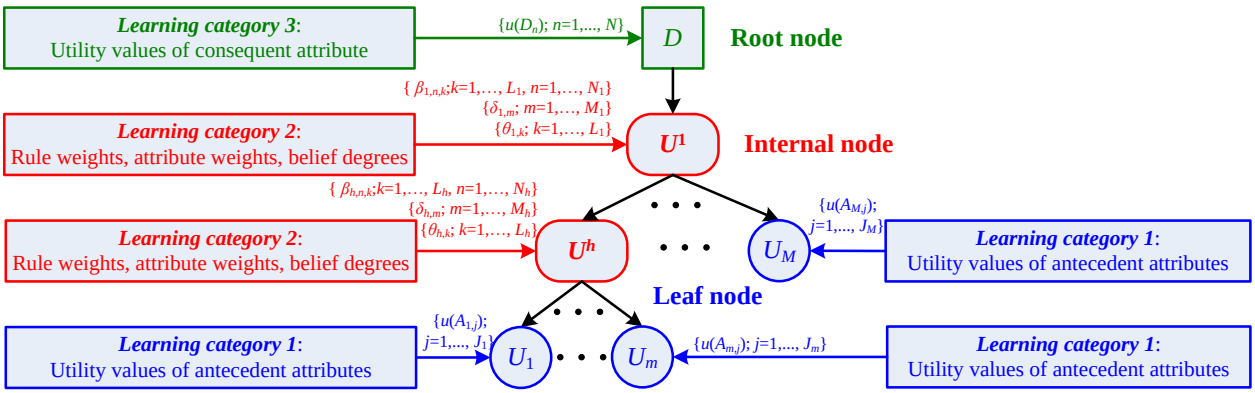


Fig. 9. Illustration of MTS-BRB learning procedure

From Fig. 9, the three learning categories can be used to group the parameters that need to be optimized in an

MTS-BRB:

Learning category 1: Optimization of utility values in leaf nodes.

For the leaf node in MTS-BRB, its utility values should be optimized. Assume that there are M antecedent attributes U_m (m=1,…, M) having J_m assessment ratings $A_{m,j}$ (j=1,…, J_m) and utility values $u(A_{m,j})$ (j=1,…, J_m). The constraints used in the MTS-BRB learning model are as follows:

$$u(A_{m,1}) = lb_m; m = 1,...,M \tag{20}$$

$$u(A_{m,J_m}) = ub_m; m = 1,...,M \tag{21}$$

$$u(A_{m,j}) \le u(A_{m,j+1}); j = 1,...,J_m - 1; m = 1,...,M \tag{22}$$

where lb_m and ub_m denote the lower and upper bounds of collected input data in the mth antecedent attribute.

Learning category 2: Optimization of weights and belief degrees in internal nodes.

For the internal node in MTS-BRB, its rule weights, attribute weights, and belief degrees should be optimized. Assume there are H internal nodes, and the hth (h= 1,…, H) internal node $\boldsymbol{U^h}$ has N_h consequents $D_{h,n}$ (n=1,…, N_h) and M_h child nodes $U_{h,m}$ (m= 1,…, M_h) with $J_{h,m}$ assessment ratings $A_{h,m,j}$ (j=1,…, $J_{h,m}$). The constraints used in the MTS-BRB learning model are as follows:

(1) The attribute weights for M_h child nodes. The attribute weight $\delta_{h,m}$ is normalized, so it is between 0 and 1, *i.e.*,

$$0 \le \delta_{h,m} \le 1; m = 1,...,M_h; h = 1,...,H \tag{23}$$

(2) The rule weights for L_h belief rules. The rule weight $\theta_{h,k}$ is normalized, so that it is between 0 and 1, *i.e.*,

$$0 \le \theta_{h,k} \le 1; k = 1,...,L_h; h = 1,...,H \tag{24}$$

(3) The belief degrees for N_h consequents in L_h belief rules. The belief degree $\beta_{h,n,k}$ is between 0 and 1, *i.e.*,

$$0 \le \beta_{h,n,k} \le 1; n = 1,...,N_h; k = 1,...,L_h; h = 1,...,H \tag{25}$$

(4) If a belief rule is complete, the total belief degree of the belief rule should be equal to 1, *i.e.*,

$$\sum_{n=1}^{N_h} \beta_{h,n,k} = 1; k = 1,...,L_h; h = 1,...,H \tag{26}$$

Learning category 3: Optimization of utility values in root node.

For the root node in the MTS-BRB, its utility values should be optimized. Assume the consequent attribute D has N consequents D_n (n=1,…, N) and utility values $u(D_n)$ (n=1,…, N). The constraints used in the MTS-BRB learning model are as follows:

$$u(D_1) = lb \tag{27}$$

$$u(D_N) = ub \tag{28}$$

$$u(D_n) \le u(D_{n+1}), n = 1,...,N-1 \tag{29}$$

where lb and ub denote the lower and upper bounds of collected output data in the consequent attribute.

Based on the above three learning categories, a MTS-BRB learning objective can be given based on T input-output data pairs <$\boldsymbol{x_t}$, y_t> (t=1,…, T), where the objective is to minimize the error between estimated outputs and actual outputs:

$$\begin{aligned} &\min\ MAE(u(A_{m,j}), u(D_n), \delta_{h,m}, \theta_{h,k}, \beta_{h,n,k}) = \sum_{t=1}^{T} \left| f(\boldsymbol{x_t}) - y_t \right| \\ &s.t.\ \text{the constains shown in Eqs.}\ (22)-(28) \end{aligned} \tag{30}$$

where $MAE(u(A_{m,j}), u(D_n), \delta_{h,m}, \theta_{n,k}, \beta_{h,n,k})$ denotes the mean absolute error (MAE) of the BRB expert system composed by the parameters $u(A_{m,j})$, $u(D_n)$, $\delta_{h,m}$, $\theta_{n,k}$, and $\beta_{h,n,k}$.

4.4. Framework of MTS-BRB expert system

When modelling complex problems and resolving the combinatorial explosion problem of conventional BRB, the MTS-BRB modelling, inferencing, and learning techniques offer an efficient way to use even fewer belief rules. These processes enable the development of a novel BRB expert system known as the MTS-BRB expert system, whose framework is depicted in Fig. 10.

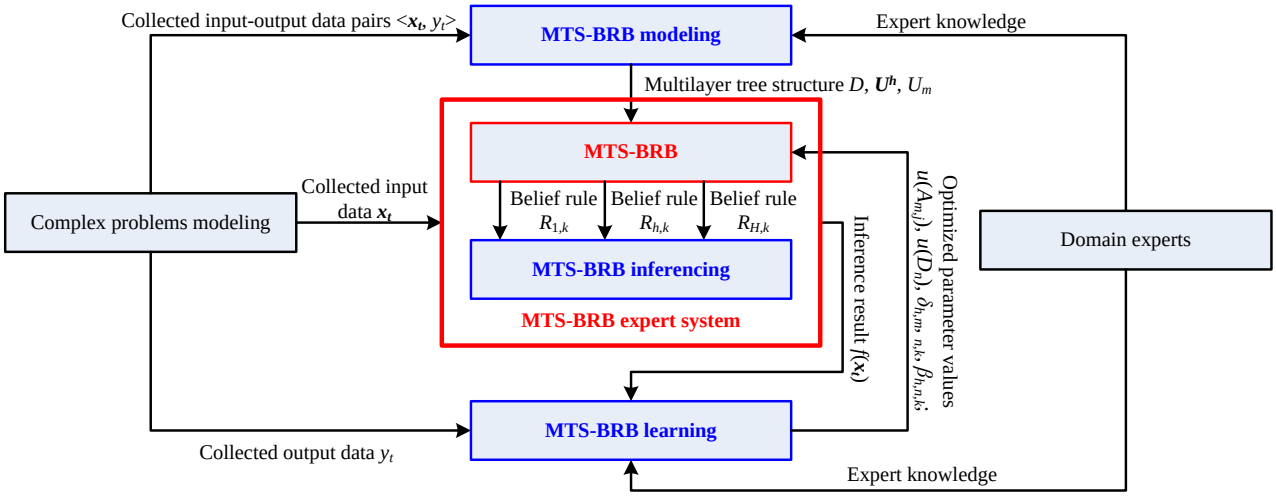


Fig. 10. Generic framework of MTS-BRB expert system

It is clear from Fig. 10 that MTS-BRB expert system consists of a MTS-BRB and its inferencing procedure, which can produce an inference result based on MTS-BRB for a given input data (Please see ***Section* 3.2** for details). For constructing a MTS-BRB, the collected input-output data pairs of complex problems and the expert knowledge obtained from domain experts should be provided (Please see ***Section* 3.1** for details). Furthermore, the MTS-BRB learning procedure can be used to guarantee an excellent performance of MTS-BRB expert system (Please see ***Section* 3.3** for details).

On the basis of above components, the advantages of MTS-BRB expert system can be summarized by comparing to some classic methods as follows:

(1) MTS-BRB expert system has a complete rule representation scheme over some fuzzy systems, because the use of belief structure in THEN part guarantees the fuzzy, probabilistic, and incomplete uncertainties coexist in belief rules.

(2) MTS-BRB expert system provides a cooperation scheme with attribute selection and extraction techniques, thus the existing techniques, *e.g.*, PCA, MDS, and GT, can be embedded into MTS-BRB modeling procedure.

(3) MTS-BRB expert system is based on the data collected from complex problems to self-organize sub-BRB and MTS, this determines the uniqueness of both sub-BRB and MTS and it is different from ensemble methods, *e.g.*, random forest.

5. Case Studies

In this section, the background of benchmark problems is introduced in ***Section* 5.1**. Next, the process of developing a MTS-BRB expert system for R&D projects risk assessment is provided in ***Section* 5.2**. Finally, the comparative analysis of the MTS-BRB expert system with some existing studies is given in ***Section* 5.3** and ***Section* 5.4**.

5.1. Background of benchmark problems

This subsection aims to introduce the background of benchmark problems sourced from R&D projects risk assessment and UCI classification database, where the former one is the questionnaire data collected from top managers, departmental

managers and project managers in Chinese industry (Yang *et al*., 2019); the latter one is a well-known database maintaining lots of classification datasets as a service to the machine learning community (Dua and Graff, 2019)

The first benchmark problem is R&D projects risk assessment, which has 13 risk factors that may affect the success of each R&D project. These risk factors include: 1) project manage competency (PMC); 2) project termination quality (PTQ); 3) formalization of portfolio management (FPM); 4) top management involvement (TMI); 5) strategic consistency (SC); 6) business cooperation (BC); 7) market uncertainty (MU); 8) technology uncertainty (TU); 9) company sales growth (CSG); 10) average net profit (ANP); 11) project success rate (PS); 12) company scales (CS); and 13) the number of ongoing projects (NOP). Based on the 13 risk factors, a total of 169 historical data collected from 169 Chinese industries and their data statistics are shown in Table 2.

Table 2. Data statistics of Chinese industries

Name	PMC	PTQ	FPM	TMI	SC	BC	MU	TU	CSG	ANP	PS	CS	NOP	PP
Minimum	1.40	1.40	1.50	1.00	1.40	1.33	1.40	1.00	1.00	1.00	1.00	2.00	3.00	2.00
Average	3.89	3.47	3.92	3.78	3.74	3.55	3.49	3.84	3.49	3.33	3.49	4.20	3.97	3.67
Maximum	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Std. dev.	0.71	0.68	0.70	0.74	0.68	0.83	0.84	0.87	1.18	1.10	0.99	0.80	0.90	0.69

The second benchmark problem is UCI classification problems and 4 commonly used datasets are used to validate the performance of MTS-BRB expert system. The main characteristics of these 4 datasets are summarized in Table 3, where "#Data" denotes the number of data, "#Attribute" denotes the number of continuous and discrete attributes, "#Continuous" denotes the number of continuous attributes, "#Discrete" denotes the number of discrete attributes, and "#Class" denotes the number of classes. Here, it is worth noting that the data of continuous attributes includes real data and the data of discrete attributes includes integer data and categorical data.

Table 3. Basic information of UCI classification datasets

Dataset	#Data	#Attribute	#Continuous	#Discrete	#Class
Appendicitis	106	7	7	0	2
Heart	270	13	1	12	2
Wine	178	13	13	0	3
Cleveland	297	13	13	0	5

5.2. Development process of the MTS-BRB expert system

In this subsection, the development process of the MTS-BRB expert is provided based on R&D project risk assessment, including MTS-BRB modeling, learning, and inferencing procedures in ***Sections* 5.2.1** to **5.2.3**, respectively. Additionally, in order to justify the advantage of the proposed procedures, the conventional BRB is used as baseline for model validation.

5.2.1. Project risk assessment with MTS-BRB modeling procedure

To perform the MTS-BRB modeling procedure, a common feature extraction technique, namely PCA, is introduced to select main PCs from 13 risk facts with more than 80% accumulative contribution ratio (ACR). Table 4 shows eigenvalues, contribution ratio (CR), and ACR for the thirteen PCs.

Table 4. Eigenvalue, CR, and ACR of thirteen PCs

Name	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13
Eigenvalue	4.937	1.566	1.494	0.902	0.812	0.645	0.607	0.482	0.395	0.359	0.289	0.283	0.231
CR	0.380	0.120	0.115	0.069	0.062	0.050	0.047	0.037	0.030	0.028	0.022	0.022	0.018
ACR	0.380	0.500	0.615	0.685	0.747	0.797	0.843	0.880	0.911	0.938	0.961	0.982	1.000

From Table 4, the top seven PCs should be extracted. Clearly, the size of conventional BRB would be 3^{13}=1,594,323 rules without PCA and 3^7=2,187 rules with PCA, respectively, while assuming 3 assessment ratings for each PC. Hence, it is necessary to construct MTS-BRB for project risk assessment. In what follows, the intermediate results of the modeling procedure are provided as follows:

***Step* 1**: Generation of a root node and the first internal node.

According to ***Step* 1** shown in ***Section* 4.1**, the project risk consequence (consequent attribute) D should be regarded as a root node and the filtered set of project risk factors (antecedent attributes) $\boldsymbol{U^1}$ obtained from PCs as the first internal node to construct a MTS-BRB. As a consequence, the root node is corresponding to PP, namely D=PP, and the first internal node is corresponding to the top seven PCs, namely $\boldsymbol{U^1}$={PC_1, PC_2, PC_3, PC_4, PC_5, PC_6, PC_7}.

***Step* 2**: Calculation of factor relations for the top seven PCs.

Based on ***Step* 2.1** at ***Section* 4.1**, the assessment ratings of those top seven PCs should be given firstly according to the domain knowledge of experts and 169 input data of R&D projects can be transformed into belief distributions Afterwards, all belief distributions are used to calculate the attribute relation of the seven PCs based on ***Definition* 6**, the corresponding attribute relations are shown in Table 5.

Table 5. Attribute relation of the top seven PCs

Factor Relation	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7
PC_1	1.000000	0.000181	0.001933	0.000712	0.000278	0.000491	0.002949
PC_2	0.000181	1.000000	0.001028	0.002567	0.001261	0.000575	0.000928
PC_3	0.001933	0.001028	1.000000	0.002531	0.001142	0.000578	0.000186
PC_4	0.000712	0.002567	0.002531	1.000000	0.001208	0.000621	0.001204
PC_5	0.000278	0.001261	0.001142	0.001208	1.000000	0.000246	0.000587
PC_6	0.000491	0.000575	0.000578	0.000621	0.000246	1.000000	0.000091
PC_7	0.002949	0.000928	0.000186	0.001204	0.000587	0.000091	1.000000

***Step* 3**: Attribute clustering for internal nodes.

Based on ***Step* 2.2** at ***Section* 4.1**, the set of top seven PCs, namely $\boldsymbol{U^1}$={PC_1, PC_2, PC_3, PC_4, PC_5, PC_6, PC_7} should be packaged into three clusters when C=3. Firstly, three pivot attribute can be selected by

$$P_1 = \arg\min_{U_m \in \boldsymbol{U}^1} \{\sum_{U_n}^{\boldsymbol{U}^1} AR(U_m, U_n)\} = \mathrm{PC}_6 \tag{31}$$

$$P_2 = \arg\min_{U_m \in \boldsymbol{U}^1 - \{\mathrm{PC}_6\}} \{\sum_{U_n}^{\{\mathrm{PC}_6\}} AR(U_m, U_n)\} = \mathrm{PC}_7 \tag{32}$$

$$P_3 = \arg\min_{U_m \in \boldsymbol{U}^1 - \{\mathrm{PC}_6, \mathrm{PC}_7\}} \{\sum_{U_n}^{\{\mathrm{PC}_6, \mathrm{PC}_7\}} AR(U_m, U_n)\} = \mathrm{PC}_3 \tag{33}$$

Afterwards, the remaining four PCs are assigned to the three clusters ***Sub-U*^{1}**, ***Sub-U*^{2}**, and ***Sub-U*^{3}**, which should be initialized using the selected three pivot attributes, namely ***Sub-U*^{1}**={PC_6}, ***Sub-U*^{2}**={PC_7}, and ***Sub-U*^{3}**={PC_3}. Taking PC_1 for example, Table 4 shows that the attribute relations among PC_1, PC_6, PC_7, and PC_3 are AR(PC_1, PC_6)=0.000491, AR(PC_1, PC_7)=0.002949, and AR(PC_1, PC_3)=0.001933, thus PC_1 should be assigned to ***Sub-U*^{2}**, namely ***Sub-U*^{2}**={PC_1, PC_7}, because

of $AR(PC_1, PC_7)>AR(PC_1, PC_3)>AR(PC_1, PC_6)$. In the same way, PC_2, PC_4, and PC_5 are assigned to $\boldsymbol{Sub\text{-}U^3}$, namely $\boldsymbol{Sub\text{-}U^3}=\{PC_2, PC_3, PC_4, PC_5\}$.

Next, for $\boldsymbol{Sub\text{-}U^1}=\{PC_6\}$, $\boldsymbol{Sub\text{-}U^2}=\{PC_1, PC_7\}$, and $\boldsymbol{Sub\text{-}U^3}=\{PC_2, PC_3, PC_4, PC_5\}$, three new pivot attributes should be selected until the three new pivot attributes are not completely equal to the three original ones. the four PCs, including PC_1, PC_2, PC_3, and PC_5 should be reassigned to the three new clusters $\boldsymbol{Sub\text{-}U^1}$, $\boldsymbol{Sub\text{-}U^2}$, and $\boldsymbol{Sub\text{-}U^3}$, which are initialized using the new pivot attribute, namely $\boldsymbol{Sub\text{-}U^1}=\{PC_6\}$, $\boldsymbol{Sub\text{-}U^2}=\{PC_7\}$, and $\boldsymbol{Sub\text{-}U^3}=\{PC_4\}$. As such, the three new clusters are $\boldsymbol{Sub\text{-}U^1}=\{PC_6\}$, $\boldsymbol{Sub\text{-}U^2}=\{PC_1, PC_7\}$, and $\boldsymbol{Sub\text{-}U^3}=\{PC_2, PC_3, PC_4, PC_5\}$. It is obviously that the clustering of $\boldsymbol{U^1}$ is finished because the new pivot clusters are completely equal to the original ones.

Finally, when the maximum number of the attributes used to construct a downsized BRB is set as 3, namely $S=3$, $\boldsymbol{Sub\text{-}U^1}$ is regarded as a leaf node of the internal node $\boldsymbol{U^1}$, $\boldsymbol{Sub\text{-}U^2}$ is regarded as the second internal node, namely $\boldsymbol{U^2}=\{PC_1, PC_7\}$, and the PCs in $\boldsymbol{Sub\text{-}U^3}$ are all regarded as the leaf nodes of $\boldsymbol{U^3}$ because of $3>|\boldsymbol{Sub\text{-}U^3}|=2>1$, and $\boldsymbol{Sub\text{-}U^3}$ is regarded as the third internal node, namely $\boldsymbol{U^3}=\{PC_2, PC_3, PC_4, PC_5\}$, which needs for attribute clustering again.

After performing the above procedure, a MTS-BRB can be constructed and it is shown in Fig. 11.

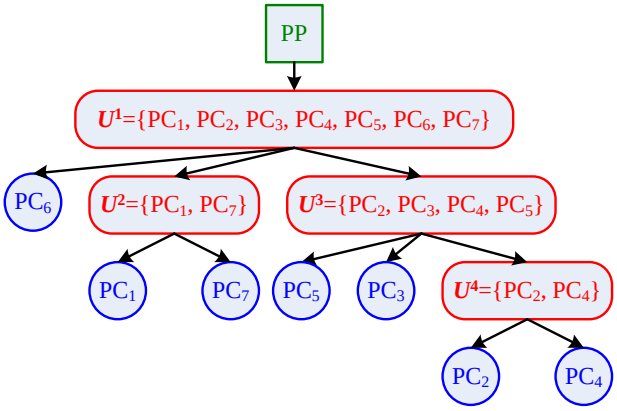


Fig. 11. MTS-BRB for R&D project risk assessment

***Step* 4**: Construction of sub-BRBs for each internal node.

Based on ***Step* 3** shown in ***Section* 4.1**, all internal nodes in MTS-BRB should be used to construct sub-BRBs. When the number of consequents used for the internal nodes $\boldsymbol{U^1}$, $\boldsymbol{U^2}$, $\boldsymbol{U^3}$, and $\boldsymbol{U^4}$ is set as 5, 3, 3, and 3, four sub-BRBs can be constructed and their size is shown in Table 6, in which the sub-BRB_1, sub-BRB_2, sub-BRB_3, and sub-BRB_4 denote the sub-BRBs regarding $\boldsymbol{U^1}$, $\boldsymbol{U^2}$, $\boldsymbol{U^3}$, and $\boldsymbol{U^4}$; the ratio denotes number of training data per local region.

Table 6. Size comparison of conventional BRB and MTS-BRB

Type	Chinese industries	Four sub-BRBs in MTS-BRB				MTS-BRB	Conventional BRB	
		Sub-BRB_1	Sub-BRB_2	Sub-BRB_3	Sub-BRB_4		Without PCA	With PCA
Size	169	27	9	27	9	72	1,594,323	2,187
No. of local regions	-	8	4	8	4	8	8,192	128
Ratio	-	21.125	42.250	21.125	42.250	21.125	0.021	1.32

From Table 6, the size of MTS-BRB 27+9+27+9=72 is far less than that of conventional BRBs, which has 2,187 rules with PCA and 1,594,323 rules without PCA. In this case, although PCA can significantly downsize a BRB, the size of downsized BRB is still greater than the number of Chinese industries. In comparison of the number of local regions, it is worth noting that the number of local regions in the MTS-BRB is obtained from the maximum one in its four sub-BRBs, so there are plenty of data to train the MTS-BRB, *i.e.*, the ratio of MTS-BRB is 21.125 which is significantly greater than that of BRBs 0.021 and 1.32.

5.2.2. Project risk assessment with MTS-BRB learning procedure

In order to obtain the optimal parameters of MTS-BRB, the collected data of Chinese industries are used to perform the MTS-BRB learning procedure. Hence, 169 Chinese industries are randomly divided into 2 parts: 90% industries as training data and 10% industries as testing data, namely 152 training data and 17 testing data.

As shown in Fig. 1, the MTS-BRB has one root node, four internal nodes, and seven leaf nodes. Hence, based on ***Learning category*** **1**, **2**, and **3**, the following BRB learning objective can be established:

$$\min MAE(u(A_{m,j}),u(D_n),\delta_{h,m},\theta_{h,k},\beta_{h,n,k}) = \sum_{t=1}^{T}\left| f(\boldsymbol{x}_t) - y_t \right| \tag{34}$$

$$s.t.\, u(A_{m,1}) = lb_m, lb_m \le u(A_{m,2}) \le ub_m, u(A_{m,3}) = ub_m; m = 1,...,M \tag{35}$$

$$u(D_1) = lb, u(D_n) \le u(D_{n+1}), u(D_5) = ub; n = 1,...,N \tag{36}$$

$$0 \le \beta_{h,n,k} \le 1, \sum_{n=1}^{N_h} \beta_{h,n,k} = 1; n = 1,...,N_h; k = 1,...,L_h; h = 1,...,H \tag{37}$$

$$0 \le \delta_{h,m}, \theta_{h,k} \le 1; m = 1,...,M_h; k = 1,...,L_h; h = 1,...,H \tag{38}$$

where the constraints shown in Eq. (35) are on the seven leaf nodes, lb_m and ub_m denote the lower and upper bounds of the mth PC, M denotes the number of leaf nodes and it is M=7; the constraints shown in Eq. (36) are on the root node, lb and ub denote the lower and upper bounds of PP, N denotes number of consequents and it is N=5; the constraints shown in Eq. (37) to (38) are on the four internal nodes, N_h denotes the number of consequents in the hth internal node and they are N_1=5, N_2=3, N_3=3, and N_4=3, L_h denotes the number of rules in the hth internal node and they are L_1=27, L_2=9, L_3=27, and L_4=9, M_h denotes the number of child nodes in the hth internal node and they are M_1=3, M_2=2, M_3=3, and M_4=2, H denotes the number of internal node and it is H=4; T denotes the number of training data and it is T=152, $f(\boldsymbol{x}_t)$ denotes the inference result of MTS-BRB for the given input data $\boldsymbol{x}_t$, y_t denotes the actual output of the given input data $\boldsymbol{x}_t$.

To solve the MTS-BRB learning model, the differential evolution (DE)-based learning algorithm (Yang *et al*., 2017) is introduced. Table 7 shows the corresponding results of MTS-BRB and the conventional BRB with and without PCA under the same learning algorithm and experimental conditions, in which NA means that the running time of learning exceeds one week and it has to be artificially terminated; baseline ratio is the result of MTS-BRB divided by the result of the baseline. Here, it is worth noting that the largest baseline ratio is considered to be the best level for MTS-BRB.

Table 7. Comparison of conventional BRB and MTS-BRB learning

Type	MTS-BRB	Conventional BRB		Baseline ratio	
		Without PCA	With PCA	Without PCA	With PCA
Learning time (HH:MM:SS)	00:05:14	NA(>>1 week)	07:21:47	NA	84.4
No. of parameters to be trained	378	9,565,995	13,155	25,306.9	34.8
MAE at training data	0.434341	NA	0.424851	NA	1.0

From Table 7, it is clear that MTS-BRB learning is faster than conventional BRB learning. The main reason is that there are significant differences in the number of parameters that need to be trained in the learning model. For example, the MTS-BRB has only 378 parameters, whereas the conventional BRB without and with PCA have 9,565,995 and 13,155 parameters, respectively, and have baseline ratios of 25,306.9 and 34.8. Additionally, the baseline ratio of the MAE between MTS-BRB and conventional BRB is similar at 1.0. As a result, it can be said that MTS-BRB can reach the same degree of

modelling skill as the conventional BRB with a lot less effort and learning.

5.2.3. Project risk assessment with MTS-BRB inferencing procedure

Continuing with the R&D project risk assessment, this subsection is to show how to produce an inference result for a given input data based on MTS-BRB inferencing procedure.

Suppose a given input data is $\boldsymbol{x}$=<x_{PC1}=2.5565, x_{PC2}=-0.7364, x_{PC3}=-0.2140, x_{PC4}=-0.7815, x_{PC5}=1.7183, x_{PC6}=0.2485, x_{PC7}=-0.7739, y=4.25>. The procedure of producing a corresponding inference result is provided as follows:

Firstly, for the seven leaf nodes PC_m (m=1,…, 7), their belief distributions can be calculated according to ***Inference category* 1**. Taking x_{PC1}=2.5565 for example, the obtained belief distribution is $S(\boldsymbol{x}, PC_1)$={(*Low*, 0), (*Middle*, 0.2330), (*High*, 0.7670)} because x_{PC1}=2.5565 is located in the range of the referential values *Middle* and *High*. Similarly, the other belief distributions can be calculated.

Secondly, for the four internal nodes, the belief distributions can be calculated by according to ***Inference category* 2**. Taking internal node $\boldsymbol{U}^4$ for example, Fig. 10 shows that $\boldsymbol{U}^4$ has two leaf nodes PC_2 and PC_4, thus the inputs of the sub-BRB_4 are $S(\boldsymbol{x}, PC_2)$ and $S(\boldsymbol{x}, PC_4)$. Afterwards, the activation weights for each rule in sub-BRB_4 can be calculated based on ***Step* 1** shown in ***Section* 4.2**. Next, the combined belief distribution $S(\boldsymbol{x}, \boldsymbol{U}^4)$={($D_{4,1}$, 0.1768), ($D_{4,2}$, 0.3448), ($D_{4,3}$, 0.4784)} can be calculated using ***Step* 2** shown in ***Section* 4.2**. Similarity, the combined belief distribution of the other internal nodes can be obtained, namely, $S(\boldsymbol{x}, \boldsymbol{U}^3)$={($D_{3,1}$, 0.2431), ($D_{3,2}$, 0.4587), ($D_{3,3}$, 0.2982)}, $S(\boldsymbol{x}, \boldsymbol{U}^2)$={($D_{2,1}$, 0.1116), ($D_{2,2}$, 0.1319), ($D_{2,3}$, 0.7565)}, and $S(\boldsymbol{x}, \boldsymbol{U}^1)$={($D_{1,1}$, 0.1029), ($D_{1,2}$, 0.1165), ($D_{1,3}$, 0.1104) , ($D_{1,4}$, 0.2029) , ($D_{1,5}$, 0.4673)}

Finally, according to ***Inference category* 3**, the belief distribution of the root node equates that of the first internal node, namely $S(\boldsymbol{x}, PP)=S(\boldsymbol{x}, \boldsymbol{U}^1)$. Such that, the output of MTS-BRB can be calculated using the belief distribution $S(\boldsymbol{x}, PP)$, namely, $f(\boldsymbol{x})$=4.1438. Thus, the error between $f(\boldsymbol{x})$ and y is 4.25-4.1438=0.1062. Correspondingly, MAPE and RMSE of MTS-BRB for testing data are shown in Table 8, where NA means that the conventional BRB without PCA fails to produce inference results because of time-consuming process.

Table 8. Performance comparison of conventional BRB and MTS-BRB

Type	MTS-BRB	Conventional BRB		Baseline ratio	
		Without PCA	With PCA	Without PCA	With PCA
MAPE at testing data	9.0139	NA	10.5899	NA	1.2
RMSE at testing data	0.4611	NA	0.4902	NA	1.1

It is clear from Table 8 that the MAPE and RMSE of MTS-BRB are slightly better than that of the conventional BRB with PCA, and their baseline ratio is 1.2 and 1.1, but MTS-BRB has significant advantages in the terms of parameters and learning time, *i.e.*, Table 7 shows a total of 7 hours, 21 minutes, and 47 seconds needed to obtain the conventional BRB with PCA, but only 5 minutes and 14 seconds for MTS-BRB. It is therefore possible to draw the conclusion that MTS-BRB is able to circumvent the combinatorial explosion problem of conventional BRB, resulting in a BRB expert system with fewer rules and lower prediction error for modelling complicated problems.

5.3. Comparative analysis with some existing project risk assessment studies

In order to validate the effectiveness of the proposed MTS-BRB expert system, some well-known feature selection or extraction techniques, existing BRB expert systems and classical methods for R&D project risk assessment are introduced to carry out three comparative experiments under ten-fold cross validation.

(1) The first experiment to compare with different feature selection techniques

According to the MTS-BRB modeling procedure shown in ***Section* 5.2.1**, four MTS-BRBs can be constructed using four feature selection or extraction techniques. Table 9 shows the corresponding comparative results of conventional BRB and MTS-BRB expert systems. Additionally, in order to investigate the influence of maximum number of attributes and groups on MTS-BRB modeling procedure, three situations, namely, C=2 and S=2, C=2 and S=3, and C=3 and S=3, are considered to construct three kinds of MTS-BRBs.

Table 9. Comparison of conventional BRB and MTS-BRB for R&D project risk assessment

Method (No. of factors used to construct BRB)	Criteria	Conventional BRB	MTS-BRB		
			(C=2, S=2)	(C=2, S=3)	(C=3, S=3)
None	MAPE	NA	15.0995	15.9181	13.8296
(13 factors)	RMSE	NA	0.6365	0.6757	0.6023
	No. of rules	1,594,323	108	126	144
	Time(HH:MM:SS)	NA(>>1 week)	01:03:12	01:04:04	01:39:41
PCA with 80% ACR	MAPE	15.7868	12.7416	12.0333	12.8112
(7 factors)	RMSE	0.7087	0.5553	0.5408	0.5630
	No. of rules	2,187	54	72	72
	Time(HH:MM:SS)	130:52:30	00:34:18	00:40:29	00:55:56
PCA with 70% ACR	MAPE	14.0415	11.8535	11.2874	12.4961
(5 factors)	RMSE	0.6404	0.5319	0.5133	0.5471
	No. of rules	243	36	45	45
	Time(HH:MM:SS)	05:51:30	00:19:20	00:23:23	00:24:54
CFS	MAPE	12.2694	12.3123	11.9063	12.0567
(5 factors)	RMSE	0.5676	0.5408	0.5231	0.5156
	No. of rules	243	36	36	45
	Time(HH:MM:SS)	02:49:20	00:17:53	00:17:54	00:28:04
FSE-RF	MAPE	12.8208	12.3333	12.2552	12.3573
(5 factors)	RMSE	0.5933	0.5436	0.5298	0.5160
	No. of rules	243	36	36	45
	Time(HH:MM:SS)	02:35:00	00:17:00	00:17:04	00:25:03

It is clear from Table 9 that the number of rules in a MTS-BRB would be increased with the increase of the maximum numbers of groups C and attributes S, since they allow more numbers of sub-BRBs or numbers of attributes used to construct sub-BRB in the MTS-BRB. But in any case, the number of rules in MTS-BRB is far less than that in conventional BRB. Owing to too many rules in a BRB, the BRB learning procedure is time-consuming, which has to spend more than 2 hours. Moreover, the BRB learning procedure using PCA with 80% ACR needs 130 hours. On the other hand, the MTS-BRB learning procedure is more efficient, taking only 34 minutes for the PCA with 80% ACR. When comparing MAPE and RMSE, MTS-BRB expert systems perform better than BRB expert systems. For example, while employing PCA with 70% ACR, the latter's MAPE and RMSE are larger than 14 and 0.6, respectively, those of MTS-BRB expert systems are less than 14 and 0.6. It can be seen from comparing MTS-BRB expert systems with and without feature selection or extraction approaches that the latter can enhance the performance of MTS-BRB expert systems by preventing the former from being deceptive owing to redundant factors or over-fitting due to irrelevant elements.

(2) The second experiment to compare with existing BRB expert systems

In earlier research, the RS-BRB expert system and its extensions were one of the key methods for assessing the risk of R&D projects (Yang *et al*., 2019). They are also used to contrast with the MTS-BRB expert system, where RS-BRB expert system designates that the BRB expert system is modeled by using random subspaces; the R-RS-BRB expert system designates that the RS-BRB expert system is used in conjunction with the random initialization method; the AF-RS-BRB expert system designates that the RS-BRB expert system is used in conjunction with the average fusion combination method; and the WAF-RS-BRB expert system designates that the RS-BRB expert system is used in conjunction with the weighted average fusion combination method. Comparing several BRB expert systems in terms of MAPE, RMSE, and number of rules is shown in Table 10. Be aware that the MTS-BRB expert system's outputs are a representation of the worst and best outcomes from numerous experiments using the four feature extraction or selection procedures

Table 10. Comparison of different BRB expert systems for R&D project risk assessment

Criteria	RS-BRB	R-RS-BRB	AF-RS-BRB	WAF-RS-BRB	MTS-BRB
MAPE	13.73	15.85	15.01	15.05	[11.28, 12.81]
RMSE	0.5737	0.6958	0.6257	0.6143	[0.5133, 0.5630]
No. of rules	270	270	270	270	[36, 72]

Table 10 makes it clear that the MTS-BRB expert system is superior to the existing BRB expert systems for assessing the risk associated with R&D projects. For example, the MTS-BRB expert system's worst MAPE is 12.81, which is lower than the MAPEs of the RS-BRB, R-RS-BRB, AF-RS-BRB, and WAF-RS-BRB expert systems. Although the worst RMSE of the MTS-BRB expert system is somewhat lower than that of the RS-BRB expert system, there are much fewer rules in the MTS-BRB expert system.

(3) The third experiment to compare with machine learning methods

In order to further compare with MTS-BRB expert system, the classical methods for R&D project risk assessment, including artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), Bayesian network (BN), multiple regression analysis (MRA), and two recent methods, including Micro-extended BRB (Micro-EBRB) expert system (Yang *et al*., 2021) and DCFS (Wang, 2020), are applied in comparative studies, in which the PCA is used together with ANFIS to avoid too many rules generated in R&D project risk assessment; the Micro-EBRB expert system has shown its potential in the balance of accuracy and efficiency; the DCFS is a hierarchical fuzzy system. Table 11 shows the comparison of MTS-BRB expert system and different methods in terms of MAPE and RMSE.

Table 11. Comparison of MTS-BRB and existing methods for R&D project risk assessment

Criteria	ANN	ANFIS	BN	MRA	DCFS	Micro-EBRB	MTS-BRB
MAPE	16.82	16.14	17.45	17.16	13.03	13.56	[11.40, 12.81]
RMSE	0.7423	0.7401	0.6965	0.6107	0.5905	0.5861	[0.5133, 0.5630]

It is clear from Table 11 that MTS-BRB expert system is better than four classical methods and two recent methods for R&D project risk assessment in Chinese industries, *i.e.*, the worst MAPE and RMSE of MTS-BRB expert system is 12.81 and 0.5630, which are smaller than those of ANN, ANFIS, NB, MRA, and hierarchical fuzzy system, and Micro-EBRB. The final ranking of these methods is MTS-BRB > DCFS > Micro-EBRB > ANFIS > ANN > MRA > BN in the terms of MAPE and RMSE.

In summary, it is evident from the above three experiments that MTS-BRB expert system performs well in predicting

project performance from the comprehensive view of both MAPE and RMSE indicators. More importantly, MTS-BRB has smaller number of rules and less parameter learning time comparing to the conventional BRB, so that it is able to overcome the combinatorial exploration problem of conventional BRB.

5.4. Comparative analysis with some existing benchmark classification studies

In this subsection, the four benchmark datasets listed in Table 3 are used to conduct experiments for comparing with classical fuzzy system-based classifiers and machine learning-based classifiers, where the fuzzy system-based classifier include structural learning algorithm on vague environment (SLAVE) (Gonzalez and Perez, 1999), fuzzy hybrid generic-based machine learning algorithm (FH-GBML) (Ishibuchi, 2005), SGREED (Mansoori *et al*., 2008), FARC-HD (Alcala-Fdez *et al*., 2011), and DCFS (Wang, 2020); the machine learning-based classifiers include K nearest neighbor (KNN), native Bayesian (NB), decision tree (DT), support vector machine (SVM), ANN, and random forest (RF). Tables 12 and 13 show the accuracy of 4 benchmark datasets related to the above classical classifiers under 10-fold cross validation.

Table 11. Comparison of MTS-BRB expert system and fuzzy system-based classifiers

Dataset	SLAVE	FH-GBML	SGRED	FARC-HD	CFAR	DCFS	MTS-BRB
Appendicitis	82.91 (7)	86.00 (3)	84.48 (5)	84.18 (6)	87.82 (2)	84.91 (4)	88.68 (1)
Heart	71.36 (7)	75.93 (5)	73.21 (6)	84.44 (1)	82.22 (3)	80.00 (4)	82.59 (2)
Wine	89.47 (7)	92.61 (4)	91.88 (5)	94.35 (2)	93.24 (3)	91.57 (6)	97.75 (1)
Cleveland	48.82 (6)	53.51 (4)	51.59 (5)	55.24 (2)	53.88 (3)	41.41 (7)	56.90 (1)
Average rank	6.75	4	5.25	2.75	2.75	5.25	1.25

From Table 11, it is clear that the MTS-BRB expert system can obtain the best accuracy on three of four datasets and the second best accuracy on the other dataset, whose number is greater than other six kinds of fuzzy system-based classifiers. Here, it should be highlighted that DCFS is a recent representation of hierarchical fuzzy systems and FARC-HD is the fuzzy system aimed to solve the high-dimensional classification problems, all of them were proposed to improve the modeling ability of fuzzy systems when facing complex problem modeling. In the comparison of average rank, the MTS-BRB expert system outperforms all fuzzy system-based classifiers and the order of average rank is MTS-BRB (1.25) > FARC-HD (2.75) = CFAR (2.75) > FH-GBML (4) > DCFS (5.25) = SGRED (5.25) > SLAVE (6.75).

Table 12. Comparison of MTS-BRB expert system and machine learning-based classifiers

Dataset	KNN	NB	DT	SVM	ANN	RF	MTS-BRB
Appendicitis	82.08 (6)	85.85 (3)	85.85 (3)	80.19 (7)	85.85 (3)	84.91 (5)	88.68 (1)
Heart	74.81 (6)	83.70 (1)	77.41 (5)	55.56 (7)	82.22 (3)	81.48 (4)	82.59 (2)
Wine	97.19 (3)	96.63 (5)	92.13 (6)	44.38 (7)	97.17 (4)	98.31 (1)	97.75 (2)
Cleveland	55.56 (4)	54.88 (5)	56.57 (2)	53.87 (6)	52.53 (7)	56.23 (3)	56.90 (1)
Average rank	4.75	3.5	4	6.75	4.25	3.25	1.5

From Table 12, although the accuracy of the MTS-BRB expert system is even worse than some machine learning-based classifiers on datasets like Heart and Wine, where the best classifier of these two datasets is NB and RF, respectively, it is still possible to find an acceptable accuracy. This is because, on one hand, the MTS-BRB expert system is the second best accuracy on Heart and Wine. On the other hand, the MTS-BRB expert system has the maximum number of the best accuracy over other machine learning-based classifiers. Consequently, in the comparison of average rank, the MTS-BRB

expert system outperforms all machine learning-based classifiers and the order of average rank is MTS-BRB (1.5) > RF (3.25) > NB (3.5) > DT (4) > ANN (4.25) > KNN (4.75) > SVM (6.75).

In summary, for the comparison of fuzzy system-based and machine learning-based classifiers on some classification problems, the comparative results demonstrated that MTS-BRB expert system not only can be used for complex problems modelling, but also has desired classification accuracy better than some existing classifiers.

6. Conclusions

The conventional BRB's combinatorial explosion problem makes it difficult to incorporate expert knowledge and could jeopardize the interpretability of the BRB expert system. This study developed MTS to offer a novel representation scheme of hierarchical BRB for the first time, referred to as MTS-BRB, in order to address this issue. To build and optimize an MTS -BRB from data, as well as to respond to input data given to the MTS-BRB, the relevant MTS-BRB modelling, inferencing, and learning processes are also proposed. The key findings can be distilled into the three categories shown below:

(1) In order to find a solution to the combinatorial explosion problem, MTS was utilized to provide a panorama for displaying hierarchical BRB. It was also shown that the MTS-BRB could have a reliable size while handling complex problems when compared to the conventional BRB.

(2) For the dilemma in existing studies that the hierarchical BRB should be given by the expert in advance but it may be impossible for complex problems, a MTS-BRB modeling procedure was proposed to construct self-organized hierarchical BRB from the collected data of complex problems.

(3) In order to upgrade conventional BRB expert system as MTS-BRB expert system, a MTS-BRB inferencing and a MTS-BRB learning procedures were further proposed to optimize the parameters of the MTS-BRB and produce inferential output for replying given input data using the MTS-BRB.

(4) Five benchmark challenges regarding the risk assessment of R&D projects and UCI datasets were introduced for system development and comparison with the purpose of verifying the proposed MTS-BRB expert system. The outcomes showed that MTS-BRB expert system performance outperformed previously published studies in this area.

For this study, it should be noted that the data used in MTS-BRB modeling procedure is mainly based on continuous data. Future research in this area could include the use of MTS-BRB expert system for complex problems with discrete data. On a similar note, this study was limited to considering only single feature extraction technique, and only using subjective preference to determine the size and number of sub-BRB in MTS-BRB. Both of these topics contain numerous existing works which could be applied to this type of expert system, improving its interpretability, efficiency, and accuracy. There remains significant room for further research in the area of hierarchical BRB or FRB for complex problems modeling, especially in relation to refining the modeling process, and exploring additional structure optimization strategies.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Nos. 72001043 and 72001042), the Humanities and Social Science Foundation of the Ministry of Education of China (No. 20YJC630188), and the National Science Foundation of Fujian Province of China (Nos. 2020J05122 and 2022J01178).

References

Abudahab, K., Xu, D. L., & Chen, Y. W., (2016). A new belief rule base knowledge representation scheme and inference methodology using the evidential reasoning rule for evidence combination. *Expert Systems with Applications*, 51: 218-230.

Ahmed, T. U., Jamil, M. N., Hossain, M. S., Islam R. U., & Andersson, K., (2022). An Integrated Deep Learning and Belief Rule Base Intelligent System to Predict Survival of COVID-19 Patient under Uncertainty. *Cognitive Computation*, 14: 660-676.

Alcala-Fdez, J., Alcala, R., & Herrera, F., (2011). A Fuzzy Association Rule-Based Classification Model for High-Dimensional Problems with Genetic Rule Selection and Lateral Tuning. *IEEE Transactions on Fuzzy Systems*, 19(5): 857-872.

Aminravan, F., Sadiq, R., Hoorfar, M., Rodriguez, M.J., & Najjaran, H., (2015). Multi-level information fusion for patiotemporal monitoring in water distribution networks. *Expert Systems with Applications*, 42: 3813-3831.

Bart, K., (2018). Additive Fuzzy Systems: From Generalized Mixtures to Rule Continua. *International Journal of Intelligent Systems*. *33*(8): 1573-1623.

Cao, Y., Zhou, Z. J., Hu, C. H., He, W., & Tang, S. W., (2021). On the interpretability of belief rule-based expert systems. *IEEE Transactions on Fuzzy Systems*. 29(11): 3489-3505.

Cao, Y., Zhou, Z. J., Hu, C. H., Tang, S. W., & Wang, J., (2021). A New Approximate Belief Rule Base Expert System for Complex System Modelling. *Decision Support Systems*. 150: 113558.

Chang, L. L., Fu, C., Wu, Z. J., & Liu, W. Y., (2022). A Data-Driven Method using BRB with Data Reliability and Expert Knowledge for Complex Systems Modeling. *IEEE Transactions on Systems Man Cybernetics-Systems*, 52(11): 6729-6743.

Chang, L. L., Xu, X. J., Liu, Z. G., Qian, B., Xu, X. B., & Chen, Y. W., (2021) BRB Prediction with Customized Attributes Weights and Tradeoff Analysis for Concurrent Fault Diagnosis. *IEEE System Journal*. 15(1): 1179-1190.

Chang, L. L., Zhou, Y., Jiang, J., Li, M. J., & Zhang, X. H., (2013). Structure learning for belief rule base expert system: A comparative study. *Knowledge - Based Systems*. 39: 159-172.

Chang, L. L., Zhou, Z. J., Liao, H. C., Chen, Y. W., Tan, X., & Herrera, F., (2019). Generic Disjunctive Belief Rule Base Modeling, Inferencing, and Optimization. *IEEE Transactions on Fuzzy Systems*. 27(9): 1866-1880.

Chen, Y., Chen, Y. W., Xu, X. B., Pan, C. C., Yang, J. B., & Yang, G. K., (2015). A data-driven approximate causal inference model using the evidential reasoning rule. *Knowledge-Based Systems*, 88: 264-272.

Chen, Y. W., Yang, J. B., Xu, D. L., & Yang, S. L., (2013). On the inference and approximation properties of belief rule based systems. *Information Sciences*. 234: 121-135.

Chen, Y. W., Yang, J. B., Xu, D. L., Zhou, Z. J., & Tang, D. W., (2011). Inference analysis and adaptive training for belief rule based systems, *Expert Systems with Applications*. 38(10): 12845-12860.

Cheng, C., Qiao, X., Teng, W., Gao, M., Zhang, B., Yin, X., & Luo, H., (2020). Principal component analysis and belief-rule-base aided health monitoring method for running gears of high-speed train. *Science China Information Science*. 63(9): 1-3.

Cordón, O., Jesus, M. J. D., Herrera, F., & Lozano, M., (1999). MOGUL: A Methodology to Obtain Genetic Fuzzy Rule-Based Systems under the Iterative Rule Learning Approach. *International Journal of Intelligent Systems*, 14(11): 1123-1153.

Diao, H. Y., Lu, Y. F., Deng, A. S., Zou, L., & Li, X. F., (2022). Convolutional rule inference network based on belief rule-based system using an evidential reasoning approach. *Knowledge-Based Systems*, 237: 107713.

Du, Y. K., Han, S. H., Kim, H., & Part, H., (2009). Structuring the prediction model of project performance for international construction projects: A comparative analysis, *Expert Systems with Applications*. 36(2): 1961-1971.

Dua, D., & Graff, C., (2019). UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Sciences.

Elkano, M., Galar, M., Sanz, J., & Bustince, H., (2016). Fuzzy Rule-Based Classification Systems for multi-class problems using binary decomposition strategies: On the influence of n-dimensional overlap functions in the Fuzzy Reasoning Method. *Information Sciences*, 332: 94-114.

Fang, W. J., Gong, X. T., Liu, G. G., Wu, Y. J., & Fu, Y. G., (2020). A Balance Adjusting Approach of Extended Belief-Rule-Based System for Imbalanced Classification Problem. *IEEE Access*. 8: 41201-41212.

Fernández, A., Calderon, M., Barrenechea, E., Bustince, H., & Herrera, F., (2010). Solving multi-class problems with linguistic fuzzy rule based classification systems based on pairwise learning and preference relations. *Fuzzy Sets and Systems*, 161: 3064-3080.

Gao, F., Zhang, A., & Ma, J. W., (2021). A greedy belief rule base generation and learning method for classification problem, *Applied Soft Computing*. 98: 106856.

González, A., & Pérez, R., (1999). SLAVE: A Genetic Learning System Based on an Iterative Approach. *IEEE Transactions on Fuzzy Systems*,

7(2): 176-191.

Gu, X. & Angelov, P. P., (2020). Highly interpretable hierarchical deep rule-based classifier. *Applied Soft Computing*, 92: 106310.

He, W., Hu, G. Y., Zhou, Z. J., Qiao, P. L., Han, X. X. Qu, Y. Y., Wei, H., & Shi, C., (2018). A new hierarchical belief-rule-based method for reliability evaluation of wireless sensor network. *Microelectronics Reliability*, 87: 33-51.

Hossain, M. S., Rahaman, S., Kor, A., Andersson, K., & Pattinson, C., (2017). A Belief Rule Based Expert System for Datacentor PUE Prediction under Uncertainty. *IEEE Transactions on Sustainable Computing*, 2(2): 140-153.

Hu, G. Y., Zhou, Z. J., Hu, C. H., Zhang, B. C., Zhou, Z. G., Zhang, Y., & Wang, G. Z., (2020). Hidden behavior prediction of complex system based on time-delay belief rule base forecasting model. *Knowledge-Based Systems*. 203: 106147.

Hu, G. Y., Zhou, Z. J., Zhang, B. C., Yin, X. J., Gao, Z., & Zhou, Z. G., (2016). A Method for Predicting the Network Security Situation Based on Hidden BRB Model and Revised CMA-ES Algorithm. *Applied Soft Computing*, 48: 404-418.

Ishibuchi, H., Yamamoto, T., & Nakashima, T., (2005). Hybridization of Fuzzy GBML Approaches for Pattern Classification Problems. *IEEE Transactions on Systems Man Cybernetics-Part B: Cybernetics*, 35(2): 359-365.

Jiao, L. M., Denœux, T., & Pan, Q., (2016). A Hybrid Belief Rule-Based Classification System Based on Uncertain Training Data and Expert Knowledge. *IEEE Transactions on Systems Man Cybernetics: Systems*. 46(12): 1711-1723.

Kerr-Wilson, J., & Pedrycz, W., (2020). Generating a hierarchical fuzzy rule-based model. *Fuzzy Sets and Systems*, 381: 124-139.

Li, G. L., Zhou, Z. J., Hu, C. H., Chang, L. L., Zhou, Z. G., & Zhao, F. J., (2017). A new safety assessment model for complex system based on the conditional generalized minimum variance and the belief rule base. *Safety Science*. 93: 108-120

Li, J., Zhang, J., Ning, P., & Xiao, Q., (2020). Weighted Outlier Detection of High-Dimensional Categorical Data Using Feature Grouping. *IEEE Transactions on Systems Man Cybernetics Systems*. 50(11): 4295-4308.

Liu, J., Martínez, L., Calzada, A., & Wang, H., (2013). A novel belief rule base representation, generation and its inference methodology, *Knowledge-Based Systems*. 53: 129-141.

MacKay, D. J. C., (2003). Information Theory, Inference, and Learning Algorithms. Cambridge, U.K.: Cambridge Univ. Press.

Mansoori, E. G., Zolghadri, M. J., & Katebi, S. D., (2008). SGERD: A Steady-State Genetic Algorithm for Extracting Fuzzy Classification Rules From Data. *IEEE Transactions on Fuzzy Systems*, 16(4): 1061-1071.

Mendel, J. M. & Bonissone, P. P., (2021). Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems. IEEE Transactions on Fuzzy Systems, 29(12): 3579-3593.

Mir, F. A., & Pinnington, A. H., (2014). Exploring the value of project management: Linking project management performance and project success, *International Journal of Project Management*. 32(2): 202-217.

Mishra, A., Das, S. R., & Murray, J. J., (2016). Risk, process maturity, and project performance: an empirical analysis of us federal government technology projects. *Production and Operations Management*. 25(2): 210-232.

Mohagheghi, V., Mousavi, S. M., Vahdani, B., & Shahriari, M. R., (2017). R&D project evaluation and project portfolio selection by a new interval type-2 fuzzy optimization approach. *Neural Computing & Applications*, 28: 3869-3888.

Sachan, S., Yang, J. B., Xu, D. L., Benavides, D. E., & Li, Y., (2020). An explainable AI decision-support-system to automate loan underwriting, *Expert Systems with Applications*, 144: 113100.

Pedrycz, W., (1996). Classification of relational patterns as a decomposition problem. *Pattern Recognition Letter*, 17(1): 91-99.

Pedrycz, W., Al-Hmouz, R., Balamash, A. S., & Morfeq, A., (2015). Designing granular fuzzy models: A hierarchical approach to fuzzy modeling. *Knowledge-Based Systems*, 76: 42-52.

Sun, R., (1995). Robust reasoning: integrating rule-based and similarity-based reasoning. *Artificial Intelligence*,75: 241-295.

Takagi, T., & Sugeno, M., (1993). Fuzzy identification of systems and its applications to modeling and control. *Readings in Fuzzy Sets for Intelligent Systems*, 15(1): 387-403.

Wang, L. X., (2020). Fast Training Algorithms for Deep Convolutional Fuzzy Systems with Application to Stock Index Prediction. *IEEE Transactions on Fuzzy Systems*. 18(7): 1-14.

Wang, Y., Liu, H., Jia W., Guan S., Liu X., & Duan, X., (2022). Deep Fuzzy Rule-Based Classification System with Improved Wang-Mendel Method. *IEEE Transactions on Fuzzy Systems*, 30(8): 2957-2970.

Wang, Y. M., Yang, J. B., Xu, D. L., & Chin, K. S., (2009). Consumer preference prediction by using a hybrid evidential reasoning and belief rule-based methodology. *Expert Systems with Applications*. 36(4): 8421-8430.

Yang, J. B., Liu, J., Wang, J., Sii, H. S., & Wang, H. W., (2006). Belief rule-base inference methodology using the evidential reasoning approach - RIMER. *IEEE Transactions on Systems Man Cybernetics- Part A*. 36(2): 266-285.

Yang, J. B., Wang, Y. M., Xu, D. L., Chin, K. S., & Chatton, L., (2012). Belief rule-based methodology for mapping consumer preferences and

setting product targets. *Expert Systems with Applications*. 39(5): 4749-4759.

Yang, L. H., Liu, J., Wang, Y. M., & Martínez, L., (2021). A Micro-Extended Belief Rule-Based System for Big Data Multiclass Classification Problems. *IEEE Transactions Systems Man Cybernetics: Systems*. 51(1): 420-440.

Yang, L. H., Wang, Y. M., Chang, L. L., & Fu, Y. G., (2017). A disjunctive belief rule-based expert system for bridge risk assessment with dynamic parameter optimization model. *Computers & Industrial Engineering*. 113: 459-474.

Yang, L. H., Wang, Y. M., Liu, J., & Martínez, L., (2018). A joint optimization method on parameter and structure for belief-rule- based systems, *Knowledge - Based Systems*. 142: 220-240.

Yang, L. H., Ye, F. F., & Wang, Y. M., (2020). Ensemble belief rule base modeling with diverse attribute selection and cautious conjunctive rule for classification problems. *Expert Systems with Applications*. 146: 113161.

Yang, Y., Fu, C., Chen, Y. W., Xu, D. L., & Yang, S. L., (2016). A belief rule based expert system for predicting consumer preference in new product development. *Knowledge-Based Systems*. 94: 105-113.

Yang, Y., Wang, J., Wang, G., & Chen, Y. W., (2019). Research and development project risk assessment using a belief rule-based system with random subspaces. *Knowledge-Based Systems*. 178: 51-60.

Yet, B., Constantinou, A., Fenton, N., Neil, M., Luedeling, E., & Shepherd, K., (2016). A Bayesian network framework for project cost, benefit and risk analysis with an agricultural development case study. *Expert Systems with Applications*. 60: 141-155.

You, Y. Q., Sun, J. B., Chen, Y. W., Niu, C. Y., & Jiang, J., (2021). Ensemble Belief Rule-Based Model for complex system classification and prediction. *Expert Systems with Applications*. 2021, 164: 113952.

Zhang, B. C., Hu, G. Y., Zhou, Z. J., Zhang, Y. M., Qiao, P. L., & Chang, L. L., (2017). Network Instrusion Detection Based on Directed Acyclic Graph and Belief Rule Base. *ETRI Journal*, 39(4): 592-604.

Zhang, X., Onieva, E., Perallos, A., Osaba, E., & Lee V. C. S., (2014). Hierarchical fuzzy rule-based system optimized with genetic algorithms for short term traffic congestion prediction. *Transportation Research Part C: Emerging Technologies*, 43: 127-142.

Zhang, Z., Xu, X., Chen, P., Wu, X., Xu, X., Wang, G., & Dustdar, S., (2021). A novel nonlinear causal inference approach using vector-based belief rule base. *International Journal of Intelligent Systems*. 36: 5005-5027.

Zhou, Z. G., Liu, F., Li, L. L., Jiao, L. C., Zhou, Z. J., Yang, J. B., & Wang, Z. L., (2015). A cooperative belief rule based decision support system for lymph node metastasis diagnosis in gastric cancer. *Knowledge-Based Systems*. 2015, 85: 62-70.

Zhou, Z. J., Feng, Z. C., Hu, C. H., Han, X. X., Zhou, Z. G., Li, & G. L., (2019). A hidden fault prediction model based on the belief rule base with power set and considering attribute reliability. *Science China Information Sciences*, 62(10): 1-16.

Zhou, Z. J., Feng, Z. C., Hu, C. H., Hu, G. Y., He, W., & Han, X. X., (2020). Aeronautical relay health state assessment model based on belief rule base with attribute reliability. *Knowledge-Based Systems*. 197: 105869.

Zhou, Z. J., Hu, G. Y., Hu, C. H., Wen, C. L., & Chang, L. L., (2021). A Survey of Belief Rule-Base Expert System. *IEEE Transactions on Systems Man Cybernetics: Systems*. 51(8): 4944-4958.

Zhuang, J. H., Ye, J. F., Chen, N. N., Fang, W. J., Fan, X. C., & Fu, Y. G., (2021). Extended Belief Rule-Base Optimization Base on Clustering Tree and Parameter Optimization, *IEEE Access*, 9: 12533-12544.

Zielinski, K., & Laur, R., (2008). Stopping Criteria for Differential Evolution in Constrained Single-Objective Optimization, In: Chakraborty U.K. (eds) Advances in Differential Evolution, 143, Springer, Berlin, Heidelberg.