International Journal of Approximate Reasoning 82 (2017) 119–137

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Local multigranulation decision-theoretic rough sets ☆

Yuhua Qian [a,b,*], Xinyan Liang [a,b], Guoping Lin [c], Qian Guo [a,b], Jiye Liang [b]

[a] School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China
[b] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006 Shanxi, China
[c] School of Mathematics and Statistics, Minnan Normal University, Zhangzhou, 363000, Fujian, China

A R T I C L E I N F O

Article history:
Received 28 April 2016
Received in revised form 12 December 2016
Accepted 14 December 2016
Available online 23 December 2016

Keywords:
Decision-theoretic rough set
Limited labeled data
Local rough set
Multigranulation
Rough set

A B S T R A C T

Multigranulation rough sets (MGRSs) where a target concept is approximated by granular structures induced by multiple binary relations have been applied successfully in many domains but they are still affected by two issues. First, similar to other rough set models, calculating the approximation of a target set is extremely time-consuming for larger scale data. Second, MGRSs comprise a supervised learning method, so they often require a large amount of labeled data. However, in the era of big data, labeling all data is almost infeasible in some cases. In this study, to address these issues, we propose the combination of local rough sets with multigranulation decision-theoretic rough sets to obtain local multigranulation decision-theoretic rough sets (LMG-DTRSs) as a semi-unsupervised learning method. We also explore a number of important properties of LMG-DTRSs. In addition, we verify the efficiency of a concept approximation algorithm designed with LMG-DTRS based on theoretical and experimental analyses. Furthermore, we present two types of local MGRS frameworks under PRSs and variable precision rough sets, where the relationships between them and the LMG-DTRS model are also discussed. We show that many local MGRS models can be derived from the LMG-DTRS framework.

1. Introduction

Rough set theory was proposed by Pawlak in 1982 [24] and it has become an extremely useful tool for handling and modeling the vagueness and uncertainty of data. We do not need to have any transcendental knowledge regarding the data when utilizing rough set model to handle data [25]. Thus, this method has been applied in many areas, such as feature selection [30,48,56], rule extraction [2,9], uncertainty reasoning [27,39], granular computing [32], pattern recognition [38, 42], classification [59], and knowledge discovery [1,20,57,58].

Many extensions of the rough set model have been proposed to meet various requirements, which can be classified into two types depending on whether the decision risk is considered: rough set models without decision risk and rough set models with decision risk. The former type includes tolerance rough sets [46], dominance rough sets [5], neighborhood rough sets [8], and fuzzy rough sets [23]. These models do not consider the decision risk so they have no capacity for fault tolerance and they are extremely sensitive to noisy data. The second type includes decision-theoretic rough sets (DTRSs)

☆ This paper is part of the virtual special issue on tri-partition, edited by Davide Ciucci and Yiyu Yao.

* Corresponding author at: School of Computer and Information Technology, Shanxi University, Taiyuan, 030006 Shanxi, China.
E-mail addresses: jinchengqyh@126.com (Y. Qian), xinyanliang@email.sxu.edu.cn (X. Liang), guoplin@163.com (G. Lin), czguoqian@163.com (Q. Guo), ljy@sxu.edu.cn (J. Liang).

http://dx.doi.org/10.1016/j.ijar.2016.12.008
0888-613X/

[12–14,41,50,55], the probabilistic rough set model [51,52,61], the Bayesian rough set model [40,54], the variable precision rough set model [62], game-theoretic rough sets [7], parameterized rough set models [6], and double-quantitative rough set models [3,60] which are robust when dealing with noisy data.

In rough set theory, concept approximation using two subsets known as the lower and upper approximations is a cardinal task. It has been applied successfully in image segmentation [21,22,26]. Peters et al. [26] proposed an image segmentation approach where a set of pixel values were partitioned into bins that represent equivalence classes using a new indiscernibility relation based on k-means clustering of pixel values. Then, a form of upper and lower approximation specialized relative to sets of pixel values. Mohabey et al. [21] proposed an interesting method for color image segmentation introducing histon that correlates with upper approximation of a set such that all elements belonging to this set are classified as possibly belonging to the same segment or segments showing similar color value. According to how many granular structures are used for concept approximation, rough sets can be divided into two classes: single granulation rough sets and multigranulation rough sets (MGRSs). In single granulation rough sets, the concept approximation is calculated by a single relation (a single granular structure). By contrast, a target set is approximated by multiple equivalence relations (multiple granular structures) in MGRSs. The MGRS was first proposed by Qian in 2006 [29] and it has became a major research area in rough set theory. There are many real-life applications of the MGRSs, such as multi-source information systems, distributive information systems, and multiagent systems [28]. For a multi-source information system, each source used by the information system can be regarded as a granular structure. From this viewpoint, a multi-source information system can also be called a multigranulation (MG) information system. In theory, MGRSs have been generalized into fuzzy rough sets [49], DTRSs [28,37,47], variable precision rough sets [4], neighborhood rough sets [19], and other types in recent years. Moreover, Wu et al. proposed a granular computing method that uses multi-scale data measured at different levels of granulation [44]. Liang et al. [15] also presented an efficient feature selection algorithm for large-scale data based on a MG strategy. Moreover, She et al. [39] studied the structure of MGRSs. To extend the theory of DTRSs, Li et al. [10] investigated MG-DTRSs in ordered information systems. Recently, Li et al. [11] explored the relationship between MGRSs and concept lattices based on rule acquisition. In 1989, Rasiowa et al. [35] attempted to allow the total agents to reach conclusions based on a consensus, i.e., by agreeing to common statements. Rauszer [36] proposed a complete formal system for reasoning based on incomplete information in a multi-agent system. These tasks can be regarded as MG fusion problems in MGRSs.

In the present study, we combine the optimistic MGRS [33] and DTRS [55] in the same representative rough set model, which we refer to as the global optimistic (GO) MG-DTRS (GOMG-DTRS) [28]. Throughout this study, we assume that (U, AT) is an approximation space, where $R_1, R_2, ..., R_m \subseteq AT$ are m equivalence relations on U, U is a nonempty finite object set, R_i is an equivalence relation on U, $U/R_i = \{[x]_{R_i} : x \in U\}$ is an equivalence class induced by R_i, and $0 \leq \beta < \alpha \leq 1$. Then, $\forall X \subseteq U$, the lower and upper approximations of X are defined as follows:

$$\underline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \geq \alpha \vee P(X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, x \in U\}, \tag{1}$$

$$\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\ \beta)}(X) = U - \{x : P(X|[x]_{R_1}) \leq \beta \wedge P(X|[x]_{R_2}) \leq \beta \wedge \cdots \wedge P(X|[x]_{R_m}) \leq \beta, x \in U\}. \tag{2}$$

From the definition given above, we can observe two limitations of the global MG-DTRS, as follows.

(1) Computational inefficiency

Similar to other rough set models, the MG-DTRS is highly time-consuming because it is necessary to scan all the objects when approximating a concept. Its time complexity is $\mathcal{O}(n^2)$ without pre-ranking and $\mathcal{O}(n \log n)$ with pre-ranking. In particular, in the era of big data, the number of data objects is huge. The existing methods based on classical rough sets also have difficulty dealing with huge-scale data. Thus, reducing the computational time is our first motivation.

(2) Semi-supervised property of big data

The increasing scale of data makes it impossible to label all of the data in some cases, whereas collecting them is relatively easy, so big data possess a semi-supervised property, i.e., some data are labeled whereas others are not. These types of data are often represented as a data table, as shown in Table 1 [34], where $U = \{x_1, x_2, \cdots, x_n\}$ is the set of all objects, $U' = \{x_1, x_2, \cdots, x_p\}$ is an object set labeled with class labels $\{d_1, d_2, \cdots, d_r\}$, and $U - U' = \{x_{p+1}, x_{p+2}, \cdots, x_n\}$ is an unlabeled object set. However, existing rough sets-based models often require large amounts of labeled data when they are used for analyzing rough data. Thus, even if we obtain a large-scale data set comprising the objects from U, we can only use the granules induced by objects from the set U' to compute the lower and upper approximations of a target concept. Thus, much of the information related to objects in the set $U - U'$ is wasted. Thus, when calculating the lower and upper approximations of a target concept, making full use of the information related to the objects in $U - U'$ by applying the semi-supervised concept to reconstruct MG-DTRSs is the second motivation of this study, i.e., building models based on MG-DTRSs to handle data characterized by the semi-supervised property of big data.

According to the discussion above, it is necessary to develop generalized MG-DTRSs called local MG-DTRSs (LMG-DTRS) in order to improve the global MG-DTRSs by addressing the two disadvantages described above based on theoretical and experimental analyses.

Table 1
Data table containing a limited number of labeled objects.

Objects	x_1	$\cdots$	$\cdots$	$\cdots$	$\cdots$	x_p	$\cdots$	x_{n-1}	x_n
a_1	$a_1(x_1)$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$a_1(x_p)$	$\cdots$	$a_1(x_{n-1})$	$a_1(x_n)$
a_2	$a_2(x_1)$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$a_2(x_p)$	$\cdots$	$a_2(x_{n-1})$	$a_2(x_n)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
a_k	$a_k(x_1)$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$a_k(x_p)$	$\cdots$	$a_k(x_{n-1})$	$a_k(x_n)$
Class labels	d_1	d_1	d_2	$\cdots$	d_r	d_r	$***$	$*$	$*$

The remainder of this paper is organized as follows. First, some basic concepts related to Pawlak's rough sets (PRSs), DTRSs, and MGRSs are briefly reviewed in Section 2. Section 3 defines LMG-DTRSs and their properties are considered. In Section 4, we present two types of local MGRS models and the relationships between them. Finally, we give our conclusions and suggestions for further research.

2. Preliminary knowledge about rough sets

In this section, we briefly review some basic concepts related to PRSs, DTRSs, local rough sets (LRSs), MGRSs, and MG-DTRSs.

2.1. PRSs

In rough set theory, real-world data are represented by an information system. An information system is a tuple: $S = (U, AT, \{V_a | a \in AT\}, \{f_a | a \in AT\})$, where U is a finite nonempty set of objects, AT is a finite nonempty set of attributes, V_a is a nonempty set of values of $a \in AT$, and $f_a : U \to V_a$ is an information function, which maps an object from U to exactly one value in V_a.

When a decision is added to an information system, we refer to it as a decision information system, which is given by $S = (U, AT \cup D, \{V_a | a \in AT\}, \{f_a | a \in AT\})$, where AT is a set of condition attributes characterizing the objects and D is a set of decision attributes indicating the classes of objects.

Each nonempty subset $B \subseteq AT$ determines an indiscernibility relation, which is defined as $R_B = \{(x, y) \in U \times U \mid f_a(x) = f_a(y), \forall a \in B\}$.

The relation R_B partitions U into some equivalence classes given by $U/R_B = \{[x]_B | x \in U\}$, where $[x]_B = \{y \in U | (x, y) \in R_B\}$.

For any $X \subseteq U$, the sets $\underline{R_B} X = \cup\{Y \in U/IND(B) \mid Y \subseteq X\}$ and $\overline{R_B} X = \cup\{Y \in U/IND(B) \mid Y \cap X \neq \phi\}$ are called the lower and the upper approximations of X with respect to B, respectively.

The *area of uncertainty or boundary region* is $Bn(X) = \overline{R_B} X - \underline{R_B} X$.

2.2. DTRS

In order to consider the decision risk or enhance the robustness (to noisy data) of a rough-set-based decision, Yao proposed the DTRS model [55]. We briefly review some basic concepts regarding the DTRS model with an equivalence relation.

In the Bayesian decision procedure, a finite set of states can be written as $\Omega = \{\omega_1, \omega_2, \cdots, \omega_s\}$, and a finite set of m possible actions can be denoted by $A = \{a_1, a_2, \cdots, a_r\}$. Let $P(\omega_j|\mathbf{x})$ be the conditional probability of an object $\mathbf{x}$ being in state ω_j given that the object is described by $\mathbf{x}$. Let $\lambda(a_i|\omega_j)$ denote the loss, or cost, for taking action a_i when the state is ω_j, then the expected loss associated with taking action a_i is given by

$$R(a_i|\mathbf{x}) = \sum_{j=1}^{s} \lambda(a_i|\omega_j) P(\omega_j|\mathbf{x}). \tag{3}$$

An example is provided as follows to better understand formula (3). For example, in medical decision making, a subject denoted by $\mathbf{x}$ has a medical examination. The set of states is given by $\Omega = \{Health, Ill\}$, where *Health* and *Ill* represent the two states of being healthy and ill, respectively. In addition, the set of actions given by $A = \{Treatment,\ No\ treatment\}$ indicates that the doctors take measures for the subject. Let $\lambda(Treatment|Health)$ denote the loss incurred for taking action *Treatment* when the subject is healthy, $\lambda(Treatment|ill)$ denote the loss incurred for taking the same action when the subject is ill, $\lambda(No\ treatment|Health)$ denote the loss incurred for taking action *No treatment* when the subject is healthy, and $\lambda(No\ treatment|ill)$ denote the loss incurred for taking action *No treatment* when the subject is ill. These states and losses are shown in Table 2. In general, the losses induced by different actions are provided by experts. Losses may be estimated using techniques such as cost-effectiveness analysis and cost-benefit analysis [52]. $P(Health|\mathbf{x})$ and $P(Ill|\mathbf{x})$ are the probabilities that the subject is healthy and ill, respectively. The expected losses $R(Treatment|\mathbf{x})$ and $R(No\ treatment|\mathbf{x})$ associated with taking *Treatment* and *No treatment*, respectively, can be written as:

Table 2
Data set description.

	Health	Ill
Treatment	$\lambda(Treatment\|Health)$	$\lambda(Treatment\|Ill)$
No treatment	$\lambda(No\ Treatment\|Health)$	$\lambda(No\ Treatment\|Ill)$

$$R(Treatment|\mathbf{x}) = \lambda(Treatment|Health)P(Health|\mathbf{x}) + \lambda(Treatment|Ill)P(Ill|\mathbf{x}),$$
$$R(No\ treatment|\mathbf{x}) = \lambda(No\ treatment|Health)P(Health|\mathbf{x}) + \lambda(No\ treatment|Ill)P(Ill|\mathbf{x}).$$

The losses may be interpreted based on the harm caused by treating patients who do not have the disease and the benefits obtained by treating patients with the disease.

In classical rough set theory, the approximation operators partition the universe into three disjoint classes: $POS(A)$, $NEG(A)$, and $BND(A)$. By using the conditional probability $P(X|[x])$, the Bayesian decision procedure can decide how to assign x to these three disjoint regions. The expected loss $R(a_i|[x])$ associated with taking each of the individual actions can be expressed as [53,55]

$$\begin{aligned} R(a_1|[x]) &= \lambda_{11}P(X|[x]) + \lambda_{12}P(X^c|[x]), \\ R(a_2|[x]) &= \lambda_{21}P(X|[x]) + \lambda_{22}P(X^c|[x]), \\ R(a_3|[x]) &= \lambda_{31}P(X|[x]) + \lambda_{32}P(X^c|[x]), \end{aligned} \tag{4}$$

where $\lambda_{i1} = \lambda(a_i|X)$, $\lambda_{i2} = \lambda(a_i|X^c)$, and $i = 1, 2, 3$. When $\lambda_{11} \le \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \le \lambda_{32} < \lambda_{12}$, the Bayesian decision procedure leads to the following minimum-risk decision rules [50]:

(P) If $P(X|[x]) \ge \gamma$ and $P(X|[x]) \ge \alpha$, then decide $POS(X)$;
(N) If $P(X|[x]) \le \beta$ and $P(X|[x]) \le \gamma$, then decide $NEG(X)$;
(B) If $\beta \le P(X|[x]) \le \alpha$, then decide $BND(X)$;

where

$$\alpha = \frac{\lambda_{12} - \lambda_{32}}{(\lambda_{31} - \lambda_{32}) - (\lambda_{11} - \lambda_{12})},$$
$$\gamma = \frac{\lambda_{12} - \lambda_{22}}{(\lambda_{21} - \lambda_{22}) - (\lambda_{11} - \lambda_{12})},$$
$$\beta = \frac{\lambda_{32} - \lambda_{22}}{(\lambda_{21} - \lambda_{22}) - (\lambda_{31} - \lambda_{32})}.$$

If a loss function with $\lambda_{11} \le \lambda_{31} < \lambda_{21}$ and $\lambda_{22} \le \lambda_{32} < \lambda_{12}$ also satisfies the condition:

$$(\lambda_{12} - \lambda_{32})(\lambda_{21} - \lambda_{31}) \ge (\lambda_{31} - \lambda_{11})(\lambda_{32} - \lambda_{22}),$$

then $\alpha \ge \gamma \ge \beta$.

When $\alpha > \beta$, we have $\alpha > \gamma > \beta$. The DTRS has the following decision rules:

(P1) If $P(X|[x]) \ge \alpha$, then decide $POS(X)$;
(N1) If $P(X|[x]) \le \beta$, then decide $NEG(X)$;
(B1) If $\beta < P(X|[x]) < \alpha$, then decide $BND(X)$.

Using these three decision rules, we obtain the probabilistic approximation:

$$\underline{R}_{(GRS,\alpha)}(X) = \{x \mid P(X|[x]) \ge \alpha,\ x \in U\}, \tag{5}$$
$$\overline{R}_{(GRS,\beta)}(X) = \{x \mid P(X|[x]) > \beta,\ x \in U\}, \tag{6}$$

where $0 \le \beta < \alpha \le 1$.

In the DTRS framework, the PRS model, variable precision rough set model, Bayesian rough set model, and the 0.5-probabilistic rough set model can be pooled together and studied based on the notion of a conditional function.

2.3. LRS

In order to obtain a rough set $< lowerapproximation, upperapproximation >$ of any subset on sample set, all the information granules are first computed by comparing the difference between any two objects from a given data set. This implies that a global rough set must observe the relationships between a target concept and each of the information granules. However, this is not a good strategy for approximating a target concept $X \subseteq U$. In fact, the information granules $\{[x] : [x] \cap X = \phi,\ x \in U\}$ are not useful for computing the lower/upper approximation of X. Indeed, we only need to calculate the information granules related to the target concept X. In particular, for a very large-scale data set, it is often true that $n \gg |X|$, where n and $|X|$ are the number in the data set and the number of X, respectively. This time reduction improvement would be very useful for rough data analysis based on big data. According to this consideration, Qian et al. [34] reconstructed the rough set model as follows.

Definition 1. Let (U, R) be an approximation space and $\mathcal{D}$ be an inclusion degree defined on $\mathcal{P}(U) \times \mathcal{P}(U)$. Then, for any $X \subseteq U$, the α-lower and β-upper approximations are defined by

$$\underline{R}_{(LRS,\alpha)}(X) = \{x \mid \mathcal{D}(X/[x]_R) \geq \alpha,\ x \in X\}, \tag{7}$$

$$\overline{R}_{(LRS,\beta)}(X) = \{x \mid \mathcal{D}(X/[x]_R) > \beta,\ x \in U\} = \cup\{[x]_R \mid \mathcal{D}(X/[x]_R) > \beta,\ x \in X\}. \tag{8}$$

The pair $\langle \underline{R}_{(LRS,\alpha)}(X), \overline{R}_{(LRS,\beta)}(X)\rangle$ is called the LRS.

The boundary of X is denoted by $BN_R(X) = \overline{R}_{(LRS,\beta)}(X) - \underline{R}_{(LRS,\alpha)}(X)$, which we refer to as the local boundary region of X.

We should note that when $\alpha = 1$ and $\beta = 0$, the LRS above will degenerate into a PRS. Thus, the LRS model will have the same form and semantics as the original PRS model (it can be viewed as a type of global rough set). From this viewpoint, the reconstruction does not change the original idea of rough set theory proposed by Pawlak and it has a consistent capacity to deal with uncertainty.

Unlike a global rough set, according to the definition of the rough set given above, the computation of its lower/upper approximation is based only on the information granules determined by objects within a target concept, rather than those of all objects from a given universe, which can significantly reduce the time consumed when computing approximations. Thus, we refer to these types of rough sets as *LRSs*. For a local rough set, we refer to $\underline{R}_{(LRS,\alpha)}(X)$ as the local lower approximation of X and $\overline{R}_{(LRS,\beta)}(X)$ as the local upper approximation of X.

Proposition 1. *Let (U, R) be an approximation space and R be an equivalence relation on U, then for any $X \subseteq U$, the β-upper approximation of GRS is equal to that of the LRS model, i.e., $\overline{R}_{(LRS,\beta)}(X) = \overline{R}_{(GRS,\beta)}(X)$.*

Proof. For any $x \in \overline{R}_{(LRS,\beta)}(X)$, based on the definition of $\overline{R}_{(LRS,\beta)}(X)$, we can obtain $[x]_R \subseteq \overline{R}_{(LRS,\beta)}(X)$ and it is easy to see that $x \in \overline{R}_{(GRS,\beta)}(X)$, so $\overline{R}_{(LRS,\beta)}(X) \subseteq \overline{R}_{(GRS,\beta)}(X)$.

By contrast, for any $x \in \overline{R}_{(GRS,\beta)}(X)$, based on the definition of $\overline{R}_{(GRS,\beta)}(X)$, it is easy to find that $x \in \overline{R}_{(LRS,\beta)}(X)$, and thus $\overline{R}_{(GRS,\beta)}(X) \subseteq \overline{R}_{(LRS,\beta)}(X)$. Hence, $\overline{R}_{(LRS,\beta)}(X) = \overline{R}_{(GRS,\beta)}(X)$. $\square$

2.4. MGRS

In 2006, Qian et al. [29] proposed an extended model of the PRS called the MGRS where a target concept is approximated by multiple binary relations. Furthermore, two types of common MGRSs were defined according to optimistic and pessimistic strategies, i.e., the optimistic MGRS and pessimistic MGRS, respectively [33,31].

Definition 2. Let $S = (U, AT, f)$ be an information system, $A_1, A_2, \cdots, A_m \subseteq AT$, and $X \subseteq U$. The optimistic lower approximation and upper approximation of X with respect to $A_1, A_2, \cdots, A_m$ are denoted by $\underline{\sum_{i=1}^m A_i}^O X$ and $\overline{\sum_{i=1}^m A_i}^O X$, respectively, where

$$\underline{\sum_{i=1}^m A_i}^O (X) = \bigcup\{x \in U \mid [x]_{A_i} \subseteq X, \text{for some } i \leq m\}, \tag{9}$$

$$\overline{\sum_{i=1}^m A_i}^O (X) = \sim \underline{\sum_{i=1}^m A_i}^O (\sim X). \tag{10}$$

Then, $(\underline{\sum_{i=1}^m A_i}^O X, \overline{\sum_{i=1}^m A_i}^O X)$ is the optimistic MGRS [33]. The word "optimistic" is used to express the idea that in multiple independent granular structures, we only need at least one granular structure to satisfy the inclusion condition between an equivalence class and a target concept. The upper approximation of the optimistic MGRS is defined by the complement of the lower approximation. In addition, the *area of uncertainty* or *boundary region* in the MGRS is $Bn^O_{\sum_{i=1}^m}(X) = \overline{\sum_{i=1}^m A_i}^O X \setminus \underline{\sum_{i=1}^m A_i}^O X$.

The pessimistic MGRS is defined as follows:

$$\underline{\sum_{i=1}^{m} A_i}^{P}(X) = \{x \in U \mid [x]_{A_1} \subseteq X \wedge [x]_{A_2} \subseteq X \wedge \cdots \wedge [x]_{A_m} \subseteq X\}, \tag{11}$$

$$\overline{\sum_{i=1}^{m} A_i}^{P}(X) = \sim \underline{\sum_{i=1}^{m} A_i}^{P}(\sim X). \tag{12}$$

Then, $(\underline{\sum_{i=1}^{m} A_i}^{P} X, \overline{\sum_{i=1}^{m} A_i}^{P} X)$ is the pessimistic MGRS [31]. The word "pessimistic" is used to express the idea that in multiple independent granular structures, we need all the granular structures to satisfy the inclusion condition between an equivalence class and a target concept. The upper approximation of the optimistic MGRS is also defined by the complement of the lower approximation. In addition, the *area of uncertainty* or *boundary region* in the MGRS is

$$BN^{P}_{\sum_{i=1}^{m} A_i}(X) = \overline{\sum_{i=1}^{m} A_i}^{P}(X) \setminus \underline{\sum_{i=1}^{m} A_i}^{P}(X).$$

Selecting relevant granules is a key issue with MGRSs. Information granularity can help to choose more suitable granular structures for approximating a target concept with much higher approximation accuracy. Several definitions of information granularity have been proposed according to various perspectives and viewpoints. Liang et al. [17,18] used two forms of information granularity for measuring complete data and incomplete data. Wierman [43] gave a so-called granulation measure for evaluating the uncertainty of knowledge in a knowledge base, where its formulation is the same as that of Shannon entropy to some extent. The combination granulation proposed by Qian and Liang [16] can also be used to measure the granulation degree of knowledge in a knowledge base. Xu et al. [45] improved the concept of roughness in the rough set theory proposed by Pawlak [25], which also can be viewed as an information granularity. Qian et al. [1] first presented a new concept called the fuzzy granular structure distance for differentiating two fuzzy granular structures from the same universe.

2.5. MG-DTRS

By combining the DTRS and MGRS, Qian et al. [28] developed a new MGRS model called the MG-DTRS. In this subsection, we review two MG-DTRS models called the optimistic MG-DTRS (OMG-DTRS, also called GOMG-DTRS) and the pessimistic MG-DTRS (PMG-DTRS, also called the global pessimistic MG-DTRS or GPMG-DTRS).

Definition 3. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations and $\forall X \subseteq U$, the optimistic MG lower and upper approximations are denoted by $\underline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\alpha)}(X)$ and $\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X)$, respectively,

$$\underline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\alpha)}(X) = \{x : P(X|[x]_{R_1}) \geq \alpha \vee P(X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, x \in U\}, \tag{13}$$

$$\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X) = U - \{x : P(X|[x]_{R_1}) \leq \beta \wedge P(X|[x]_{R_2}) \leq \beta \wedge \cdots \wedge P(X|[x]_{R_m}) \leq \beta, x \in U\}, \tag{14}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, $[x]_{R_i} (1 \leq i \leq m)$ is the equivalence class of x induced by R_i, $P(X|[x]_{R_i})$ is the conditional probability of the equivalent class $[x]_{R_i}$ with respect to X, and α, β are two probability constraints.

By the lower approximation $\underline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\alpha)}(X)$ and upper approximation $\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X)$, the optimistic MG boundary region of X is

$$BN^{GOMG-DTRS}_{\sum_{i=1}^{m} R_i}(X) = \overline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\beta)}(X) - \underline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\alpha)}(X). \tag{15}$$

Proposition 2. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, then for $\forall X \subseteq U$ and $0 \leq \beta < \alpha \leq 1$, the following properties hold:*

$$\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\ \beta)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{(GRS,\beta)}(X), \tag{16}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, $[x]_{R_i} (1 \leq i \leq m)$ is the equivalence class of x induced by R_i, $P(X|[x]_{R_i})$ is the conditional probability of the equivalent class $[x]_{R_i}$ with respect to X, $\overline{R_i}_{(GRS,\beta)}(X) = \{x : P(X|[x]_{R_i}) > \beta, x \in U\}$, and α and β are two probability constraints.

Proof. For any $x \in \overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X)$, $X \subseteq U$, we have

$$\begin{aligned}
\forall x \in \overline{\sum\nolimits_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X) &\Leftrightarrow x \in U - \{x : P(X|[x]_{R_1}) \le \beta \wedge (X|[x]_{R_2}) \le \beta \wedge \cdots \wedge P(X|[x]_{R_m}) \le \beta, x \in U\}, \\
&\Leftrightarrow x \in \{x : P(X|[x]_{R_1}) > \beta \vee (X|[x]_{R_2}) > \beta \vee \cdots \vee P(X|[x]_{R_m}) > \beta, x \in U\}, \\
&\Leftrightarrow x \in \{x : P(X|[x]_{R_1}) > \beta, x \in U\} \text{ or } x \in \{x : P(X|[x]_{R_2}) > \beta, x \in U\} \text{ or } \cdots \text{ or} \\
&\quad x \in \{x : P(X|[x]_{R_m}) > \beta, x \in U\}, \\
&\Leftrightarrow x \in \overline{R_1}_{(GRS,\beta)}(X) \text{ or } x \in \overline{R_2}_{(GRS,\beta)}(X) \text{ or } \cdots \text{ or } x \in \overline{R_m}_{(GRS,\beta)}(X), \\
&\Leftrightarrow x \in \bigcup_{i=1}^{m} \overline{R_i}_{(GRS,\beta)}(X).
\end{aligned}$$

Thus, we can obtain $\overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\ \beta)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{(GRS,\beta)}(X)$.

By Proposition 2, we see that the optimistic MG lower approximation can be obtained by the union of the corresponding multiple single granulation lower approximations. □

Definition 4. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations and $\forall X \subseteq U$, the optimistic MG lower and upper approximations are denoted by $\underline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\alpha)}(X)$ and $\overline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\beta)}(X)$, respectively,

$$\underline{\sum\nolimits_{i=1}^{m} R_i}^{(GPMG-DTRS,\alpha)}(X) = \{x : P(X|[x]_{R_1}) \ge \alpha \wedge P(X|[x]_{R_2}) \ge \alpha \wedge \cdots \wedge P(X|[x]_{R_m}) \ge \alpha, x \in U\}, \tag{17}$$

$$\overline{\sum\nolimits_{i=1}^{m} R_i}^{(GPMG-DTRS,\beta)}(X) = U - \{x : P(X|[x]_{R_1}) \le \beta \vee P(X|[x]_{R_2}) \le \beta \vee \cdots \vee P(X|[x]_{R_m}) \le \beta, x \in U\}, \tag{18}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations.

By the lower approximation $\underline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\alpha)}(X)$ and upper approximation $\overline{\sum_{i=1}^{m} R_i}^{(GPMG-DTRS,\beta)}(X)$, the optimistic MG boundary region of X is

$$BN_{\sum_{i=1}^{m} R_i}^{GPMG-DTRS}(X) = \overline{\sum\nolimits_{i=1}^{m} R_i}^{(GPMG-DTRS,\beta)}(X) - \underline{\sum\nolimits_{i=1}^{m} R_i}^{(GPMG-DTRS,\alpha)}(X). \tag{19}$$

3. LMG-DTRS

In this section, we extend the classical/global MG-DTRS to a LMG-DTRS. In fact, different fusion strategies yield various versions of LMG-DTRSs, including local optimistic (LO), local pessimistic, and local mean MG-DTRSs. For convenience, we mainly consider the LO version in this study. In the following, we provide a framework for the LOMG-DTRS.

3.1. LOMG-DTRS

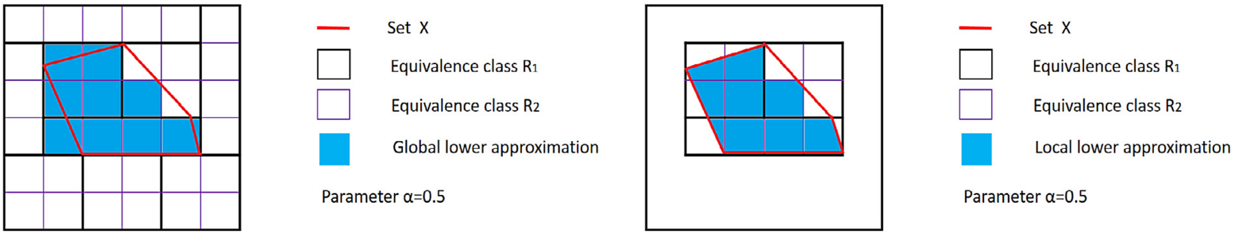
The classical MG-DTRS can approximate a concept by using multiple equivalence relations, but these objects need to be labeled. Thus, this can be regarded as a supervised learning method in some sense. However, the existing MG version cannot deal with data sets that contain a limited number of labels. Therefore, it is necessary to develop a new rough set based on multiple relations to handle big data with limited labels. In this subsection, we provide a definition of the LOMG-DTRS and investigate some of its properties.

Definition 5. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, and $\forall X \subseteq U$, the lower and upper approximations of X in a LOMG-DTRS are denoted by $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X)$ and $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X)$, respectively.

$$\underline{\sum\nolimits_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \ge \alpha \vee P(X|[x]_{R_2}) \ge \alpha \vee \cdots \vee P(X|[x]_{R_m}) \ge \alpha, x \in X\}, \tag{20}$$

$$\overline{\sum\nolimits_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{(LRS,\beta)}(X), \tag{21}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, $[x]_{R_i} (1 \le i \le m)$ is the equivalence class of x induced by R_i, $P(X|[x]_{R_i})$ is the conditional probability of the equivalent class $[x]_{R_i}$ with respect to X, $\overline{R_i}_{(LRS,\beta)}(X) = \cup\{[x]_{R_i} : P(X|[x]_{R_i}) > \beta, x \in X\}$, and α and β are two probability constraints.



(a) GOMG-DTRS (b) LOMG-DTRS

Fig. 1. Difference between the GOMG-DTRS and LOMG-DTRS.

The area of the boundary or region of uncertainty is defined as:

$$BN^{LOMG-DTRS}_{\sum_{i=1}^m R_i}(X) = \overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X) - \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X). \tag{22}$$

From the definition we obtain the following interpretations:

(1) The lower approximation of a set X with respect to $\sum_{i=1}^m R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^m R_i$ with the probability greater than or equal to α and be from X.
(2) The upper approximation of a set X with respect to $\sum_{i=1}^m R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^m R_i$ with the probability greater than β.
(3) The boundary of a set X with respect to $\sum_{i=1}^m R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^m R_i$ with the probability greater than β and less than α.

The difference between GOMG-DTRS and LOMG-DTRS is shown in Fig. 1. According to Definition 1 and Fig. 1, the lower/upper approximation is only related to the equivalence classes of some objects from a target concept. However, in global rough sets, the lower/upper approximation always depends on the equivalence classes of all the objects. When calculating approximations, the most time-consuming process is obtaining the equivalence classes. Considering the amount of equivalence classes, the LOMG-DTRS has a great advantage compared with the GOMG-DTRS, particularly when the scale of the observed concept set is relatively small compared with that of the entire data set. Fig. 1(a) shows that the equivalence classes of all the objects need to be computed. Moreover, we note that its 0.5-lower approximation set overflows the target concept X. Thus, in the GOMG-DTRS, there is a complex relationship between the lower approximation set and a target set (including $<, =, >$). By contrast, for the LOMG-DTRS, their relation only includes $\geq$ because the lower approximation set is included in the target concept X. This is a useful property because it makes the LOMG-DTRS more robust when handling noisy data.

The properties of the GOMG-DTRS model were described in a previous study [28]. In the following, we provide a theoretical comparison between the LOMG-DTRS and GOMG-DTRS.

Proposition 3. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, then for $\forall X \subseteq U$ and $0 \leq \beta < \alpha \leq 1$, the following properties hold:*

(1) $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) \subseteq \underline{\sum_{i=1}^m R_i}^{(GOMG-DTRS,\ \alpha)}(X)$,
(2) $\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X) = \overline{\sum_{i=1}^m R_i}^{(GOMG-DTRS,\ \beta)}(X)$.

Proof. (1) $\forall x \in \overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X), \forall X \subseteq U$, we have

$$\begin{aligned}
\forall x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X), X \subseteq U \Leftrightarrow\ & x \in \{x : P(X|[x]_{R_1}) \geq \alpha \vee (X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, \\
& x \in X\}, x \in X \subseteq U, \\
\Rightarrow\ & x \in \{x : P(X|[x]_{R_1}) \geq \alpha \vee (X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, x \in U\}, \\
\Rightarrow\ & x \in \underline{\sum_{i=1}^m R_i}^{(GOMG-DTRS,\alpha)}(X),
\end{aligned}$$

and thus we can obtain $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X) \subseteq \underline{\sum_{i=1}^m R_i}^{(GOMG-DTRS,\alpha)}(X)$.

(2) According to Proposition 1, we can obtain $\overline{R}_{(LRS,\beta)}(X) = \overline{R}_{(GRS,\beta)}(X)$. Furthermore, based on Definition 5 and Proposition 2, we have

Table 3
A toy information system comprising data with limited labels.

Project	Locus	Investment	Population density	Decision
x_1	Good	Very High	Big	Yes
x_2	Good	Very High	Big	Yes
x_3	Good	High	Big	$*$
x_4	Common	Medium	Medium	No
x_5	Common	Medium	Medium	$*$
x_6	Common	Medium	Medium	No
x_7	Bad	Low	Small	No
x_8	Bad	Very low	Small	$*$

$$\begin{aligned}\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(X) &= \overline{R_1}_{(LRS,\beta)}(X) \cup \overline{R_2}_{(LRS,\beta)}(X) \cup \cdots \cup \overline{R_m}_{(LRS,\beta)}(X),\\ &= \overline{R_1}_{(GRS,\beta)}(X) \cup \overline{R_2}_{(GRS,\beta)}(X) \cup \cdots \cup \overline{R_m}_{(GRS,\beta)}(X),\\ &= \bigcup_{i=1}^{m} \overline{R_i}_{(GRS,\beta)}(X),\\ &= \overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X),\end{aligned}$$

and thus we can find that $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(X) = \overline{\sum_{i=1}^{m} R_i}^{(GOMG-DTRS,\beta)}(X)$. □

Proposition 3 shows the relationships between the lower and upper approximations for the LOMG-DTRS and GOMG-DTRS, which indicates that the LO lower approximation is not more than the GO lower approximation, and the LO upper approximation is equal to the GO upper approximation.

Next, we provide an example to illustrate the process used for computing the upper and low approximations of a target X for both the LOMG-DTRS and GOMG-DTRS, as well as their difference.

Example 1. A toy information system is shown in Table 3. Suppose that $\alpha = 0.6, \beta = 0.4, X = \{x_1, x_2, x_4, x_7\}$, and $R_1 = \{Locus\}, R_2 = \{Investment\}$. For the attribute Decision, its attribute domain is {Yes, No}. The objects with decision values of $*$ are not labeled.

For the LOMG-DTRS, we only need to obtain equivalence classes for these objects from X. After a computation, we have

$$\begin{aligned}&[x_1]_{R_1} = \{x_1, x_2, x_3\}, [x_2]_{R_1} = \{x_1, x_2, x_3\}, [x_4]_{R_1} = \{x_4, x_5, x_6\}, [x_7]_{R_1} = \{x_7, x_8\},\\ &[x_1]_{R_2} = \{x_1, x_2\}, [x_2]_{R_2} = \{x_1, x_2\}, [x_4]_{R_2} = \{x_4, x_5, x_6\}, [x_7]_{R_2} = \{x_7\},\\ &P(X|[x_1]_{R_1}) = \frac{2}{3}, P(X|[x_2]_{R_1}) = \frac{2}{3}, P(X|[x_4]_{R_1}) = \frac{1}{3}, P(X|[x_7]_{R_1}) = \frac{1}{2},\\ &P(X|[x_1]_{R_2}) = 1, P(X|[x_2]_{R_2}) = 1, P(X|[x_4]_{R_2}) = \frac{1}{3}, P(X|[x_7]_{R_2}) = 1.\end{aligned}$$

Thus,

$$\begin{aligned}&\underline{\sum_{i=1}^{2} R_i}^{(LOMG-DTRS,\ 0.6)}(X) = \{x_1, x_2, x_7\},\\ &\overline{\sum_{i=1}^{2} R_i}^{(LOMG-DTRS,\ 0.4)}(X) = [x_1]_{R_1} \cup [x_2]_{R_1} \cup [x_7]_{R_1} \cup [x_1]_{R_2} \cup [x_2]_{R_2} \cup [x_7]_{R_2} = \{x_1, x_2, x_3, x_7, x_8\}.\end{aligned}$$

In the GOMG-DTRS, we need to calculate the equivalence classes of all the objects from U. By a computation, we have

$$\begin{aligned}&[x_1]_{R_1} = \{x_1, x_2, x_3\}, [x_2]_{R_1} = \{x_1, x_2, x_3\}, [x_3]_{R_1} = \{x_1, x_2, x_3\}, [x_4]_{R_1} = \{x_4, x_5, x_6\},\\ &[x_5]_{R_1} = \{x_4, x_5, x_6\}, [x_6]_{R_1} = \{x_4, x_5, x_6\}, [x_7]_{R_1} = \{x_7, x_8\}, [x_8]_{R_1} = \{x_7, x_8\},\\ &[x_1]_{R_2} = \{x_1, x_2\}, [x_2]_{R_2} = \{x_1, x_2\}, [x_3]_{R_2} = \{x_3\}, [x_4]_{R_2} = \{x_4, x_5, x_6\},\\ &[x_5]_{R_2} = \{x_4, x_5, x_6\}, [x_6]_{R_2} = \{x_4, x_5, x_6\}, [x_7]_{R_2} = \{x_7\}, [x_8]_{R_2} = \{x_8\}.\end{aligned}$$

Then,

$$P(X|[x_1]_{R_1}) = \frac{2}{3},\ P(X|[x_2]_{R_1}) = \frac{2}{3},\ P(X|[x_3]_{R_1}) = \frac{2}{3},\ P(X|[x_4]_{R_1}) = \frac{1}{3},$$
$$P(X|[x_5]_{R_1}) = \frac{1}{3},\ P(X|[x_6]_{R_1}) = \frac{1}{3},\ P(X|[x_7]_{R_1}) = \frac{1}{2},\ P(X|[x_8]_{R_1}) = \frac{1}{2},$$
$$P(X|[x_1]_{R_2}) = 1,\ P(X|[x_2]_{R_2}) = 1,\ P(X|[x_3]_{R_2}) = 0,\ P(X|[x_4]_{R_2}) = \frac{1}{3},$$
$$P(X|[x_5]_{R_2}) = \frac{1}{3},\ P(X|[x_6]_{R_2}) = \frac{1}{3},\ P(X|[x_7]_{R_2}) = 1,\ P(X|[x_7]_{R_2}) = 0.$$

Thus, we can obtain

$$\underline{\sum\nolimits_{i=1}^{2} R_i}^{(GOMG-DTRS,\ 0.6)}(X) = \{x_1, x_2, x_3, x_7\},$$
$$\overline{\sum\nolimits_{i=1}^{2} R_i}^{(GOMG-DTRS,\ 0.4)}(X) = U - \{x_4, x_5, x_6\} = \{x_1, x_2, x_3, x_7, x_8\}.$$

In Example 1, in order to obtain the lower and upper approximations of X for the LOMG decision sets, eight equivalence classes need to be computed. However, we need to calculate 16 equivalence classes for the GOMG decision sets. Obviously, the computational time is reduced by half with respect to the number of equivalence classes. Thus, the LOMG-DTRS model can reduce the computational time, which is the first motivation of this study. Furthermore, based on the results above, we have $\underline{\sum_{i=1}^{2} R_i}^{(LOMG-DTRS,\ 0.6)}(X) \subseteq \underline{\sum_{i=1}^{2} R_i}^{(GOMG-DTRS,\ 0.6)}(X)$ and $\overline{\sum_{i=1}^{2} R_i}^{(LOMG-DTRS,\ 0.4)}(X) = \overline{\sum_{i=1}^{2} R_i}^{(GOMG-DTRS,\ 0.4)}(X)$. This observation is consistent with the finding of Proposition 3.

We illustrate the second motivation based on Example 1. According to Table 3, we know that the objects x_3, x_5, x_8 are not labeled. The equivalence classes need to be computed in advance when we calculate the lower and upper approximations of a target concept X. To calculate the equivalence class for each element of X, we use the labeled data x_1, x_2, x_4, x_6, x_7, as well as the unlabeled data x_3, x_5, x_8. Therefore, the reconstructed MG-DTRS model, i.e., LOMG-DTRS, can make full use of the information hidden in unlabeled data.

In the following, we consider some interesting properties of the LOMG-DTRS.

Proposition 4. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m granular structures on U, where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m granular structures, then for $\forall X \subseteq U$ and $0 \le \beta < \alpha \le 1$, the following properties hold:*

(1) $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) \subseteq X$,

(2) $\beta \in [0, min\{P(X|[x]_{R_i} : x \in X, i = 1, 2, \cdots, m\}) \Rightarrow X \subseteq \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X)$,

(3) $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(\phi) = \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(\phi) = \phi$,

(4) $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(U) = \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(U) = U$.

Proof. (1) By the LOMG lower approximation in Definition 5, we can easily obtain $\forall X \subseteq U$, $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) \subseteq X$.

(2) For $\forall x \in U$, we can obtain $x \in [x]_{R_i}$ and $[x]_{R_i} \cap X \neq \phi$. Hence, we have $P(X|[x]_{R_i}) > 0$. Thus, when $\beta \in [0, min\{P(X|[x]_{R_i} : x \in X, i = 1, 2, \cdots, m\})$, we find that $\forall x \in X, P(X|[x]_{R_i}) > \beta$ holds. This also implies that $x \in \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\}$. Furthermore, $x \in \bigcup_{i=1}^{m} \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\}$. Therefore, $X \subseteq \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(X)$ is true when $\beta \in [0, min\{P(X|[x]_{R_i} : x \in X, i = 1, 2, \cdots, m\})$.

(3) For $\forall x \in U, x \notin \phi$ and thus $P(\phi|[x]_{R_i}) = 0 < \alpha$ and $P(\phi|[x]_{R_i}) = 0 \le \beta$. Therefore, for any $0 \le \beta < \alpha \le 1$, we have $x \notin \{x : \vee P(\phi|[x]_{R_i}) \ge \alpha, x \in X, i = 1, 2, \cdots, m\}$ and $x \notin \bigcup_{i=1}^{m} \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\}$, i.e., $\{x : \vee P(\phi|[x]_{R_i}) \ge \alpha, x \in U, i = 1, 2, \cdots, m\} = \phi$ and $\bigcup_{i=1}^{m} \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\} = \phi$. According to Definition 5, we have $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(\phi) = \phi$ and $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(\phi) = \phi$. Thus, $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(\phi) = \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(\phi) = \phi$.

(4) For $\forall x \in U, [x]_{R_i} \subseteq U$ and thus $P(U|[x]_{R_i}) = 1 \ge \alpha$ and $P(U|[x]_{R_i}) = 1 > \beta$. Therefore, for any $0 \le \beta < \alpha \le 1$, we have $x \in \{x : \vee P(U|[x]_{R_i}) \ge \alpha, x \in X, i = 1, 2, \cdots, m\}$ and $x \in \bigcup_{i=1}^{m} \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\}$, i.e., $\{x : \vee P(U|[x]_{R_i}) \ge \alpha, x \in U, i = 1, 2, \cdots, m\} = U$ and $\bigcup_{i=1}^{m} \cup\{[x]_i : P(X|[x]_{R_i}) > \beta, x \in X\} = U$. According to Definition 5, we have $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(U) = U$ and $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(U) = U$. Thus, $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(U) = \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(U) = U$.

By term (1) in Proposition 4, we can find that the lower approximation of the LOMG-DTRS must be included in a target concept when $0 < \alpha \le 1$, whereas it may not for the GOMG-DTRS. Similar to the GOMG-DTRS, the upper approximation of the LOMG-DTRS may overflow the range of the target concept. Furthermore, the lower and upper approximations of two special sets (empty set and universal set) are the lower and upper approximations of themselves. □

Proposition 5. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, and $\forall X \subseteq U$, then the following properties hold:*

(1) $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) \supseteq \underline{R_i}_{(LRS,\ \alpha)}(X), i \in \{1, 2, \cdots, m\}$,
(2) $\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X) \supseteq \overline{R_i}_{(LRS,\ \beta)}(X), i \in \{1, 2, \cdots, m\}$,
(3) $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \cup_{i=1}^m \underline{R_i}_{(LRS,\ \alpha)}(X)$.

Proof. (1) $\forall x \in \underline{R_i}_{(LRS,\ \alpha)}(X)$, we have

$$\begin{aligned} x \in \underline{R_i}_{(LRS,\ \alpha)}(X), i \in \{1, 2, \cdots, m\} &\Leftrightarrow x \in X, P(X|[x]_{R_i}) \geq \alpha, i \in \{1, 2, \cdots, m\} \\ &\Rightarrow x \in X, P(X|[x]_{R_1}) \geq \alpha \vee P(X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_i}) \geq \alpha \vee \cdots \\ &\quad \vee P(X|[x]_{R_m}) \geq \alpha, \\ &\Rightarrow x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X), \end{aligned}$$

and thus we can obtain $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) \supseteq \underline{R_i}_{(LRS,\ \alpha)}(X), i \in \{1, 2, \cdots, m\}$.

(2) $\forall x \in \overline{R_i}_{(LRS,\ \beta)}(X)$, we have

$$\begin{aligned} x \in \overline{R_i}_{(LRS,\ \beta)}(X), i \in \{1, 2, \cdots, m\} &\Rightarrow x \in \overline{R_1}_{(LRS,\ \beta)}(X) \text{ or } x \in \overline{R_2}_{(LRS,\ \beta)}(X) \text{ or } \cdots \text{ or } x \in \overline{R_m}_{(LRS,\ \beta)}(X), \\ &\Rightarrow x \in \{\overline{R_1}_{(LRS,\ \beta)}(X) \cup \overline{R_2}_{(LRS,\ \beta)}(X) \cup \cdots \cup x \in \overline{R_m}_{(LRS,\ \beta)}(X)\}, \\ &\Rightarrow x \in \overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X), \end{aligned}$$

and thus we can obtain $\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) \supseteq \overline{R_i}_{(LRS,\ \beta)}(X), i \in \{1, 2, \cdots, m\}$.

(3) $\forall x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X)$, we have

$$\begin{aligned} x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X) &\Leftrightarrow x \in X, P(X|[x]_{R_1}) \geq \alpha \vee P(X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, \\ &\Leftrightarrow x \in X, P(X|[x]_{R_1}) \geq \alpha \text{ or } x \in X, P(X|[x]_{R_2}) \geq \alpha \text{ or } \cdots \text{ or } x \in X, P(X|[x]_{R_m}) \geq \alpha, \\ &\Leftrightarrow x \in \underline{R_1}_{(LRS,\ \alpha)}(X) \text{ or } x \in \underline{R_2}_{(LRS,\ \alpha)}(X) \text{ or } \cdots \text{ or } x \in \underline{R_m}_{(LRS,\ \alpha)}(X), \\ &\Leftrightarrow x \in \bigcup_{i=1}^m \underline{R_i}_{(LRS,\ \alpha)}(X), \end{aligned}$$

from which we can obtain $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \bigcup_{i=1}^m \underline{R_i}_{\alpha}(X)$. □

Proposition 5 shows the relationships between the approximations of the local DTRSs and the approximations of the LOMG-DTRSs. By term (3) in Proposition 5, we can see the LOMG decision-theoretic lower approximation can also be obtained by the union of the local decision-theoretic lower approximation in terms of each granular structure.

Proposition 6. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations. Suppose that $\forall X_1, X_2 \subseteq U$ and $X_1 \subseteq X_2$. Then,*

(1) $\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X_1) \subseteq \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X_2)$,
(2) $\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\beta)}(X_1) \subseteq \overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\beta)}(X_2)$.

Proof. (1) Suppose that $X_1 \subseteq X_2$, then for any $x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X_1)$, we have

$$\begin{aligned} \forall x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X_1), X_1 \subseteq X_2 &\Leftrightarrow x \in \{x : P(X_1|[x]_{R_1}) \geq \alpha \vee (X_1|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X_1|[x]_{R_m}) \geq \alpha, \\ &\quad x \in X_1\}, P(X_2|[x]_{R_i}) \geq P(X_1|[x]_{R_i}), \\ &\Rightarrow x \in \{x : P(X_2|[x]_{R_1}) \geq \alpha \vee (X_2|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X_2|[x]_{R_m}) \geq \alpha, \\ &\quad x \in X_2\}, \\ &\Rightarrow x \in \underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\alpha)}(X_2). \end{aligned}$$

Hence, $\underline{\sum_{i=1}^{m} R_i}^{LOMG-DTRS,\alpha}(X_1) \subseteq \underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\alpha)}(X_2)$.

(2) Suppose that $X_1 \subseteq X_2$, for any $x \in \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X_1)$, then according to Definition 5, we have

$$\begin{aligned}
\forall x \in \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(X_1), X_1 \subseteq X_2 &\Rightarrow x \in \cup\{[x]_{R_1} : P(X_1|[x]_{R_1}) > \beta\} \text{ or } x \in \cup\{[x]_{R_2} : P(X_1|[x]_{R_2}) > \beta\},\\
&\quad \text{or } \cdots \text{ or } x \in \cup\{[x]_{R_m} : P(X_1|[x]_{R_m}) > \beta\}, P(X_2|[x]_{R_i}) \geq P(X_1|[x]_{R_i}),\\
&\Rightarrow x \in \cup\{[x]_{R_1} : P(X_2|[x]_{R_1}) > \beta\} \text{ or } x \in \cup\{[x]_{R_2} : P(X_2|[x]_{R_2}) > \beta\},\\
&\quad \text{or } \cdots \text{ or } x \in \cup\{[x]_{R_m} : P(X_2|[x]_{R_m}) > \beta\},\\
&\Rightarrow x \in \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\beta)}(X_2),
\end{aligned}$$

so we can obtain $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X_1) \subseteq \overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X_2)$. □

By Proposition 6, we can see that the lower and upper approximations of a target concept become much larger as the size of the target concept increases. Thus, if an inclusion relationship exists between the two sets, their lower and upper approximations also satisfy this relationship.

3.2. Computing the LOMG decision-theoretic lower approximation of a target concept

Next, we provide an algorithm for computing the LOMG lower approximation of a target concept.

Algorithm 1. Computing the LOMG decision-theoretic lower approximation of a target concept (LOMGDTLAC).

Input: Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and a target concept set $X \subseteq U$
Output: The local lower approximation LLA of a target concept X

```
Step 1: For i = 1 to m
            For j = 1 to |X|
              compute [x_j]_{R_i};
            end
        end
Step 2: LLA = φ, j = 1;
Step 3: For j = 1 to |X|
            If ∨P(X|[x_j]_{R_i}) ≥ α, where i = 1, 2, ···, m
              LLA ← LLA ∪ x_j;
            end
        end
Step 4: Return LLA.
```

In addition, we provide the algorithm for computing the GOMG lower approximation of a target concept in Algorithm 2.

Algorithm 2. Computing the GOMG decision-theoretic lower approximation of a target concept (GOMGDTLAC).

Input: Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and a target concept set $X \subseteq U$
Output: The global lower approximation GLA of a target concept X

```
Step 1: For i = 1 to m
            For j = 1 to |U|
              compute [x_j]_{R_i};
            end
        end
Step 2: GLA = φ, j = 1;
Step 3: For j = 1 to |U|
            If ∨P(X|[x_j]_{R_i}) ≥ α, where i = 1, 2, ···, m
              GLA ← GLA ∪ x_j;
            end
        end
Step 4: Return GLA.
```

Table 4
Time complexity of computing the LOMG decision-theoretic lower approximation of a target concept.

Algorithms	Step 1	Step 3	Other steps	Total time consumption
GOMGDTLAC	$\mathcal{O}(\|m\|\|U\|^2)$	$\mathcal{O}(\|m\|\|X\|\|U\|)$	Constant	$\mathcal{O}(\|m\|\|U\|^2 + \|m\|\|X\|\|U\|)$
LOMGDTLAC	$\mathcal{O}(\|m\|\|X\|\|U\|)$	$\mathcal{O}(\|m\|\|X\|^2)$	Constant	$\mathcal{O}(\|m\|\|X\|\|U\| + \|m\|\|X\|^2)$

In the LOMGDTLAC algorithm, Step 1 needs to compute $|m||X|$ equivalence classes by scanning the entire universe U, so the time complexity for computing $|X|$ equivalence classes on $|U|$ based on $|m|$ binary relations is $\mathcal{O}(|m||X||U|)$. However, for a GOMG decision-theoretic lower approximation, the time complexity for computing all the equivalence classes on $|U|$ based on $|m|$ binary relations is $\mathcal{O}(|m||U|^2)$. In Step 3, computing the lower approximation requires comparisons of $|m||X|^2$ equivalence classes with the target concept X. Thus, its time complexity is $\mathcal{O}(|m||X|^2)$. However, the time complexity of the global approximation process is $\mathcal{O}(|m||X||U|)$. The time complexity of the other steps can be treated as constant. To illustrate the high efficiency of the LOMGDTLAC algorithm, we denote the algorithm for computing the GOMG lower approximation by GOMGDTLAC, and their time complexities are shown in Table 4. Next, we quantitatively analyze the relationship between the time complexity of the LOMGDTLAC and GOMGDTLAC algorithms.

Theorem 1. *Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U, where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations, and a target concept $X \subseteq U$, the time reduction ratio with the LOMGDTLAC algorithm is $p = 1 - \frac{|X|}{|U|}$ relative to the GOMGDTLAC algorithm. The time efficiency improvement with the LOMGDTLAC algorithm is $t = \frac{|U|}{|X|}$ relative to the GOMGDTLAC algorithm.*

Proof. From Table 4, we can obtain

$$
\begin{aligned}
p &= \mathcal{O}(|m||U|^2 + |m||X||U|) - \mathcal{O}(|m||X||U| + |m||X|^2)/\mathcal{O}(|m||U|^2 + |m||X||U|) \\
&= \mathcal{O}(|m||U|^2 + |m||X||U| - |m||X||U| - |m||X|^2))/\mathcal{O}(|m||U|^2 + |m||X||U|) \\
&= \mathcal{O}(|m||U|^2 - |m||X|^2))/\mathcal{O}(|m||U|^2 + |m||X||U|) \\
&= \mathcal{O}(|m|(|U| + |X|)(|U| - |X|))/\mathcal{O}(|m|(|U| + |X|)|U|) \\
&= \mathcal{O}(|U| - |X|))/\mathcal{O}(|U|) \\
&= 1 - \frac{|X|}{|U|}, \\
t &= \mathcal{O}(|m||U|^2 + |m||X||U|)/\mathcal{O}(|m||X||U| + |m||X|^2) \\
&= \mathcal{O}(|m|(|U| + |X|)|U|)/\mathcal{O}(|m|(|U| + |X|)|X|) \\
&= \frac{|U|}{|X|}.
\end{aligned}
$$

Thus, the time reduction ratio with the LOMGDTLAC algorithm is $p = 1 - \frac{|X|}{|U|}$ relative to the GOMGDTLAC algorithm, and the time efficiency improvement with the LOMGDTLAC algorithm is $t = \frac{|U|}{|X|}$ relative to the GOMGDTLAC algorithm. This completes the proof. □

Based on the analysis using the theorem above, we can see that the LOMGDTLAC algorithm is much more efficient at obtaining a local lower approximation. In the big data era, the scale required to deal with the target concept is often smaller compared with the scale of a given data set. In particular, when $|X| \ll |U|$, the LOMGDTLAC algorithm is still efficient, whereas the GOMGDTLAC algorithm may have difficulty to calculate the lower approximation. From Theorem 1, we can see that the LOMGDTLAC algorithm is more efficient when the scale of a data set is bigger.

3.3. Experimental analysis

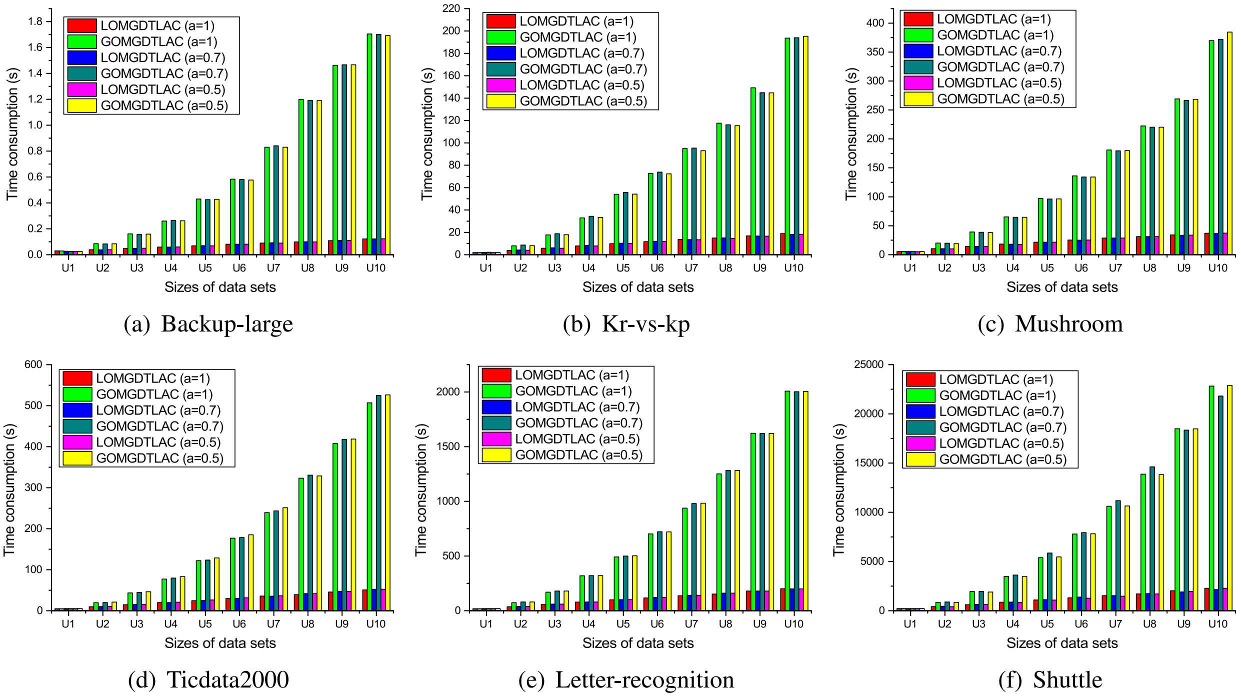
In this subsection, we illustrate the efficiency of the LOMGDTLAC algorithm by using the six UCI data sets shown in Table 5 to compare the computational performance of the LOMGDTLAC and GOMGDTLAC algorithms.

The experimental setting comprised a personal computer operating Windows 10 with an Intel(R)Core i7-4790 CPU 3.6 GHz processor and 16 GB DDR3 memory. The software used was Matlab (2012a). Without loss of generality, we used the parameters of $\alpha = 0.5$, $\alpha = 0.7$, and $\alpha = 1$.

The data sets all comprised categorical data. Among the six data sets, Mushroom was a data set with missing values. To ensure the uniform treatment of all the data sets, we removed the objects with missing values. In the experimental design, we fixed the size of a target concept, which comprised the first 10% of the objects in each of the six data sets. To distinguish the computational times, we divided each of the six data sets into 10 parts of equal size. The first part was

Table 5
Data set descriptions.

Id	Data sets	Cases	Features	Classes
1	Backup-large	376	35	19
2	Kr-vs-kp	3196	36	2
3	Mushroom	5644	22	2
4	Ticdata2000	5822	85	2
5	Letter-recognition	20000	16	26
6	Shuttle	58000	9	7



(a) Backup-large (b) Kr-vs-kp (c) Mushroom
(d) Ticdata2000 (e) Letter-recognition (f) Shuttle

Fig. 2. Times required by LOMGDTLAC and GOMGDTLAC versus the size of the universe.

Table 6
Computational times required for concept approximation using LOMGDTLAC and GOMGDTLAC.

Data sets	$\alpha = 0.5$		$\alpha = 0.7$		$\alpha = 1$	
	LOMGDTLAC	GOMGDTLAC	LOMGDTLAC	GOMGDTLAC	LOMGDTLAC	GOMGDTLAC
1	0.123179	1.692173	0.122745	1.700576	0.122273	1.704785
2	18.2299	195.3078	18.1549	194.0444	18.9128	193.6883
3	37.2664	384.5981	36.612398	371.9906	36.9675	369.7628
4	52.3869	526.3573	52.3747	524.7541	50.7401	507.0200
5	199.7684	2005.2692	199.9358	2003.1935	200.6743	2007.9204
6	2273.5318	22895.8649	2146.9253	21811.1123	2273.5317	22895.8649

regarded as the first data set, the combination of the first part and the second part was treated as the second data set, the combination of the second data set and the third part was regarded as the third data set, . . . , and the combination of all 10 parts was treated as the tenth data set. Furthermore, the attribution sets were split into two parts, which were regarded as two binary relations. These data sets were used to calculate time required by the LOMGDTLAC and GOMGDTLAC algorithms, and to represent this time in terms of the size of the universe based on two binary relations. To ensure fairness, the two algorithms were not optimized.

The experimental results obtained using the six data sets are shown in Fig. 2 (the x-coordinate denotes the size of the data set starting from the smallest and the y-coordinate represents the computational time) and Table 6. Fig. 2 shows that the computational time increased with the size of the data set using each of the two algorithms. In addition, for each of values of the parameter α, the LOMGDTLAC algorithm was consistently faster than the GOMGDTLAC algorithm when using the same universe and attribute set. Moreover, as the scale of the data set increased, the time consumption required by the LOMGDTLAC algorithm increased in a linear manner, whereas the time consumption by the GOMGDTLAC algorithm increased according to the square of the number of objects. Table 6 shows the computational times required by the LOMGDTLAC and GOMGDTLAC algorithms using the tenth data set based on the six original data sets, which indicates that LOMGDTLAC only required one-tenth of the execution time consumed by GOMGDTLAC. Thus, in terms of computing the local lower approximation of a target concept, the LOMGDTLAC algorithm based on LRSs is an efficient tool for handling big data.

4. Two types of local MGRS model as well as the relationships between them

In this section, we consider two more types based on the local MGRS framework using PRSs and variable precision rough sets, which we denote as local MG-PRS model and local MG variable precision rough set model, respectively. For convenience, we mainly describe the optimistic versions. The pessimistic versions can be obtained in a similar manner so we omit them.

4.1. LOMG-PRS model

First, we give a definition of the optimistic MG PRS model.

Definition 6. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and $\forall X \subseteq U$, the LOMG Pawlak's lower and upper approximations of X are denoted by $\underline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X)$ and $\overline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X)$.

$$\underline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X) = \{x : [x]_{R_1} \subseteq X \vee [x]_{R_2} \subseteq X \vee \cdots \vee [x]_{R_m} \subseteq X, x \in X\}, \tag{23}$$

$$\overline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X) = \bigcup_{i=1}^m \overline{R_i}_\beta(X), \tag{24}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations and $\overline{R_i}_\beta(X) = \cup\{[x]_{R_i} : [x]_{R_i} \cap X \neq \phi, x \in X\}$.

From the definition we obtain the following interpretations:

(1) The LOMG-PRS lower approximation of a set X with respect to $\sum_{i=1}^m R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^m R_i$ with the probability equal to 1 and being from X.
(2) The LOMG-PRS upper approximation of a set X with respect to $\sum_{i=1}^m R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^m R_i$ with the probability greater than 0.

By Definitions 5 and 6, we can easily obtain the relationship between the LOMG-DTRS and LOMG-PRS.

$$\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha)}(X) \supseteq \underline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X),$$
$$\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta)}(X) \subseteq \overline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X).$$

Furthermore, when $\alpha = 1, \beta = 0$, the LOMG-DTRS becomes the LOMG-PRS model.

$$\underline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \alpha=1)}(X) = \{x : P(X|[x]_{R_1}) \geq 1 \vee P(X|[x]_{R_2}) \geq 1 \vee \cdots \vee P(X|[x]_{R_m}) \geq 1, x \in X\},$$
$$\overline{\sum_{i=1}^m R_i}^{(LOMG-DTRS,\ \beta=0)}(X) = \bigcup_{i=1}^m (\cup\{[x]_{R_i} : P(X|[x]_{R_i}) > \beta, x \in X\}),$$
$$\Rightarrow \underline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X) = \{x : [x]_{R_1} \subseteq X \vee [x]_{R_2} \subseteq X \vee \cdots \vee [x]_{R_m} \subseteq X, x \in X\},$$
$$\overline{\sum_{i=1}^m R_i}^{LOMG-PRS}(X) = \bigcup_{i=1}^m (\cup\{[x]_{R_i} : [x]_{R_i} \cap X \neq \phi, x \in X\}),$$

Definition 7. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and $\forall X \subseteq U$, the GOMG Pawlak's lower and upper approximations of X are denoted by $\underline{\sum_{i=1}^m R_i}^{GOMG-PRS}(X)$ and $\overline{\sum_{i=1}^m R_i}^{GOMG-PRS}(X)$.

$$\underline{\sum_{i=1}^{m} R_i}^{GOMG-PRS}(X) = \{x : [x]_{R_1} \subseteq X \vee [x]_{R_2} \subseteq X \vee \cdots \vee [x]_{R_m} \subseteq X, x \in U\}, \quad (25)$$

$$\overline{\sum_{i=1}^{m} R_i}^{GOMG-PRS}(X) = \{x : [x]_{R_1} \cap X \neq \phi \wedge [x]_{R_2} \cap X \neq \phi \wedge \cdots \wedge [x]_{R_m} \cap X \neq \phi, x \in U\}, \quad (26)$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations.

From the definition we obtain the following interpretations:

(1) The GOMG-PRS lower approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^{m} R_i$ with the probability equal to 1.
(2) The GOMG-PRS upper approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^{m} R_i$ with the probability greater than 0.

In fact, the LOMG-PRS model is equal to the GOMG-PRS model.

$$\underline{\sum_{i=1}^{m} R_i}^{LOMG-PRS}(X) = \{x : [x]_{R_1} \subseteq X \vee [x]_{R_2} \subseteq X \vee \cdots \vee [x]_{R_m} \subseteq X, x \in X\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{LOMG-PRS}(X) = \bigcup_{i=1}^{m}(\cup\{[x]_{R_i} : [x]_{R_i} \cap X \neq \phi, x \in X\}),$$
$$\Rightarrow \underline{\sum_{i=1}^{m} R_i}^{GOMG-PRS}(X) = \{x : [x]_{R_1} \subseteq X \vee [x]_{R_2} \subseteq X \vee \cdots \vee [x]_{R_m} \subseteq X, x \in U\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{GOMG-PRS}(X) = \{x : [x]_{R_1} \cap X \neq \phi \wedge [x]_{R_2} \cap X \neq \phi \wedge \cdots \wedge [x]_{R_m} \cap X \neq \phi \wedge, x \in U\}.$$

According to this perspective, LOMG PRSs do not change the semantics of classical rough sets.

4.2. LOMG variable precision rough set model (LOMG-VPRS)

In this subsection, we first define the LOMG-VPRS model.

Definition 8. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and $\forall X \subseteq U$. The LOMG variable precision lower and upper approximations of X are denoted by $\underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X)$ and $\overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X)$.

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \geq \alpha \vee P(X|[x]_{R_2}) \geq \alpha \vee \cdots \vee P(X|[x]_{R_m}) \geq \alpha, x \in X\}, \quad (27)$$

$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{\alpha}(X), \quad (28)$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations and $\overline{R_i}_{\alpha}(X) = \cup\{[x_{R_i} : P(X|[x]_{R_i}) > 1-\alpha, x \in X\}$.

From the definition we obtain the following interpretations:

(1) The LOMG-VPRS lower approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be clearly classified as X using $\sum_{i=1}^{m} R_i$ with the probability greater than or equal to α and being from X.
(2) The LOMG-VPRS upper approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^{m} R_i$ with the probability greater than $1-\alpha$.

By Definitions 5 and 8, we can easily determine the relationship between LOMG-DTRS and LOMG-VPRS. When $\alpha + \beta > 1$ and $0 \leq \beta < \alpha \leq 1$,

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X),$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X) \subseteq \overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X).$$

When $\alpha + \beta < 1$ and $0 \leq \beta < \alpha \leq 1$,

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X),$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta)}(X) \supseteq \overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X).$$

When $\alpha + \beta = 1$ and $0 \le \beta < \alpha \le 1$, the LOMG decision-theoretic rough set LOMG-DTRS becomes the LOMG-VPRS.

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \ge \alpha \vee P(X|[x]_{R_2}) \ge \alpha \vee \cdots \vee P(X|[x]_{R_m}) \ge \alpha, x \in X\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-DTRS,\ \beta=1-\alpha)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{\beta}(X),$$
$$\Rightarrow \underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \ge \alpha \vee P(X|[x]_{R_2}) \ge \alpha \vee \cdots \vee P(X|[x]_{R_m}) \ge \alpha, x \in X\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{\alpha}(X).$$

Definition 9. Given $R_1, R_2, \cdots, R_m \subseteq \mathbf{R}$ m equivalence relations on U and $\forall X \subseteq U$, the LOMG variable precision lower and upper approximations of X are denoted by $\underline{\sum_{i=1}^{m} R_i}^{(GOMG-VPRS,\ \alpha)}(X)$ and $\overline{\sum_{i=1}^{m} R_i}^{(GOMG-VPRS,\ \alpha)}(X)$.

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_1}) \ge \alpha \vee P(X|[x]_{R_2}) \ge \alpha \vee \cdots \vee P(X|[x]_{R_m}) \ge \alpha, x \in U\}, \tag{29}$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha)}(X) = \{x : P(X|[x]_{R_i}) > 1-\alpha \vee P(X|[x]_{R_i}) > 1-\alpha \vee \cdots \vee P(X|[x]_{R_i}) > 1-\alpha, x \in U\}, \tag{30}$$

where $\mathbf{R} = \{R_1, R_2, \cdots, R_m\}$ is a set containing m equivalence relations.

From the definition we obtain the following interpretations:

(1) The GOMG-VPRS lower approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be clearly classified as X using $\sum_{i=1}^{m} R_i$ with the probability greater than or equal to α.
(2) The GOMG-VPRS upper approximation of a set X with respect to $\sum_{i=1}^{m} R_i$ is the set of all elements, which can be classified as X using $\sum_{i=1}^{m} R_i$ with the probability greater than $1-\alpha$.

When $\alpha = 1$, the LOMG-VPRS is equal to the GOMG-VPRS.

$$\underline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha=1)}(X) = \{x : P(X|[x]_{R_1}) \ge 1 \vee P(X|[x]_{R_2}) \ge 1 \vee \cdots \vee P(X|[x]_{R_m}) \ge 1, x \in X\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{(LOMG-VPRS,\ \alpha=1)}(X) = \bigcup_{i=1}^{m} \overline{R_i}_{\alpha}(X),$$
$$\Rightarrow \underline{\sum_{i=1}^{m} R_i}^{(GOMG-VPRS,\ \alpha=1)}(X) = \{x : P(X|[x]_{R_1}) \ge 1 \vee P(X|[x]_{R_2}) \ge 1 \vee \cdots \vee P(X|[x]_{R_m}) \ge 1, x \in U\},$$
$$\overline{\sum_{i=1}^{m} R_i}^{(GOMG-VPRS,\ \alpha=1)}(X) = \{x : P(X|[x]_{R_1}) > 0 \vee P(X|[x]_{R_2}) > 0 \vee \cdots \vee P(X|[x]_{R_m}) > 0, x \in U\}.$$

Based on the analysis above, we can see that the LOMGRS model degrades into different forms of optimistic MGRS models when the two parameters α and β satisfy diverse constraint conditions. Thus, the LOMGRS model is robust and flexible for complex real-world systems.

According to the discussion above, we can determine the relationships among LOMG-DTRS, GOMG-DTRS, LOMG-PRS, GOMG-PRS, LOMG-VPRS, and GOMG-VPRS as shown in Fig. 3.

5. Conclusions

MG-DTRSs where the lower and upper approximations are constructed using multiple granular structures derived from multiple binary relations enhance the flexibility when employing rough set theory for analyzing data. Using the same data background, related granular structures can be selected for specific tasks. However, as the information acquisition technology improves and the number of information sources increases, constructing a raw database is becoming increasingly easier (raw

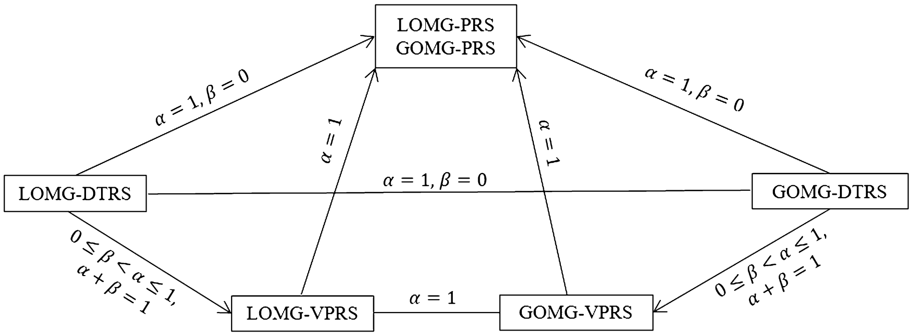


Fig. 3. Relationships among LOMG-DTRS, GOMG-DTRS, LOMG-PRS, GOMG-PRS, LOMG-VPRS, and GOMG-VPRS ("—" and "⟶" indicate that two models are equal and that one model degenerates into the other when the constraint is satisfied).

data comprise data that have not been processed, such as labeling normalization). It is known that MGRS models require a large amount of labeled data. Thus, if only use the data with labels, a vast amount of unlabeled data will be wasted. The LRS proposed by Qian et al. is an excellent solution for dealing with data that have limited labels. In this study, we proposed a new MG-DTRS by combining the MG-DTRS and LRS, which we refer to as the LMG-DTRS. When calculating the upper and lower approximations, most of the time required is consumed obtaining the information granules. Unlike the GOMG-DTRS, the LOMG-DTRS does need not to obtain information granules for all the objects in a given data set in advance, but instead it only computes them for objects from a specific target concept, which is an efficient computation strategy. Moreover, it should be noted that the proposed method can employ more of the information provided by unlabeled data. Indeed, there are many different versions of LMG-DTRSs. For convenience, we considered the LOMG-DTRS as an example to investigate some of its properties and design an algorithm for computing the LOMG lower approximation of a target concept. We also described other types of local MGRS models and explored the relationships among the LOMG-PRS, LOMG-VPRS, and LOMG-DTRS. Furthermore, the proposed approach has several applications. It is well known that the MG-DTRS can be applied to special information systems, such as multi-source information systems, distributive information systems, and groups of multiagent systems. However, if some objects are not labeled in each source, distributive system, or agent in these information systems, as shown in Table 1 (the advent of the big data era means that labeling data is an expensive and laborious task, and sometimes even infeasible, whereas unlabeled data are cheap and easy to obtain, so this situation is very common), the global MG-DTRS cannot deal with these tasks, whereas the LMG-DTRS can.

This study provides a bridge between the MG-DTRS and LRSs, thereby enriching both their theories simultaneously as well as being very helpful for knowledge discovery in various data sets in the context of MG and data with limited labels. Various issues need to be studied in the future, such as LMG-DTRSs for other types of data sets, attribute reduction, and rule extraction.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Nos. 61672332, 61322211, 61432011, U1435212, 11671006 and 61603173), the Program for New Century Excellent Talents in University (No. NCET-12-1031), the Program for the Outstanding Innovative Teams of Higher Learning Institutions of Shanxi (No. 02150116072021), the Program for the Young San Jin Scholars of Shanxi, and the National Key Basic Research and Development Program of China (973) (Nos. 2013CB329404 and 2013CB329502).

References

[1] H.M. Chen, T.R. Li, C. Luo, S.J. Horng, G.Y. Wang, A decision-theoretic rough set approach for dynamic data mining, IEEE Trans. Fuzzy Syst. 23 (6) (2015) 1958–1970.
[2] Y. Du, Q.H. Hu, P.F. Zhu, P.J. Ma, Rule learning for classification based on neighborhood covering reduction, Inf. Sci. 181 (24) (2011) 5457–5467.
[3] B.W. Fang, B.Q. Hu, Probabilistic graded rough set and double relative quantitative decision-theoretic rough set, Int. J. Approx. Reason. 74 (2016) 1–12.
[4] T. Feng, J.S. Mi, Variable precision multigranulation decision-theoretic fuzzy rough sets, Knowl.-Based Syst. 91 (2016) 93–101.
[5] S. Greco, B. Matarazzo, R. Słowiński, Rough approximation by dominance relations, Int. J. Intell. Syst. 17 (2) (2002) 153–171.
[6] S. Greco, B. Matarazzo, R. Słowiński, Parameterized rough set model using rough membership and Bayesian confirmation measures, Int. J. Approx. Reason. 49 (2) (2008) 285–300.
[7] J.P. Herbert, J.T. Yao, Game-theoretic rough sets, Fundam. Inform. 108 (3–4) (2011) 267–286.
[8] Q.H. Hu, D.R. Yu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Inf. Sci. 178 (18) (2008) 3577–3594.
[9] J.H. Li, C.L. Mei, Y.J. Lv, Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction, Int. J. Approx. Reason. 54 (1) (2013) 149–165.
[10] W.T. Li, W.H. Xu, Multigranulation decision-theoretic rough set in ordered information system, Fundam. Inform. 139 (1) (2015) 67–89.
[11] J.H. Li, C.L. Mei, Y. Ren, C.L. Mei, Y.H. Qian, X.B. Yang, A comparative study of multigranulation rough sets and concept lattices via rule acquisition, Knowl.-Based Syst. 91 (2016) 152–164.
[12] D.C. Liang, D. Liu, A novel risk decision making based on decision-theoretic rough sets under hesitant fuzzy information, IEEE Trans. Fuzzy Syst. 23 (2) (2015) 237–247.

[13] D.C. Liang, D. Liu, A. Kobina, Three-way group decisions with decision-theoretic rough sets, Inf. Sci. 345 (2016) 46–64.
[14] D.C. Liang, W. Pedrycz, D. Liu, Determining three-way decisions with decision-theoretic rough sets using a relative value approach, IEEE Trans. Syst. Man Cybern. 99 (2016) 1–15.
[15] J.Y. Liang, F. Wang, C.Y. Dang, Y.H. Qian, An efficient rough feature selection algorithm with a multi-granulation view, Int. J. Approx. Reason. 53 (6) (2012) 912–926.
[16] J.Y. Liang, Y.H. Qian, Information granules and entropy theory in information systems, Sci. China, Ser. F Inform. Sci. 38 (12) (2008) 2048–2065.
[17] J.Y. Liang, Z.Z. Shi, The information entropy, rough entropy and knowledge granulation in rough set theory, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 12 (01) (2004) 37–46.
[18] J.Y. Liang, Z.Z. Shi, D.Y. Li, M.J. Wierman, The information entropy, rough entropy and knowledge granulation in incomplete information system, Int. J. Gen. Syst. 35 (6) (2006) 641–654.
[19] G.P. Lin, Y.H. Qian, J.J. Li, NMGRS: neighborhood-based multigranulation rough sets, Int. J. Approx. Reason. 53 (7) (2012) 1080–1093.
[20] D. Liu, T.R. Li, D.C. Liang, Incorporating logistic regression to decision-theoretic rough sets for classifications, Int. J. Approx. Reason. 55 (1) (2014) 197–210.
[21] A. Mohabey, A.K. Ray, Fusion of rough set theoretic approximations and FCM for color image segmentation, in: IEEE International Conference on Systems, Man, and Cybernetics, vol. 2, 2000, pp. 1529–1534.
[22] M.M. Mushrif, A.K. Ray, Color image segmentation: rough-set theoretic approach, Pattern Recognit. Lett. 29 (4) (2008) 483–493.
[23] S. Nanda, S. Majumdar, Fuzzy rough sets, Fuzzy Sets Syst. 45 (2) (1992) 157–160.
[24] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (5) (1982) 341–356.
[25] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Dordrecht, 1991.
[26] J.F. Peters, M. Borkowski, K-means indiscernibility relation over pixels, in: International Conference on Rough Sets and Current Trends in Computing, Uppsala, in: Lecture Notes in Computer Science, vol. 3066, Springer, 2004, pp. 580–585.
[27] L. Polkowski, A. Skowron, Rough mereology: a new paradigm for approximate reasoning, Int. J. Approx. Reason. 15 (4) (1996) 333–365.
[28] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 225–237.
[29] Y.H. Qian, J.Y. Liang, C.Y. Dang, Rough set method based on multi-granulations, in: Proceedings of 5th IEEE Conference on Cognitive Informatics, vol. 1, 2006, pp. 297–304.
[30] Y.H. Qian, J.Y. Liang, W. Pedrycz, Positive approximation: an accelerator for attribute reduction in rough set theory, Artif. Intell. 174 (9) (2010) 597–618.
[31] Y.H. Qian, J.Y. Liang, W. Wei, Pessimistic rough decision, in: Second International Workshop on Rough Sets Theory, 2010, pp. 440–449.
[32] Y.H. Qian, J.Y. Liang, W.Z. Wu, Information granularity in fuzzy binary GrC model, IEEE Trans. Fuzzy Syst. 19 (2) (2011) 253–264.
[33] Y.H. Qian, J.Y. Liang, Y.Y. Yao, MGRS: a multi-granulation rough set, Inf. Sci. 180 (6) (2010) 949–970.
[34] Y.H. Qian, X.Y. Liang, J.Y. Liang, B. Liu, A.S. Kowron, Y.Y. Yao, J.M. Ma, C.Y. Dang, Local rough set, IEEE Trans. Fuzzy Syst., submitted for publication.
[35] H. Rasiowa, W. Marek, On reaching consensus by groups of intelligent agents, in: Methodologies for Intelligent Systems, vol. 4, 1989, pp. 234–243.
[36] C.M. Rauszer, Rough logic for multi-agent systems, in: International Conference on Logic at Work, Springer, Berlin, Heidelberg, 1992, pp. 161–181.
[37] Y.L. Sang, J.Y. Liang, Y.H. Qian, Decision-theoretic rough sets under dynamic granulation, Knowl.-Based Syst. 91 (2016) 84–92.
[38] R.W. Swiniarski, A. Skowron, Rough set methods in feature selection and recognition, Pattern Recognit. Lett. 24 (2003) 833–849.
[39] Y.H. She, X.L. He, On the structure of the multigranulation rough set model, Knowl.-Based Syst. 36 (2012) 81–92.
[40] D. Slezak, W. Ziarko, The investigation of the Bayesian rough set model, Int. J. Approx. Reason. 40 (1) (2005) 81–91.
[41] B.Z. Sun, W.M. Ma, X. Xiao, Three-way group decision making based on multigranulation, Int. J. Approx. Reason. 81 (2017) 87–102.
[42] J.M. Wei, S.Q. Wang, X.J. Yuan, Ensemble rough hypercuboid approach for classifying cancers, IEEE Trans. Knowl. Data Eng. 22 (3) (2010) 381–391.
[43] M.J. Wierman, Measuring uncertainty in rough set theory, Int. J. Gen. Syst. 28 (4–5) (1999) 283–297.
[44] W.Z. Wu, Y. Leung, Optimal scale selection for multi-scale decision tables, Int. J. Approx. Reason. 54 (8) (2013) 1107–1129.
[45] B.W. Xu, Y.M. Zhou, H.M. Lu, An improved accuracy measure for rough sets, J. Comput. Syst. Sci. 71 (2) (2005) 163–173.
[46] W.H. Xu, Q.R. Wang, X.T. Zhang, Multi-granulation rough sets based on tolerance relations, Soft Comput. 17 (7) (2013) 1241–1252.
[47] H.L. Yang, Z.L. Guo, Multigranulation decision-theoretic rough sets in incomplete information systems, Int. J. Mach. Learn. Cybern. 6 (6) (2015) 1005–1018.
[48] X.A. Yang, G.Y. Wang, H. Yu, T.R. Li, Decision region distribution preservation reduction in decision-theoretic rough set model, Inf. Sci. 278 (2014) 614–640.
[49] X.B. Yang, X.N. Song, H.L. Dou, J.Y. Yang, Multi-granulation rough set: from crisp to fuzzy case, Ann. Fuzzy Math. Inform. 1 (1) (2011) 55–70.
[50] Y.Y. Yao, Decision-theoretic rough set models, in: International Conference on Rough Sets and Knowledge Technology, Springer, Berlin, Heidelberg, 2007, pp. 1–12.
[51] Y.Y. Yao, Probabilistic rough set approximations, Int. J. Approx. Reason. 49 (2) (2008) 255–271.
[52] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. 180 (3) (2010) 341–353.
[53] Y.Y. Yao, Probabilistic approaches to rough sets, Expert Syst. 20 (5) (2003) 287–297.
[54] Y.Y. Yao, B. Zhou, Two bayesian approaches to rough sets, Eur. J. Oper. Res. 251 (3) (2016) 904–917.
[55] Y.Y. Yao, S.K.M. Wong, P. Lingras, A decision-theoretic rough set model, in: Z.W. Ras, M. Zemankova, M.L. Emrich (Eds.), Methodologies for Intelligent Systems, vol. 5, North-Holland, New York, 1990, pp. 17–24.
[56] Y.Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, Inf. Sci. 178 (17) (2008) 3356–3373.
[57] T. Young, Data mining and machine oriented modeling: a granular computing approach, Appl. Intell. 13 (2) (2000) 113–124.
[58] H. Yu, Z.G. Liu, G.Y. Wang, An automatic method to determine the number of clusters using decision-theoretic rough set, Int. J. Approx. Reason. 55 (1) (2014) 101–115.
[59] X.D. Yue, Y.F. Chen, D.Q. Miao, J. Qian, Tri-partition neighborhood covering reduction for robust classification, Int. J. Approx. Reason. (2017), http://dx.doi.org/10.1016/j.ijar.2016.11.010.
[60] X.Y. Zhang, D.Q. Miao, Two basic double-quantitative rough set models of precision and grade and their investigation using granular computing, Int. J. Approx. Reason. 54 (8) (2013) 1130–1148.
[61] W. Ziarko, Probabilistic approach to rough sets, Int. J. Approx. Reason. 49 (2) (2008) 272–284.
[62] W. Ziarko, Variable precision rough set model, J. Comput. Syst. Sci. 46 (1) (1993) 39–59.