Int. J. Mach. Learn. & Cyber. (2018) 9:1941–1954
DOI 10.1007/s13042-017-0729-x

ORIGINAL ARTICLE

Decision-theoretic rough set model of multi-source decision systems

Binbin Sang[1] · Yanting Guo[2] · Derong Shi[1] · Weihua Xu[1]

Received: 4 February 2017 / Accepted: 16 September 2017 / Published online: 5 October 2017

Abstract Through the complementary integration of information from different sources, information fusion can improve the decision-making process in the increasing uncertain environments. How to make full use of the information from various sources to make decisions is a key problem in multi-source decision-theoretic. The more accurate and comprehensive information is, the easier the decision will be. Thus the uncertainty of decision-making is an objective criterion for evaluating the fusion effect. Therefore, three kinds of multi-source decision methods are proposed based on considering the uncertainty of decision-making process, which are the conditional entropy multi-source decision (CE-MSD) method, the decision support degree multi-source decision (DS-MSD) method and the mean multi-source decision (M-MSD) method. The CE-MSD method based on taking into account the uncertainty of each condition attribute for decisions aims to select the most reliable source for each attribute according to the conditional entropy, and then make final decisions under a new restructuring decision table. The DS-MSD method proposed by considering the uncertainty of condition attribute set for decisions aims to make the final decision through the decision support degree of all the sources to each object. The M-MSD method and the approximate precision index are introduced as reference standards to measure the effectiveness of CE-MSD and DS-MSD in the multi-source decision system. Meanwhile, three corresponding algorithms are designed to verify the effectiveness and feasibility of the proposed decision methods. Finally, in order to verify the validity of methods, approximation accuracies of CE-MSD, DS-MSD and M-MSD are compared in multi-source decision systems which are generated by adding Gauss noise and random noise to Data set downloaded from UCI. In sum, the decision theory of multi-source decision systems is a generalization of the decision-theoretic rough set, which is worthy of further research.

Keywords Decision-theoretic rough set · Multiple-source decision systems · Information fusion · Information system

✉ Weihua Xu
chxuwh@gmail.com; datongxuweihua@126.com

[1] School of Science, Chongqing University of Technology, 400054 Chongqing, People's Republic of China

[2] Faculty of Information Technology, Macau University of Science and Technology, Taipa, Macau, People's Republic of China

1 Introduction

With the development of information technology, data size is increasing and the amount of data is also growing. One of the most urgent problems is how to use data from multiple sources to make decisions. Through the integration of data from different sources, we can make up for the deficiency of the single data to achieve the mutual complement and mutual confirmation of a variety of data sources. In this way, it not only expands the application range of the data, but also improves the accuracy of the analysis. Therefore, it is especially important that taking full advantage of multi-source information decision-making.

Decision-theoretic rough set (DTRS) theory first proposed by Professor Yao [28]. DTRS is a prominent probabilistic rough set model, in which thresholds can be calculated by the decision risk minimization based on Bayesian decision theory [33], conditional probabilities can be estimated by Naive Bayesian model [31]. And the positive region, negative region, boundary region of probabilistic rough sets can be regarded as a application of three-way decisions theory [29, 30]. The DTRS model has solid theoretical foundation and practical value [6, 14, 16, 23, 27, 37]. There are a lot of relevant research. Based on DTRS, Li and Zhou studied multi-view decision rule extraction [38] and proposed a new decision model by considering different risk preferences of decision makers [18]. Liu et al. [11] discussed multiple-category three-way decisions on the basis of decision-theoretic rough sets. Yang and Yao put forward a multi-agent decision-theoretic rough set model [36]. Ma and Sun [21] systematically researched the decision-theoretic rough set model over double universes. Jia et al. [5] explored minimum cost attribute reduction in the decision-theoretic rough set model. Liang et al. [15] investigated triangular fuzzy decision-theoretic rough sets. Wei [26] proposed interval valued hesitant fuzzy uncertain linguistic aggregation operators in multiple attribute decision making. Qi et al. [24] proposed multiple attribute group decision making based on generalized power aggregation operators under interval-valued dual hesitant fuzzy linguistic environment. Ma et al. [20] proposed multiple attribute reduction models based on DTRS. Yu et al. [32] presented an automatic method to determine the number of clusters by using decision-theoretic rough sets. Deng et al. [3] provided a method for automatically calculating thresholds alpha and beta by means of combining information entropy. Liu et al. [12] raised a new classification method on account of regression analysis and decision-theoretic rough sets. Qian et al. [20] explored decision-theoretic rough sets under multi-granularity. In fact, there are many applications about DTRS, such as text classification [35], oil exploitation [17], policy decisions [13], web-based medical decision support systems [39], email filtering [40] and so on.

From the perspective of information sources, above researches are based on the single information source. In many circumstances, it is necessary that integrating all the information from diverse sources to make decisions. There are many studies on multi-source information fusion [25, 41]. In particular, Khan and Banerjee [9, 10] studied rough sets and notions of approximations based on views of membership of objects in multiple-source approximation systems (MSAS). Besides, Md and Khan [19] proposed a modal logic for multiple-source tolerance approximation spaces (MTAS) in view of the principle of only considering the information of sources about objects. Khaleghi et al. [8] studied a comprehensive review of the data fusion state of the art, exploring its conceptualizations, benefits, and challenging aspects, as well as existing methodologies. Yao et al. [34] studied advances in the field of Web information fusion and integration. Recently, Yao et al. [35] considered approaches to rough set approximations in a multigranulation space.

Then we will study decision-theoretic rough sets under multiple-source decision systems (MSDS) which have the same universe and attributes and different information functions (namely Isomorphic multiple-source decision systems). It should be pointed out that isomorphic multiple-source information systems refer to the same cardinality of the partition generated by attribute set on the universe in each decision system. For heterogeneous multiple-source information systems, we can find the ultimate goal as a middle bridge to establish the relationship between different sources. However, on this basis will be required to achieve a higher goal for isomorphic multiple-source information systems. Because the more information you have on the same thing, learned knowledge should be more accurate. Therefore, the research of multiple-source decision systems (MSDS) which have the same universe and attributes and different information functions is meaningful. The most important issue which how to make full use of the information provided by each source in multiple-source decision systems (MSDS). Two examples are introduced to highlight the motivation of multiple-source decisions.

1. A person suspected that she had lung cancer because she often cough and feels persistent chest pain. In order to diagnose, she went to three hospitals which have different precision instrument and equipment and professional medical level to do some related checks such as the imaging of chest X-ray and CT. If the final diagnosis of the three hospitals are the same, we can determine whether the person is suffering from lung cancer or not. When three hospital diagnosis results are not identical, how to determine the final diagnosis result needs to be further explored. For example, you can get three results of CT through the check from above hospitals. Which result of CT is more able to reflect the true state of her body and is more helpful to make the final diagnosis? These phenomena are thought-provoking.
2. The evaluation of scholarship need to consider many aspects annually, such as academic achievement, scientific research and moral performance. And each aspect may need to consider two semesters. For example, each student need to be considered academic performance of two semesters before appraisal organization makes the final decision. This may remind us that the final decision of some events need to consider the decisions in different stages.

In this paper, three kinds of multi-source decision methods are proposed by considering the uncertainty of decision making. The first method is the conditional entropy multi-source decision (CE-MSD) method, which fully takes into account the uncertainty of each condition attribute for decisions. The second method is the decision support degree multi-source decision (DS-MSD) method, which fully takes into account the uncertainty of condition attribute set for decisions. The third method is the most common mean fusion namely mean multi-source decision (M-MSD) method, which is mainly used as a reference standard to measure the effectiveness of the CE-MSD and DS-MSD methods. There are three motivations for studying this topic: (1) multiple-source information systems can provide more information, which is more accurate. Therefore, the research of multiple-source decision systems (MSDS) which have the same universe and attributes and different information functions is meaningful. (2) Multivariate fusion is a necessary method for multiple-source information systems. Therefore, the CE-MDS and M-MDS methods are proposed in this paper. (3) Multi-source decision-theoretic can be effectively applied in wider areas, such as, medical diagnosis, academic evaluation and risk investment et al. So, the research of multiple-source decision systems (MSDS) and decision-theoretic rough set model which are very meaningful.

The rest of this paper is organized as follows. Section 2 provides a review of basic concepts of decision systems, rough sets theory, decision-theoretic rough sets. In Sect. 3, firstly, the definition of the multi-source decision system (MSDS) is proposed. Then multi-source decision methods based on conditional entropy, decision support degree and mean method are proposed successively, which are the CE-MSD, the DS-MSD and the M-MSD method (a reference standard). Meanwhile, three corresponding algorithms are designed to verify the effectiveness and feasibility of the proposed methods. At the same time, the approximation accuracy of the universe U about a decision partition π_D is proposed under different methods. In Sect. 4, comparison of approximation accuracies of the CE-MSD, DS-MSD and M-MSD methods are made in different data sets to verify effectiveness of the CE-MSD and DS-MSD methods. Finally, Sect. 5 gets the conclusion.

2 Basic knowledge

In this section, some basic concepts about decision systems, rough set theory, decision-theoretic rough set theory are reviewed.

A decision system is a quadruple $S = (U, A, V, f)$,where U is a nonempty finite universe; $A = C \cup D$ is the union of condition attribute set C and decision attribute set D, and $C \cap D = \emptyset$; V is the union of attribute domains, i.e., $V = \cup_{a \in A} V_a$; $f : U \times A \to V$ is an information function, i.e., $\forall a \in A, x \in U$, that $f(x,a) \in V_a$,where $f(x, a)$ is the value of the object x under the attribute a. Generally, let $D = \{d\}$. Unless otherwise specified, all the decision systems in this paper are defined as above shown.

A new form of conditional entropy proposed by Dai [2] is a reasonable measure for the uncertainty of decision systems. Let $S = (U, C \cup D, V, f)$ be a decision system, where the universe $U = \{x_1, x_2, \dots, x_n\}$, and $U/D = \{Y_1, Y_2, \dots, Y_m\}$ is a partition of U. Then conditional entropy of D with respect to B ($B \subseteq C$) is defined by

$$H(D|B) = -\sum_{i=1}^{|U|} p([x_i]_B) \sum_{j=1}^{m} p(Y_j|[x_i]_B) logp(Y_j|[x_i]_B),$$

where $p([x_i]_B) = |[x_i]_B|/|U|$, $p(Y_j|[x_i]_B) = |[x_i]_B \cap Y_j|/|[x_i]_B|$ and $[x_i]_B = \{x_j|\forall a \in B, f(x_j,a) = f(x_i,a)\}$.

The approximation accuracy proposed by Pawlak [22] is used to measure a rough classification. The approximation accuracy offers the percentage of possible correct decisions when we classify objects by an attribute set R. Let $U/D = \{Y_1, Y_2, \dots, Y_m\}$ be a partition of the universe U. For an arbitrary attribute subset R of C, the R-lower and R-upper approximations of U / D are defined as

$$\underline{R}(U/D) = \underline{R}(Y_1) \cup \underline{R}(Y_2) \cup \cdots \cup \underline{R}(Y_m),$$
$$\overline{R}(U/D) = \overline{R}(Y_1) \cup \overline{R}(Y_2) \cup \cdots \cup \overline{R}(Y_m).$$

The approximation accuracy and the approximation roughness of U / D by R are defined as follows:

$$\alpha_R(U/D) = \frac{\sum_{Y_i \in U/D} |\underline{R}(Y_i)|}{\sum_{Y_i \in U/D} |\overline{R}(Y_i)|},$$
$$Roughness_R(U/D) = 1 - \alpha_R(U/D).$$

Here, if $\alpha_R(U/D) = 1$, then the decision system is consistent; otherwise it is inconsistent(nondeterministic and non definite). Especially, this article is carried out under the uncoordinated decision systems.

Given a decision system $S = (U, C \cup D, V, f)$, Yao provides a way about how to make decisions under minimum Bayesian expectation risk in decision-theoretic rough set model [28]. Based on the idea of three-way decisions, the decision-theoretic rough set uses a state set Ω and an action set A to describe the decision-making process. $\Omega = \{X, X^C\}$ indicating that an object is in a decision class X and not in X. The set of actions with respect to a state is given by $A = \{a_P, a_B, a_N\}$, where a_P, a_B and a_N represent three actions about deciding $x \in POS(X)$, deciding $x \in BND(X)$, and deciding $x \in NEG(X)$, respectively. Let λ_{PP}, λ_{BP} and λ_{NP} denote the costs caused by taking actions a_P, a_B and a_N, respectively, when an object belongs to X; and λ_{PN}, λ_{BN} and λ_{NN} denote the costs incurred for taking the same actions when the object does not belong to X.

Given the cost function, the expected cost associated with taking the particular actions for the objects in $[x]_R$ can be expressed as:

$$R(a_P|[x]_R) = \lambda_{PP}P(X|[x]_R) + \lambda_{PN}P(X^C|[x]_R);$$
$$R(a_B|[x]_R) = \lambda_{BP}P(X|[x]_R) + \lambda_{BN}P(X^C|[x]_R);$$
$$R(a_N|[x]_R) = \lambda_{NP}P(X|[x]_R) + \lambda_{NN}P(X^C|[x]_R).$$

where $[x]_R = \{y|f(x,a) = f(y,a), \forall a \in C\}$, and $P(X|[x]) = |X \cap [x]_R|/|[x]_R|$ represents condition probability of x with respect to X and $P(X^C|[x]_R) = 1 - P(X|[x]_R)$.

By the Bayesian decision procedure, the following minimum-risk decision rules can be obtained:

(P) If $R(a_P|[x]_R) \leq R(a_B|[x]_R)$ and $R(a_P|[x]_R) \leq R(a_N|[x]_R)$, then decide $x \in POS(X)$;
(B) If $R(a_B|[x]_R) \leq R(a_P|[x]_R)$ and $R(a_B|[x]_R) \leq R(a_N|[x]_R)$, then decide $x \in BND(X)$;
(N) If $R(a_N|[x]_R) \leq R(a_P|[x]_R)$ and $R(a_N|[x]_R) \leq R(a_B|[x]_R)$, then decide $x \in NEG(X)$.

Considering a reasonable hypothesis that the decision cost of $POS(X)$ is smallest and the decision cost of $POS(X)$ and $BND(X)$ are strictly smaller than the cost of $NEG(X)$ when $x \in X$, the reverse of the order of cost is used for $x \in X^C$, namely, $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$ and $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$. Then the Bayesian decision procedure leads to the following minimum-risk decision rules:

(P) If $P(X|[x]_R) \geq \alpha$ and $P(X|[x]_R) \geq \gamma$, then decide $x \in POS(X)$;
(B) If $P(X|[x]_R) \leq \alpha$ and $P(X|[x]_R) \geq \beta$, then decide $x \in BND(X)$;
(N) If $P(X|[x]_R) \geq \beta$ and $P(X|[x]_R) \leq \gamma$, then decide $x \in NEG(X)$.

Where the parameters α, β and γ are defined as:

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})};$$
$$\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})};$$
$$\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}.$$

If a cost function further satisfies the condition: $(\lambda_{NP} - \lambda_{BP})(\lambda_{PN} - \lambda_{BN}) > (\lambda_{BP} - \lambda_{PP})(\lambda_{BN} - \lambda_{NN}$, then we can get $0 \leq \beta < \gamma < \alpha \leq 1$. Then the DTRS model has decision rules as follows:

(P) If $p(X|[x]_R) \geq \alpha$, then decide $x \in POS(X)$;
(B) If $\beta < p(X|[x]_R) < \alpha$, then decide $x \in BND(X)$;
(N) If $p(X|[x]_R) \leq \beta$, then decide $x \in NEG(X)$.

On the one hand, we can get probabilistic approximations through the three decision rules, namely the upper and lower approximations of the DTRS model:

$$\underline{R}_{(\alpha,\beta)}(X) = \{x \in U|P(X|[x]_R) \geq \alpha\}.$$
$$\underline{R}_{(\alpha,\beta)}(X) = \{x \in U|P(X|[x]_R) \geq \alpha\}.$$

If $\underline{R}_{(\alpha,\beta)}(X) = \overline{R}_{(\alpha,\beta)}(X)$, then X is a definable set, otherwise X is a rough set. Here, $POS_{(\alpha,\beta)}(X) = \underline{R}_{(\alpha,\beta)}(X)$, $NEG_{(\alpha,\beta)}(X) =\sim \overline{R}_{(\alpha,\beta)}(X)BND_{(\alpha,\beta)}(X) = \overline{R}_{(\alpha,\beta)}(X) - \underline{R}_{(\alpha,\beta)}(X)$ are the positive region, negative region and boundary region, respectively.

By using the thresholds α and β, one can divide the universe U into three regions of a decision partition π_D based on (α, β):

$POS_{(\alpha,\beta)}(\pi_D|\pi_C) = \{x \in U|p(D_{max}([x]_C)) \geq \alpha\}$,
$BND_{(\alpha,\beta)}(\pi_D|\pi_C) = \{x \in U|\beta < p(D_{max}([x]_C)) < \alpha\}$,
$NEG_{(\alpha,\beta)}(\pi_D|\pi_C) = \{x \in U|p(D_{max}([x]_C)) \leq \beta\}$,
where $D_{max}([x]_C) = argmax_{Y_j \in \pi_D}\{|[x]_C \cap Y_j|/|[x]_C|\}$.

Therefore, the upper and lower approximations of the universe U about a decision partition π_D based on (α, β) are as follows:

$\overline{R}_{(\alpha,\beta)}(\pi_D|\pi_C) = \{x \in U|p(D_{max}([x]_C)) \geq \alpha\}$,
$\underline{R}_{(\alpha,\beta)}(\pi_D|\pi_C) = \{x \in U|p(D_{max}([x]_C)) > \beta\}$.

And the approximation accuracy of the universe U about a decision partition π_D can be defined as follows:

$$\alpha_R(U/D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D|\pi_C)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D|\pi_C)|}.$$

On the other hand, in view of the classification level of tolerance of decision-theoretic rough set model, all decision rules may bring corresponding loss due to its error rate [4, 7, 29]. So let $U/D = \{Y_1, Y_2, \dots, Y_m\}$, $P = P(D_{max}([x]_R)|[x]_R)$ where

$$D_{max}([x]_R) = argmax_{Y_j \in U/D}|[x]_R \cap Y_j|/|Y_j|$$

and $\lambda_{PP} = \lambda_{NN} = 0$. Decision loss of all rules can be get as follows:

positive rule loss: $(1 - P)\lambda_{PN}$;
boundary rule loss: $P\lambda_{BP} + (1 - P)\lambda_{BN}$;
negative rule loss: $P\lambda_{NP}$;

For a given decision system S, the decision cost of S is calculated as:

$$COST = COST_{POS} + COST_{BND} + COST_{NEG} = \sum_{P_i \geq \alpha}(1 - P_i)\lambda_{PN} + \sum_{\beta < P_j < \alpha} P_j\lambda_{BP} + (1 - P_j)\lambda_{BN} + \sum_{P_k \leq \beta} P_k\lambda_{NP}$$

where $p_i = p(D_{max}([x_i]_A)|[x_i]_A)$.

3 Multiple-source decision-theoretic rough set theory

The more accurate and comprehensive the information is, the easier the decision making is. How to make the final decision according to the information from multiple sources, there is not uniform standard. In this paper, we propose two views to make decisions. The first view is to integrate information and then make decisions. The second view is to make decisions on each source, and finally make the final decision. Then three kinds of multiple-source decision methods are proposed in the following, namely the conditional entropy multi-source decision (CE-MSD) method, the decision support degree multi-source decision (DS-MSD) method and the mean multi-source decision (M-MSD) method. The CE-MSD and M-MSD methods are based on the first view. The DS-MSD method is based on the second view.

The rapid development of information science and technology has given rise to an unprecedented volume of freely available, user-generated data. In particular, these data about the same thing are usually obtained from different information sources. In particular, making full use of the information from various sources is an important issue in making right decisions process. Multi-source information fusion is a momentous content in the field of information research. The integration of information from different sources can get more comprehensive information to make the right decisions. In this paper, we explore decision-theoretic rough sets under the multi-sources decision systems (MSDS) which have the same universe and attributes and different information functions. First of all, the definition of multi-sources decision systems is proposed.

Definition 3.1 A multi-sources decision system (MSDS) is a tuple $(U, \{S_i\}_{i\in N})$, where $S_i = (U, C \cup D, V_i, f_i)$ is a decision system, $N = \{1, 2, 3, \ldots\}$ is an initial segment of the positive integers set, which represents information sources. $\forall i \in N$, S_i represents the ith source of the multiple-source decision system (MSDS).

A brief description of notations in this section is made. Let $MS = (U, \{S_i\}_{i\in N})$ be a multiple-source decision system (MSDS), $\forall i \in N$, $S_i = (U, C \cup D, V_i, f_i)$ is a decision system, where $U/D = \{Y_1, Y_2, \ldots, Y_m\}$ for each source S_i are identical, and $N = \{1, 2, \ldots\}$ denote the number of information sources.

How to make full use of the information from different sources to make right decisions focuses on information fusion. The more accurate the information collected is, the smaller the uncertainty of decisions is. In order to get more comprehensive and accurate information to make right decisions, the reliability of each information source is considered by decision cost.

Definition 3.2 Let $MS = (U, \{S_i\}_{i\in N})$ be a multiple-source decision system (MSDS). The reliability of the information source S_i ($\forall i \in N$) in the MSDS can be defined as:

$$r(S_i) = 1 - COST_{S_i} \Big/ \sum_{i\in N} COST_{S_i}$$

where $COST_{S_i}$ denotes the decision cost of ith information source, the value of $COST_{S_i}$ can be calculated by the equation

$$COST_{S_i} = \sum_{P_{R_i} \ge \alpha} (1 - P_{R_i})\lambda_{PN} + \sum_{\beta < P_{R_i} < \alpha} P_{R_i}\lambda_{BP} + (1 - P_{R_i})\lambda_{BN} + \sum_{P_{R_i} \le \beta} P_{R_i}\lambda_{NP},$$

R_i is the indiscernibility relation generated by C in the ith information source of MSDS, and $P_{R_i} = max\{P(Y_1|[x]_{R_i}), P(Y_2|[x]_{R_i}), \ldots, P(Y_m|[x]_{R_i})\}$. $\lambda_{PN}, \lambda_{BP}, \lambda_{BN}, \lambda_{NP}$ are determined by the cost function given by experts, detailed information is shown in Table 1.

In Table 1, let λ_{PP}, λ_{BP} and λ_{NP} denote the cost incurred for taking actions a_P, a_B and a_N, respectively, when an object belongs to a decision class Y. And λ_{PN}, λ_{BN} and λ_{NN} denote the cost caused by taking the same actions when the object does not belong to Y. Correspondingly, there are $\lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$, $\lambda_{NN} \le \lambda_{BN} < \lambda_{PN}$.

It is important to point out that the reliability of each information source can be measured by different criteria which are decided by decision goal. Decision cost is being used as the criteria in this paper.

Based on the consideration of the uncertainty of decision making, two kinds of multi-source decision methods are discussed in this paper, which are conditional entropy decision fusion and support degree decision fusion.

3.1 The conditional entropy multi-source decision method

By taking into account the uncertainty of each condition attribute for decisions, the reliable source of each condition attribute is selected by a certain standard. The new form of conditional entropy proposed by Dai [2] is a reasonable uncertainty measure for a complete or incomplete decision system. The conditional entropy is more sensitive, particularly when the incomplete rate is at a high level. For one attribute (or attribute set), if the information of this attribute (or attribute set) gets coarser (there are more missing values on this attribute or attribute set), the conditional entropy of this attribute (or the attribute set) gets larger. As the division becomes finer, the value of the conditional entropy is smaller. The smaller the value of the conditional entropy

Table 1 The cost function

	Y (P)	Y^C (N)
a_P	λ_{PP}	λ_{PN}
a_B	λ_{BP}	λ_{BN}
a_N	λ_{NP}	λ_{NN}

is, the more important the attribute will be. Thus the conditional entropy can be used to evaluate the importance of attributes. Namely the uncertainty of a decision system can be measured by conditional entropy. Therefore, the conditional entropy is used as the selection criterion for information sources. In this section, under the multi-sources decision systems (MSDS), the CE-MSD method is a generalization of the DTRS-Model. We turn MSDS into a new restructuring decision table by using conditional entropy. And we made the right decision with DTRS-Model in the new decision table.

Definition 3.3 Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS). The reliability of the information source S_i ($\forall i \in N$) with respect to the attribute a in the MSDS can be defined as:

$$r(S_i|a) = r(S_i)/H_i(D|a)$$

where $r(S_i)$ is the reliability of the information source S_i ($\forall i \in N$) in the MSDS, and $H_i(D|a)$ is the conditional entropy of D with respect to a in the ith decision system of MS. And $H_i(D|a)$ can be calculated by the equation

$$H_i(D|a) = -\sum_{i=1}^{|U|} p_i([x_i]_a) \sum_{j=1}^{m} p_i(Y_j|[x_i]_a) log(p_i(Y_j|[x_i]_a))$$

where $p_i([x_i]_a) = |[x_i]_a|/|U|$ and $p_i(Y_j|[x_i]_a) = |[x_i]_a \cap Y|/|[x_i]_a|$.

After we select the most reliable source selection of each condition attribute, the MSDS can be turned into a new decision system S_0. According to the given cost function, the expected costs associated with taking different actions for objects in $[x]_{R_i}$ can be expressed as

$R(a_P|[x]_{R_i}) = \lambda_{PP} p_0(Y|[x]_{R_i}) + \lambda_{PN} p_0(Y|[x]_{R_i})$,
$R(a_B|[x]_{R_i}) = \lambda_{BP} p_0(Y|[x]_{R_i}) + \lambda_{BN} p_0(Y|[x]_{R_i})$,
$R(a_N|[x]_{R_i}) = \lambda_{NP} p_0(Y|[x]_{R_i}) + \lambda_{NN} p_0(Y|[x]_{R_i})$.

According to the Bayesian decision procedure, there are

(P) If $R(a_P|[x]_{R_i}) \le R(a_B|[x]_{R_i})$ and $R(a_P|[x]_{R_i}) \le R(a_N|[x]_{R_i})$, then decide $x \in POS(Y)$;
(B) If $R(a_B|[x]_{R_i}) \le R(a_P|[x]_{R_i})$ and $R(a_B|[x]_{R_i}) \le R(a_N|[x]_{R_i})$, then decide $x \in BND(Y)$;
(N) If $R(a_N|[x]_{R_i}) \le R(a_P|[x]_{R_i})$ and $R(a_N|[x]_{R_i}) \le R(a_B|[x]_{R_i})$, then decide $x \in NEG(Y)$.

In general, there are $\lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$, and $\lambda_{NN} \le \lambda_{BN} < \lambda_{PN}$. The decision rules can be restated as

(P) If $p_0(Y|[x]_{R_i}) \ge \alpha$ and $p_0(Y|[x]_{R_i}) \ge \gamma$, then decide $x \in POS(Y)$;
(B) If $p_0(Y|[x]_{R_i}) \le \alpha$ and $p_0(Y|[x]_{R_i}) \ge \beta$, then decide $x \in BND(Y)$;
(N) If $p_0(Y|[x]_{R_i}) \le \beta$ and $p_0(Y|[x]_{R_i}) \le \gamma$, then decide $x \in NEG(Y)$.

where the thresholds values α, β and γ are given by:

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})};$$
$$\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})};$$
$$\gamma = \frac{\lambda_{PN} - \lambda_{NN}}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}.$$

Moreover, when $(\lambda_{NP} - \lambda_{BP})(\lambda_{PN} - \lambda_{BN}) > (\lambda_{BP} - \lambda_{PP})(\lambda_{BN} - \lambda_{NN})$, there are $0 \le \beta < \gamma < \alpha \le 1$. In this case, above decision rules can be stated as

(P) If $p_0(Y|[x]_{R_i}) \ge \alpha$, then decide $x \in POS(Y)$;
(B) If $\beta < p_0(Y|[x]_{R_i}) < \alpha$, then decide $x \in BND(Y)$;
(N) If $p_0(Y|[x]_{R_i}) \le \beta$, then decide $x \in NEG(Y)$.

In this paper, the cost function satisfies the following conditions:

(1) $\lambda_{PP} \le \lambda_{BP} < \lambda_{NP}$, and $\lambda_{NN} \le \lambda_{BN} < \lambda_{PN}$;
(2) $(\lambda_{NP} - \lambda_{BP})(\lambda_{PN} - \lambda_{BN}) > (\lambda_{BP} - \lambda_{PP})(\lambda_{BN} - \lambda_{NN})$.

Therefore, the three regions of the universe about a decision partition π_D based on (α, β) are stated as

$POS_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_0(Y_1|[x]_{R_i}), p_0(Y_2, |[x]_{R_i})\}, \dots, p_0(Y_m|[x]_{R_i})\} \ge \alpha\}$,
$BND_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | \beta < max\{p_0(Y_1|[x]_{R_i}), p_0(Y_2 |[x]_{R_i})\}, \dots, p_0(Y_m|[x]_{R_i})\} < \alpha\}$,
$NEG_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_0(Y_1|[x]_{R_i}), p_0(Y_2 |[x]_{R_i})\}, \dots, p_0(Y_m|[x]_{R_i})\} \le \beta\}$,

Moreover, the upper and lower approximations of the universe U about a decision partition π_D based on (α, β) in the MSDS are as follows:

$\overline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | max\{p_0(Y_1|[x]_{R_i}), p_0(Y_2|[x]_{R_i})\}, \dots, p_0(Y_m|[x]_{R_i})\} > \beta\}$,
$\underline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | max\{p_0(Y_1|[x]_{R_i}), p_0(Y_2|[x]_{R_i})\}, \dots, p_0(Y_m|[x]_{R_i})\} \ge \alpha\}$.

And the approximation accuracy of the universe U about a decision partition π_D is as follows:

$$\alpha_R(\pi_D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D)|}.$$

Example 3.1 There are four decision sources about medical diagnosis, which constructs a multi-source decision system.

Table 2 A multi-source decision system

	1st source				2nd source				3rd source				4th source				
	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4	d
x_1	1	2	2	1	1	2	2	1	1	2	1	1	1	2	2	1	1
x_2	1	2	1	1	1	2	2	1	1	2	1	1	1	2	1	1	1
x_3	1	1	2	1	1	1	2	1	1	1	1	1	1	1	2	1	0
x_4	0	1	1	1	1	1	1	1	0	1	2	1	0	1	2	0	1
x_5	2	1	1	2	0	1	1	1	1	1	1	1	2	2	1	1	0
x_6	0	1	1	0	0	1	2	0	0	1	1	0	1	1	2	0	1
x_7	1	1	2	1	2	2	2	1	1	2	1	1	1	2	1	1	0
x_8	1	1	1	0	2	1	1	0	1	1	1	0	1	1	1	0	1
x_9	2	1	1	0	2	1	1	1	2	1	2	1	2	1	2	1	0
x_{10}	1	1	1	0	1	1	1	1	0	1	2	1	0	1	2	0	0

Table 3 The cost function

	Y (P)	Y^C (N)
a_P	0	36
a_B	8	4
a_N	24	0

Specific data information is shown in Table 2. And the cost function is shown in Table 3.

First of all, the decision cost of each information source can be obtained, namely

$COST_{S_1} = 12, COST_{S_2} = 12, COST_{S_3} = 32, COST_{S_4} = 24$. By the Definition 3.2, the reliability of the information source S_i $(i \in \{1,2,3,4\})$ in the MSDS can be obtained

$r(S_1) = 0.8500,\ r(S_2) = 0.8500,\ r(S_3) = 0.6000,\ r(S_4) = 0.7000$. Moreover, the conditional entropy of D with respect to a_i $(i \in \{1,2,3,4\})$ in each decision system can be calculated as follows

$H_1(D|a_1) = 3.6000,\quad H_1(D|a_2) = 6.1084,\quad H_1(D|a_3) = 5.6541, H_1(D|a_4) = 4.0274,$
$H_2(D|a_1) = 3.6538,\quad H_{21}(D|a_2) = 5.6541,\quad H_2(D|a_3) = 4.8548, H_2(D|a_4) = 6.1048;$
$H_3(D|a_1) = 4.4265,\quad H_3(D|a_2) = 5.6541,\quad H_3(D|a_3) = 5.6541, H_3(D|a_4) = 6.1048;$
$H_4(D|a_1) = 3.7059,\quad H_4(D|a_2) = 5.2000,\quad H_4(D|a_3) = 5.2000, H_4(D|a_4) = 4.6039.$

The reliability of the information source S_i $(i \in \{1,2,3,4\})$ with respect to the attribute a in the MSDS can be obtained, namely

$r(S_1|a_1) = 0.2361, r(S_1|a_2) = 0.1392, r(S_1|a_3) = 0.1503, r(S_1|a_4) = 0.2111;$
$r(S_2|a_1) = 0.2326, r(S_2|a_2) = 0.1503, r(S_2|a_3) = 0.1751, r(S_2|a_4) = 0.1392;$
$r(S_3|a_1) = 0.1355, r(S_3|a_2) = 0.1061, r(S_3|a_3) = 0.1061, r(S_3|a_4) = 0.0982;$
$r(S_4|a_1) = 0.1889, r(S_4|a_2) = 0.1346, r(S_4|a_3) = 0.1346, r(S_4|a_4) = 0.1520.$

Table 4 The new system after entropy fusion

U	a_1	a_2	a_3	a_4	d
x_1	1	2	2	1	1
x_2	1	2	2	1	1
x_3	1	1	2	1	0
x_4	0	1	1	1	1
x_5	2	1	1	2	0
x_6	0	1	2	0	1
x_7	1	2	2	1	0
x_8	1	1	1	0	1
x_9	2	1	1	0	0
x_{10}	1	1	1	0	0

Therefore, the most reliability source of the attribute a_1 is S_1, the most reliability source of the attribute a_2 is S_2, the most reliability source of the attribute a_3 is S_2, and the most reliability source of the attribute a_4 is S_1. So the new decision table can be obtained, specific data information is shown in Table 4.

By the information of Table 4, there are $U/D = \{D_1, D_2\}$, where $D_1 = \{x_1, x_2, x_4, x_6, x_8\}$, $D_2 = \{x_3, x_5, x_7, x_9, x_{10}\}$, and

$p_0(D_1|x_1) = 2/3,\ p_0(D_2|x_1) = 1/3;\ p_0(D_1|x_2) = 2/3,\ p_0(D_2|x_2) = 1/3;$
$p_0(D_1|x_3) = 0,\ p_0(D_2|x_3) = 1;\ p_0(D_1|x_4) = 1,\ p_0(D_2|x_4) = 0;$
$p_0(D_1|x_5) = 0,\ p_0(D_2|x_5) = 1;\ p_0(D_1|x_6) = 1,\ p_0(D_2|x_6) = 0;$
$p_0(D_1|x_7) = 2/3,\ p_0(D_2|x_7) = 1/3;\ p_0(D_1|x_8) = 1/2,\ p_0(D_2|x_8) = 1/2;$
$p_0(D_1|x_9) = 0,\quad p_0(D_2|x_9) = 1;\quad p_0(D_1|x_{10}) = 1/2,\quad p_0(D_2|x_{10}) = 1/2.$

Therefore, the three regions of the universe about a decision partition π_D based on (α, β) are stated as

$POS_{(\alpha,\beta)}(\pi_D|\pi_{R_0}) = \{x_1, x_2, x_7, x_8, x_{10}\}$, $BND_{(\alpha,\beta)}(\pi_D|\pi_{R_0}) = \{x_3, x_4, x_5, x_6, x_9\}$, $NEG_{(\alpha,\beta)}(\pi_D|\pi_{R_0}) = \emptyset$.

In order to verify the feasibility and effectiveness of the conditional entropy multi-source decision method, we design Algorithm 1 and analyze the time complexity of Algorithm 1. In steps 2–17, we compute all decision cost of information sources, and the time complexity is $O(s \times |U| \times m)$. Step 18, calculate the reliability of all information sources about all attributes, and the time complexity is $O(s)$. Steps 19–27, calculate the reliability of all information sources about single attribute, and the time complexity is $(s \times |C| \times |U|)$. Steps 28–30, we choose reliability information source of each condition attribute, and a new table can be obtained by the selected reliability information source, and the time reliability is $O(|C| \times |U|)$. Steps 31–39, calculate the upper and lower approximations of the decision partition π_D about the new table, and the time complexity is $O(|U|)$.

```
Algorithm 1: The conditional entropy multi-source decision method
   Input  : α, β, λ_PN, λ_BP, λ_BN, λ_NP, MS = (U, {S_i}_{i∈{1,2,···,s}}), U/D = {Y_1, Y_2, ···, Y_m}
   Output : The approximation accuracy of U/D
1  begin
2    for i = 1 : s do                          /* the s is the number of information sources */
3      COST_{S_i} ← 0;          /* initialize the decision cost of information source S_i as 0 */
4      for x_i ∈ U do                        /* x_i is an object of the i^th information source */
5        p_i ← 0;
6        for j = 1 : m do
7          p_i ← max(p_i, |[x_i]_{R_i} ∩ Y_j|/|[x_i]_{R_i}|);   /* R_i is the equivalent relation of S_i */
8        end
9        if p_i ≥ α then             /* p_i ≥ α is the judgement condition of positive region */
10         COST_{S_i} ← COST_{S_i} + (1 − p_i)λ_PN;
11       else if p_i > β then        /* p_i > β is the judgement condition of boundary region */
12         COST_{S_i} ← COST_{S_i} + P_iλ_BP + (1 − P_i)λ_BN;
13       else                        /* p_i ≤ β is the judgement condition of negative region */
14         COST_{S_i} ← COST_{S_i} + P_iλ_NP;
15       end
16     end
17   end
18   r(S_i) ← 1 − (COST_{S_i} / Σ_{i=1}^{s} COST_{S_i});               /* r(S_i) is the reliability of S_i */
19   for i = 1 : s do
20     for a ∈ C do
21       H_i(D|a) ← 0;               /* H_i(D|a) is the conditional entropy of D about a in S_i */
22       for x_i ∈ U do
23         H_i(D|a) ← H_i(D|a) − (|[x_i]_a|/|U|) × Σ_{j=1}^{m} p_i(Y_j|[x_i]_a)log(p_i(Y_j|[x_i]_a));
24       end
25       r(S_i|a) ← r(S_i)/H_i(D|a);         /* r(S_i|a) is the reliability of S_i about a in the MS */
26     end
27   end
28   for a ∈ C do
29     k ← arg max_{i∈{1,2,···s}} (r(S_i|a)); S_0(a) ← S_k(a);        /* S_k(a) is all the value of a in the S_k */
30   end
31   for x ∈ U do
32     \overline{R_0}_{(α,β)}(π_D) ← φ, \underline{R_0}_{(α,β)}(π_D) ← φ;
33     if P_0 > β then
34       \overline{R_0}_{(α,β)}(π_D) ← \overline{R_0}_{(α,β)}(π_D) ∪ {x};   /* P_0 is the maximum conditional probability of decision
         classes and equivalent class[x]_{R_0} */
35     end
36     if P_0 ≥ α then
37       \underline{R_0}_{(α,β)}(π_D) ← \underline{R_0}_{(α,β)}(π_D) ∪ {x};
38     end
39   end
     return : α_R(π_D) = |\underline{R_0}_{(α,β)}(π_D)| / |\overline{R_0}_{(α,β)}(π_D)|;
                                   /* α_R(π_D) is the approximation accuracy of U/D in the MS */
40 end
```

3.2 The decision support degree multi-source decision method

In many circumstances, the information from different sources may have different decisions on the same thing. What we need to do is integrate these decisions directly. By taking into account the uncertainty of condition attribute set for decisions, the support degree multi-source decision method (DS-MSD) is proposed to make the final decision by the support degree of information sources to the same decision. Firstly, the definition of decision support degree is proposed.

Definition 3.4 Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS). The decision support degree of an object x be belonging to the positive, negative and boundary region of a decision partition π_D in the MSDS can be defined as:

$$d_{(POS_{(\alpha,\beta)}(\pi_D))}(x) = |S_i : x \in POS_{(\alpha,\beta)}(\pi_D|\pi_{R_i})|/|N|,$$

$$d_{(BND_{(\alpha,\beta)}(\pi_D))}(x) = |S_i : x \in BND_{(\alpha,\beta)}(\pi_D|\pi_{R_i})|/|N|,$$

$$d_{(NEG_{(\alpha,\beta)}(\pi_D))}(x) = |S_i : x \in NEG_{(\alpha,\beta)}(\pi_D|\pi_{R_i})|/|N|.$$

In the above formula, $|\bullet|$ denotes the cardinality of a set, and $POS_{(\alpha,\beta)}(\pi_D|\pi_{R_i})$, $BND_{(\alpha,\beta)}(\pi_D|\pi_{R_i})$, and $NEG_{(\alpha,\beta)}(\pi_D|\pi_{R_i})$ can be calculated by

$POS_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_i(Y_1|[x]_{R_i}), p_i(Y_2|[x]_{R_i})\}, \ldots, p_i(Y_m|[x]_{R_i})\} \geq \alpha\}$,

$BND_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | \beta < max\{p_i(Y_1|[x]_{R_i}), p_i(Y_2|[x]_{R_i})\}, \ldots, p_i(Y_m|[x]_{R_i})\} < \alpha\}$,

$NEG_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_i(Y_1|[x]_{R_i}), p_i(Y_2|[x]_{R_i})\}, \ldots, p_i(Y_m|[x]_{R_i})\} \leq \beta\}$.

It is obvious that $0 \leq d_{(POS_{(\alpha,\beta)}(\pi_D))}(x) \leq 1$, $0 \leq d_{(BND_{(\alpha,\beta)}(\pi_D))}(x)) \leq 1$ and $0 \leq d_{(NEG_{(\alpha,\beta)}(\pi_D))}(x) \leq 1$. Besides, there are $d_{(POS_{(\alpha,\beta)}(\pi_D))}(x) + d_{(BND_{(\alpha,\beta)}(\pi_D))}(x) + d_{(NEG_{(\alpha,\beta)}(\pi_D))}(x) = 1$.

Considering different risk attitudes of decision makers, we propose general multi-source decision model. So a decision level λ ($\lambda \in (0.5, 1]$)is given to make the final decision.

Definition 3.5 Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS). The upper and lower approximations of a decision partition π_D in the MSDS can be defined as:

$$\overline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | d_{(NEG_{(\alpha,\beta)}(\pi_D)}(x) < \lambda\},$$

$$\underline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | d_{(POS_{(\alpha,\beta)}(\pi_D)}(x) \geq \lambda\},$$

where λ is the decision level.

It is easy to prove that the lower approximation is contained in the upper approximation. Accordingly, the positive, negative and boundary region of a decision partition π_D are defined as follows:

$POS_{(\alpha,\beta)}(\pi_D) = \underline{R}_{(\alpha,\beta)}(\pi_D)$, $NEG_{(\alpha,\beta)}(\pi_D) = \sim \overline{R}_{(\alpha,\beta)}(\pi_D)$, $BND_{(\alpha,\beta)}(\pi_D) = \overline{R}_{(\alpha,\beta)}(\pi_D) - \underline{R}_{(\alpha,\beta)}(\pi_D)$.

And the approximation accuracy of the universe U about a decision partition π_D is as follows:

$$\alpha_R(\pi_D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D)|}.$$

By the definition of support degree and upper and lower approximations, the following conclusions hold trivially.

Proposition 3.1 *Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS), $\lambda \in (0.5, 1]$. The upper and lower approximations of a decision partition π_D can be expressed* as:

$$\overline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U : |S_i : x \in \overline{R}_{i(\alpha,\beta)}(\pi_D|\pi_{R_i})|/|N| > 1 - \lambda\},$$
$$\underline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U : |S_i : x \in \underline{R_i}_{(\alpha,\beta)}(\pi_D|\pi_{R_i})|/|N| \geq \lambda\},$$

where $\overline{R}_{i(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_i(Y_1|[x]_{R_i}), p_i(Y_2|[x]_{R_i})\}, \ldots, p_i(Y_m|[x]_{R_i})\} > \beta\}$ *and* $\underline{R_i}_{(\alpha,\beta)}(\pi_D|\pi_{R_i}) = \{x \in U | max\{p_i(Y_1|[x]_{R_i}), p_i(Y_2|[x]_{R_i})\}, \ldots, p_i(Y_m|[x]_{R_i})\} \geq \alpha\}$ *denote the upper and lower approximations of a decision partition π_D in the ith information source of MSDS.*

By observing the Proposition 3.1, a new perspective of multi-source fusion can de obtained. Generalized multigranulation idea can be introduced into the multi-source information fusion. That is to say, a multiple-source decision system can be regarded as a knowledge base, and each information source can be seen as a knowledge granularity. Therefore, the second method of multi-source decisions has certain theoretical basis.

Proposition 3.2 *Let $MS = (U, \{S_i\}_{i \in N})$ be a Multiple-source decision system (MSDS), $U/D = \{Y_1, Y_2, \ldots, Y_m\}$, $\lambda_1, \lambda_2 \in (0.5, 1]$ and $\lambda_1 < \lambda_2$. It is true that*

$$\alpha_R^{\lambda_1}(U/D) > \alpha_R^{\lambda_2}(U/D)$$

where $\alpha_R^{\lambda}(U/D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D)|}$.

Proof When $\lambda_1 < \lambda_2$, it is true that $\overline{R}^{\lambda_1}_{(\alpha,\beta)}(\pi_D) \subseteq \overline{R}^{\lambda_2}_{(\alpha,\beta)}(\pi_D)$ and $\underline{R}^{\lambda_1}_{(\alpha,\beta)}(\pi_D) \supseteq \underline{R}^{\lambda_2}_{(\alpha,\beta)}(\pi_D)$. The conclusion $\alpha_R^{\lambda_1}(U/D) > \alpha_R^{\lambda_2}(U/D)$ holds trivially

By observing the Proposition 3.2, we can know that classification approximation accuracy of the second kind of multi-source decision methods varies with the value of λ. The decision level λ can be set according to the actual need. With the increase of the decision level λ, classification approximation accuracy will decrease.

In order to verify the feasibility and effectiveness of the decision support degree multi-source decision method, we design Algorithm 2 and analyze the complexity of Algorithm 2. In steps 2–13, we compute the positive and negative regions of the decision partition π_D in information sources, and the time complexity is $O(s \times |U|)$. The steps 14 and 15, initialize the upper and lower approximations of π_D in the *MS*, and the time complexity is $O(1)$. Steps 16–25, calculate the upper and lower approximations of the decision partition π_D and the time complexity is $O(|U|)$.

Algorithm 2: The decision support degree multi-source decision method

```
Input  : α, β, λ, MS = (U, {S_i}_{i∈{1,2,⋯,s}}), U/D = {Y_1, Y_2, ⋯, Y_m}
Output : The approximation accuracy of U/D
1  begin
2      for i = 1 : s do
3          POS_(α,β)(π_D|π_{R_i}) ← φ;
4          NEG_(α,β)(π_D|π_{R_i}) ← φ;   /* POS_(α,β)(π_D|π_{R_i}) and NEG_(α,β)(π_D|π_{R_i}) are the positive and
           negative regions of the decision partition π_D in information source S_i */
5          for x ∈ U do
6              p_i(x) = max{p(Y_1|[x]_{R_i}), p(Y_2|[x]_{R_i}), ⋯, p(Y_m|[x]_{R_i})};  /* P_i(x) is the maximum conditional
               probability of decision classes and equivalent class[x]_{R_i} */
7              if p_i(x) ≥ α then
8                  POS_(α,β)(π_D|π_{R_i}) ← POS_(α,β)(π_D|π_{R_i}) ∪ {x};
9              else if p_i(x) < β then
10                 NEG_(α,β)(π_D|π_{R_i}) ← NEG_(α,β)(π_D|π_{R_i}) ∪ {x};
11             end
12         end
13     end
14     R̄_(α,β)(π_D) ← φ;          /* R̄_(α,β)(π_D) is the upper approximation of π_D in the MS */
15     R_(α,β)(π_D) ← φ;          /* R_(α,β)(π_D) is the lower approximation of π_D in the MS */
16     for x ∈ U do
17         d_(POS_(α,β)(π_D))(x) ← |{S_i|x ∈ POS_(α,β)(π_D|π_{R_i})}|/s;
18         d_(NEG_(α,β)(π_D))(x) ← |{S_i|x ∈ NEG_(α,β)(π_D|π_{R_i})}|/s;   /* d_(POS_(α,β)(π_D))(x) and d_(NEG_(α,β)(π_D))(x) are the
           decision support degree of the positive and negative regions */
19         if d_(NEG_(α,β)(π_D))(x) < λ then
20             R̄_(α,β)(π_D) ← R̄_(α,β)(π_D) ∪ {x};          /* λ is the decision level */
21         end
22         if d_(POS_(α,β)(π_D))(x) ≤ λ then
23             R_(α,β)(π_D) ← R_(α,β)(π_D) ∪ {x};
24         end
25     end
       return : α_R(π_D)) = |R_(α,β)(π_D)| / |R̄_(α,β)(π_D)|;
                              /* α_R(π_D)) is the approximation accuracy of U/D in the MS */
26 end
```

3.3 The mean multi-source decision method

Because the noise is generally obey the normal distribution, so mean value fusion is one of the most common methods of information fusion. It takes attribute value of each object as the basic point, and finally takes the average value of the basic point as the result of the fusion. Then the decision of MSDS is made on the new restructuring decision system. In this paper, M-MSD is introduced as a reference standard to measure the effectiveness of CE-MSD and DS-MSD.

Definition 3.6 Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS). The value of an object x under the attribute a in the MSDS can be defined as:

$f(x, a) = \sum_{i \in N} f_i(x, a)/|N|$,

where $f_i(x, a)$ denotes the value of the object x under the attribute a in the ith information source S_i of *MS*.

After the mean fusion, a new table S_0 can be obtained. Let R_0 be the equivalent relation generated by the condition attribute set of S_0.

Let $MS = (U, \{S_i\}_{i \in N})$ be a multiple-source decision system (MSDS). The upper and lower approximations of a decision partition π_D based on (α, β)in the MSDS can be defined as:

$$\overline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | max\{p(Y_1|[x]_{R_0}), p(Y_2|[x]_{R_0})\}, \dots, p(Y_m|[x]_{R_0})\} > \beta\},$$
$$\underline{R}_{(\alpha,\beta)}(\pi_D) = \{x \in U | max\{p(Y_1|[x]_{R_0}), p(Y_2|[x]_{R_0})\}, \dots, p(Y_m|[x]_{R_0})\} \geq \alpha\}.$$

And the approximation accuracy of the universe U about a decision partition π_D is as follows:

$$\alpha_R(\pi_D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D)|}.$$

In order to verify the feasibility and effectiveness of the CE-MSD and DS-MSD methods, the method of M-MSD is introduced as a reference standard. We design Algorithm 3 (Algorithm of M-MSD) and analyze the time complexity of Algorithm 3. In steps 2–10, we compute mean value of each object about each attribute under all information sources, and the time complexity is $O(|U| \times |C| \times s)$. Steps 11–22, calculate the upper and lower approximations of the decision partition π_D, and the time complexity is $O(m \times |U|)$.

Algorithm 3: The mean multi-source decision method

Input : $\alpha, \beta, \lambda, MS = (U, \{S_i\}_{i \in \{1,2,\cdots,s\}}), U/D = \{Y_1, Y_2, \cdots, Y_m\}$
Output : The approximation accuracy of U/D

1 **begin**
2 **for** $x \in U$ **do**
3 **for** $a \in C$ **do**
4 $temp \leftarrow 0$;
5 **for** $i = 1 : s$ **do**
6 $temp \leftarrow temp + S_i(x, a)$; /* $S_i(x, a)$ is the value of x under a in the S_i */
7 **end**
8 $S_0(x, a) \leftarrow \lfloor \frac{temp}{s} + 0.5 \rfloor$; /* $\lfloor \bullet \rfloor$ indicates rounding down */
9 **end**
10 **end**
11 **for** $j = 1 : m$ **do**
12 $\overline{R}_{(\alpha,\beta)}(\pi_D) \leftarrow \phi, \underline{R}_{(\alpha,\beta)}(\pi_D) \leftarrow \phi$; /* $\overline{R}_{(\alpha,\beta)}(\pi_D)$ and $\underline{R}_{(\alpha,\beta)}(\pi_D)$ are the upper and lower approximations of π_D in the $MSDS$ */
13 **for** $x \in U$ **do**
14 $p_0(x) = max\{p(Y_1|[x]_{R_0}), p(Y_2|[x]_{R_0}), \cdots, p(Y_m|[x]_{R_0})\}$; /* R_0 is the equivalent relation of the table S_0 obtained by mean fusion */
15 **if** $p_0(x) > \beta$ **then**
16 $\overline{R}_{(\alpha,\beta)}(\pi_D) \leftarrow \overline{R}_{(\alpha,\beta)}(\pi_D) \cup \{x\}$;
17 **end**
18 **if** $p_0(x) \geq \alpha$ **then**
19 $\underline{R}_{(\alpha,\beta)}(\pi_D) \leftarrow \underline{R}_{(\alpha,\beta)}(\pi_D) \cup \{x\}$;
20 **end**
21 **end**
22 **end**
return : $\alpha_R(\pi_D) = \frac{|\underline{R}_{(\alpha,\beta)}(\pi_D)|}{|\overline{R}_{(\alpha,\beta)}(\pi_D)|}$;
/* $\alpha_R(\pi_D)$ is the approximation accuracy of U/D in the $MSDS$ */
23 **end**

4 Comparison and analysis

In this paper, three kinds of multi-source decision methods are proposed, namely the conditional entropy multi-source decision (CE-MSD) method, the decision support degree multi-source decision (DS-MSD) method ,and the mean multi-source decision (M-MSD) method. The CE-MSD take every condition attribute as a basic point. Then the reliable source of each condition attribute is selected by conditional entropy. Finally, we make decisions on the new restructuring decision system can obtained after fusion. The DS-MSD take every decision system as a basic point. Based on the decision degree and the decision level, the decision is made in the multi-source decision system (MSDS). Besides, mean value fusion is one of the most common methods of information fusion. The M-MSD takes attribute value of each object as the basic point, and then takes the average value of the basic point as the result of the fusion. It is necessary to point out that the M-MSD method is introduced as reference standards to measure the effectiveness of CE-MSD and DS-MSD in the multi-source decision system (MSDS).

In order to evaluate the effectiveness of CE-MSD and DS-MSD, a series of experiments are conducted to show their classify advantage compared with mean multi-source decision method in a multi-source decision system. It is well known that directly download multi-source data sets is very difficult from the network. Therefore, this paper obtains multi-source data sets by adding Gauss noise and random noise to each original data set. We download four original data sets from the machine learning data repository, University of California at Irvine (http://archive.ics.uci.edu/ml/dat-asets.html). Detailed information is shown in Table 5. These experiments are implemented through using Visual C++ 6.0 and performed on a personal computer with an Intel Core i3-370, 2.40 GHz CPU, 2.0 GB of memory, and 32-bit Windows 7.

Firstly, the method of obtaining multi-source data sets is introduced, which is to add Gauss noise and random noise to the original data set. Let $MS = (U, \{S_i\}_{i \in \{1,2,\dots,s\}})$ be a multiple-source decision system (MSDS) constructed by the original decision table S'. First of all, s numbers $(g_1, g_2, \dots, g_s)$ are generated and these numbers all obey the $N(0, \sigma)$ distribution, where σ is the standard deviation. The method of adding Gauss noise to the original data is as follows:

$$f_i(x, a) = \begin{cases} f'(x, a) + g_i, & if\ (0 \leq f'(x, a) + g_i \leq 1); \\ f'(x, a), & else. \end{cases}$$

where $f'(x, a)$ represents the value of object x under attribute a in the original system s', $f_i(x, a)$ represents the value of object x under attribute a in the ith $(i \in N)$ decision system ofMS.

The method of adding random numbers to the original data is similar, and the specific process is as follows:

$$f_j(x, a) = \begin{cases} f'(x, a) + e_j, & if\ (0 \leq f'(x, a) + e_j \leq 1); \\ f'(x, a), & else. \end{cases}$$

Table 5 Experiment data sets

No.	Data set name	Abbreviation	Objects	Attributes	Decision classes	Elements of MS
1	*Balance Scale*	BS	625	4	3	25,000
2	*Wine Quality – red*	WQ-r	1599	12	6	191,880
3	*Wine Quality – white*	WQ-w	4898	12	7	587,760
4	*Page – blocks*	P-b	5473	11	5	602,030

where $f'(x, a)$ represents the value of object x under attribute a in the original system s', $f_j(x, a)$ represents the value of object x under attribute a in the jth $(j \in N)$decision system of MS, and $(e_1, e_2, \dots, e_s)$ are random numbers which are all between $-e$ and e, where e is a random error threshold.

Secondly, Gauss noise is added to 40% objects which is randomly selected from the original system s' and random noise is added 20% objects that is randomly selected from the rest of the original system s'. Finally, a multi-source decision system $MS = (U, \{S_i\}_{i\in\{1,2,\dots,s\}})$ can be obtained.

Table 6 The approximation accuracies of DS-MSD about different λ

λ	0.6	0.7	0.8	0.9	1.0
Balance-Scale	0.450	0.435	0.323	0.208	0.059
Wine quality-red	0.344	0.344	0.342	0.317	0.276
Wine quality-white	0.428	0.426	0.416	0.390	0.345
Page-blocks	0.849	0.835	0.794	0.647	0.535

According to Proposition 3.2, the approximation accuracy of DS-MSD dose not increase with the increase of the value of λ. In order to verify the fact, an illustrative experiment is conducted under multi-source data sets "Balance-Scale", "Wine quality-red", "Wine quality-white" and "Page-blocks". Accuracy results of DS-MSD are shown in Table 6.

Then in order to verify the validity of the CE-MSD method and DS-MSD methods in classification approximation accuracy, other fusion method M-MSD is introduced to compare. Then in order to test the effectiveness of the CE-MSD method and DS-MSD methods proposed by us, approximation accuracies of the three methods are compared in different data sets, which are shown in Table 7. Considering the fact that the approximation accuracy of DS-MSD is the highest when $\lambda = 0.6$ by Proposition 3.2. Need to point out that the approximation accuracy of DS-MSD is calculated under the condition $\lambda = 0.6$ in the Table 7. According to the user's different requirements, the decision level λ can take different values.

Table 7 Approximation accuracies of DS-MSD, CE-MSD, M-MSD under different data sets

No.	Balance-Scale			Wine quality-red			Wine quality-white			Page-blocks		
	DS	CE	M	DS	CE	M	DS	CE	M	DS	CE	M
1	0.435	0.468	0.350	0.352	0.299	0.066	0.444	0.405	0.127	0.845	0.897	0.785
2	0.457	0.423	0.416	0.345	0.296	0.083	0.426	0.401	0.140	0.845	0.836	0.779
3	0.467	0.454	0.364	0.343	0.290	0.073	0.457	0.424	0.152	0.814	0.881	0.762
4	0.474	0.416	0.333	0.354	0.302	0.095	0.477	0.432	0.154	0.851	0.836	0.827
5	0.461	0.455	0.356	0.342	0.307	0.097	0.467	0.423	0.165	0.826	0.854	0.818
6	0.476	0.480	0.384	0.365	0.316	0.115	0.470	0.429	0.174	0.831	0.866	0.821
7	0.487	0.470	0.329	0.336	0.293	0.112	0.491	0.443	0.178	0.848	0.835	0.827
8	0.501	0.445	0.337	0.353	0.316	0.099	0.480	0.418	0.180	0.863	0.850	0.817
9	0.524	0.477	0.371	0.370	0.346	0.110	0.498	0.438	0.169	0.870	0.850	0.771
10	0.516	0.438	0.390	0.370	0.332	0.102	0.509	0.447	0.187	0.879	0.857	0.812
11	0.491	0.444	0.376	0.346	0.319	0.104	0.508	0.449	0.186	0.885	0.865	0.818
12	0.526	0.463	0.349	0.336	0.304	0.094	0.500	0.447	0.190	0.887	0.873	0.812
13	0.533	0.505	0.380	0.349	0.325	0.121	0.514	0.465	0.193	0.889	0.875	0.779
14	0.561	0.518	0.359	0.383	0.348	0.127	0.515	0.461	0.201	0.887	0.871	0.811
15	0.530	0.482	0.388	0.345	0.320	0.122	0.516	0.467	0.205	0.890	0.873	0.771
16	0.552	0.504	0.351	0.365	0.346	0.133	0.521	0.462	0.204	0.887	0.881	0.775
17	0.541	0.475	0.376	0.345	0.317	0.136	0.524	0.463	0.218	0.890	0.879	0.784
18	0.558	0.499	0.396	0.374	0.356	0.145	0.518	0.468	0.220	0.889	0.877	0.767
19	0.533	0.518	0.330	0.346	0.329	0.156	0.538	0.492	0.223	0.889	0.880	0.772
20	0.547	0.531	0.361	0.372	0.350	0.150	0.514	0.469	0.236	0.889	0.880	0.767

The advantage of CE-MSD and DS-MSD methods in classification is visually shown through Fig. 1. For example, either the approximation accuracy of CE-MSD or the approximation accuracy of DS-MSD are higher than the accuracy of M-MSD under data sets “Wine Quality-red” and “Wine Quality-white”. Either the approximation accuracy of CE-MSD or the approximation accuracy of DS-MSD is not lower than the accuracy of M-MSD under data sets “Balance Scale” and “Wine Quality-white”. Moreover, the accuracy of DS-MSD is not lower than the accuracy of CE-MSD in most cases. A more intuitive comparison of the three methods is given in Table 8. Among them, “≻” means better. For example, the first number in Table 8, 95% indicates that 95% of approximation accuracy of DS-MSD is better than of CE-MSD in Balance-Scale’s data. As you can see in Fig. 1 and Table 8, the DS-MSD method and the CE-MSD method are superior to the M-MSD method. So, in general, we choose the DS-MSD method or the CE-MSD method. If we get unreliable sources, then we can choose the CE-MSD method. If we want to avoid loss of source information, we can choose the DS-MSD method.

5 Conclusions

With the development of information technology, taking full use of the information from multiple sources is the key to solve the practical problems. Through information fusion, more comprehensive information can be obtained to make decisions. In this paper, three kinds of multi-source decisions are proposed, namely the conditional entropy multi-source decision method (CE-MSD), the decision support degree multi-source decision method (DS-MSD) and the mean multi-source decision method (M-MSD). Related concepts

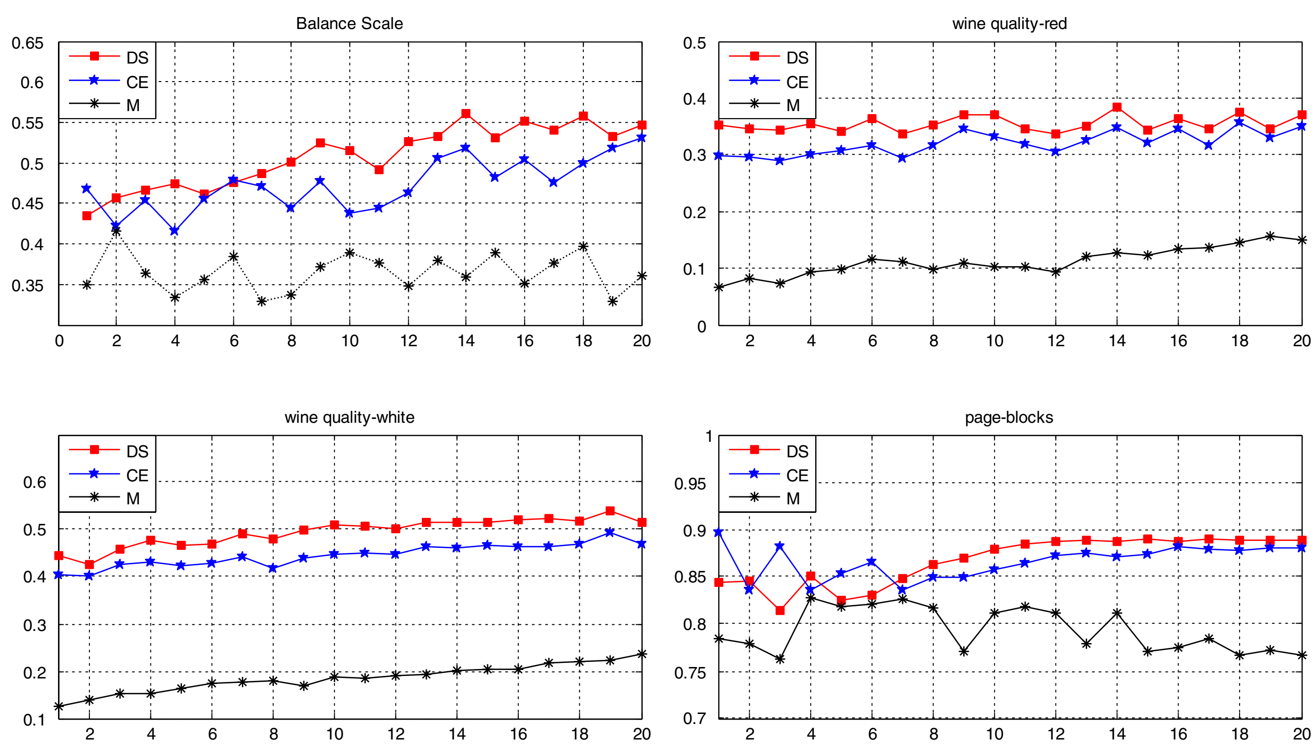

Fig. 1 Approximation accuracies of DS-MSD, CE-MSD, M-MSD under different data sets

Table 8 Comparison of the three methods

	DS ≻ CE (%)	DS ≻ M (%)	CE ≻ DS (%)	CE ≻ M (%)	M ≻ DS (%)	M ≻ CE (%)
Balance-Scale	95	100	5	100	0	0
Wine quality-red	100	100	0	100	0	0
Wine quality-white	100	100	0	100	0	0
Page-blocks	80	100	20	100	0	0

of the three kinds of multi-source decision-theoretic rough sets are described. Especially, three corresponding algorithms are designed to verify the effectiveness and feasibility of the proposed decision methods. Then the classify advantage of two methods (DS-MSD and CE-MSD) are verified by comparing with the approximate accuracy of M-MSD in multi-source decision systems which are generated by adding Gauss noise and random noise to Data set downloaded from UCI. Multi-source decision-theoretic rough set theory is a desirable research direction for decision-theoretic rough set. This paper just provides a framework of decision-theoretic rough sets in multi-source decision systems. More approaches to integration will be studied in the future.

Acknowledgements This work is supported by the Macau Science and Technology Development Fundation (No. 081/2015/A3). Natural Science Foundation of China (Nos. 61105041, 61472463, 61402064), National Natural Science Foundation of CQ CSTC (Nos. cstc2013j-cyjA40051, cstc2015jcyjA40053), and Graduate Innovation Foundation of Chongqing University of Technology (Nos. YCX2015227, YCX2014236, YCX2016227), and Graduate Innovation Foundation of CQ (No. CYS17281).

References

1. Azam N, Yao JT (2014) Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. Int J Approx Reason 55(1):142–155
2. Dai JH, Wang WT, Xu Q (2013) An uncertainty measure for incomplete decision tables and its applications. Trans Cybern 43(4):1277–1289
3. Deng XF, Yao YY (2014) A multifaceted analysis of probabilistic three-way decisions. Fundam Inf 132(3):291–313
4. Jia XY, Li WW, Shang L, Chen JJ (2011) An optimization viewpoint of decision-theoretic rough set model. Lect Notes Comput Sci 6954:457–465
5. Jia XY, Li WH, Tang ZM, Shang L (2013) Minimum cost attribute reduction in decision-theoretic rough set models. Inf Sci 219:151–167
6. Jia XY, Shang L, Zhou XZ (2012) Three-way decisions theory and its applications. Nanjing University Press, Nanjing
7. Jia XY, Tang ZM, Liao WH, Shang L (2014) On an optimization representation of decision-theoretic rough set model. Int J Approx Reason 55(1):156–166
8. Khaleghi B, Khamis A, Karry FO, Razavi SN (2013) Multisensor data fusion: a review of the state-of-the-art. Inf Fusion 14(1):28–44
9. Khan MA, Banerjee M (2010) A study of multiple-source approximation systems. Lect Notes Comput Sci 12:46–75
10. Khan MA, Banerjee M (2010) A perference-based multiple-source rough set model. Lect Notes Comput Sci 6068:247–256
11. Liu D, Li TR, Li HX (2012) A multiple-category classification approach with decision-theoretic rough sets. Fundam Inf 115(2–3):173–188
12. Liu D, Li TR, Liang DC (2014) Incorporating logistic regression to decision-theoretic rough sets for classifications. Int J Approx Reason 55(1):197–210
13. Liu D, Li TR, Liang DC (2012) Three-way government decision analysis with decision-theoretic rough sets. Int J Comput Intell Syst 20(1):119–132
14. Liu D, Li TR, Miao DQ (2013) Three-way decisions and granular computing. Science Press, Beijing
15. Liang DC, Liu D, Pedrycz W, Hu P (2013) Triangular fuzzy decision-theoretic rough sets. Int J Approx Reason 54(8):1087–1106
16. Li W, Miao DQ, Wang WL, Zhang N (2010) Hierarchical rough decision-theoretic framework for text classification. IEEE Int Conf Cogn Inf 9:484–489
17. Liu D, Yao YY, Li TR (2011) Three-way investment decisions with decision-theoretic rough sets. Int J Comput Intell Syst 4(1):66–74
18. Li HX, Zhou XZ (2011) Risk decision making based on decision-theoretic rough set: a three-way view decision model. Int J Comput Intell Syst 4(1):1–11
19. Md AK, Ma MH (2011) A modal logic for multiple-source tolerance approximation spaces. Indian Conf Logic and ITS Applica 6521:124–136
20. Ma XA, Wang GY, Yu H, Li TR (2014) Decision region distribution preservation reduction in decision-theoretic rough set model. Inf Sci 278:614–640
21. Ma WM, Sun BZ (2012) On relationship between probabilistic rough set and Bayesian risk decision over two universes. Int J Gen Syst 41(3):225–245
22. Pawlak Z (1982) Rough sets. Int J Comput Inf Sci 11(5):341–356
23. Qian YH, Zhang H, Sang YL, Liang JY (2014) Multigranulation decision-theoretic rough sets. Int J Approx Reason 55(1):225–237
24. Qi XW, Liang CY, Zhang JL (2016) Multiple attribute group decision making based on generalized power aggregation operators under interval-valued dual hesitant fuzzy linguistic environment. Int J Mach Learn Cybern 7(6):1147–1193
25. Tang YQ, Fan M, Li JH (2016) An information fusion technology for triadic decision contexts. Int J Mach Learn Cybern 7(1):13–24
26. Wei GW (2016) Interval valued hesitant fuzzy uncertain linguistic aggregation operators in multiple attribute decision making. Int J Mach Learn Cybern 7(6):1093–1114
27. Yao YY (2012) An outline of a theory of three-way decisions. In: The 8th international conference on rough sets and current trends in computing, vol 181, no. 6, pp 1–17
28. Yao YY (2007) Decision-theoretic rough set models. In: Proceeding of rough sets and knowledge technology, lecture notes in artificial intelligence, vol 4481, pp 1–12
29. Yao YY (2010) Three-way decisions with probabilistic rough sets. Inf Sci 180(3):341–353
30. Yao YY (2011) The superiority of three-way decisions in probabilistic rough set models. Inf Sci 181(6):1080–1096
31. Yao YY, Zhou B (2010) Naive Bayesian rough sets. In: The 5th international conference on rough sets and knowledge technology, pp 719–726
32. Yu H, Liu ZG, Wang GY (2014) An automatic method to determine the number of clusters using decision-theoretic rough set. Int J Approx Reason 55(1):101–115
33. Yao YY, Wong SKM, Lingras P (1990) A decision-theoretic rough set model. In: The 5th international symposium on methodologies for intelligent systems, vol 5, pp 17–25
34. Yao JT, Raghavan VV, Wu Z (2008) Web information fusion: a review of the state of the art. Inf Fusion 9(4):446–449
35. Yao YY, She YH (2016) Rough set models in multigranulation spaces. Inf Sci 327:40–56
36. Yang XP, Yao JT (2012) Modelling multi-agent three-way decisions with decision-theoretic rough sets. Fundam Inf 115(2–3):157–171
37. Yang HL, Guo ZL (2015) Multigranulation decision-theoretic rough sets in incomplete information systems. Int J Mach Learn Cybern 6(6):1005–1018
38. Zhou XZ, Li HX (2009) A multi-view decision model based on decision-theoretic rough set. In: The 4th international conference

on rough sets and knowledge technology, lecture notes in computer science, vol 5589, pp 650–657
39. Zhang Y, Yao JT (2014) Determining three-way decision regions with gini coefficients. In: The 9th international conference on rough sets and current trends in computing, lecture notes in computer science, vol 8536, pp 160–171
40. Zhou B, Yao YY, Luo JG (2014) Cost-sensitive three-way email spam filtering. J Intell Inf Syst 42(1):19–45
41. Zhang SP, Feng T (2016) Optimal decision of multi-inconsistent information systems based on information fusion. Int J Mach Learn Cybern 7(4):563–572