Granul. Comput. (2017) 2:271–288
DOI 10.1007/s41066-017-0042-9

ORIGINAL PAPER

Generalized multigranulation rough sets and optimal granularity selection

Weihua Xu[1] · Wentao Li[2] · Xiantao Zhang[1]

Received: 27 September 2016 / Accepted: 3 March 2017 / Published online: 13 April 2017

Abstract Multigranulation rough set theory is a desirable direction in the field of rough set, in which upper and lower approximations are approximated by multiple granular structures. However, classic multigranulation rough set is studied from two kinds of qualitative combination rules which were generated by pessimistic and optimistic viewpoints, respectively. The two combination rules seem to lack of practicability since one is too restrictive and the other too relaxed. To overcome this disadvantage, we propose a generalized multigranulation rough set model in this paper. First, we discuss upper and lower approximation sets of a generalized multigranulation rough set by introducing a support characteristic function and an information level. Then, as one of the most important problems in granular computing, we carefully study how to select optimal granularity in generalized multigranulation rough sets. Furthermore, algorithms of optimal granularity selection are constructed, by which we can provide an efficient approach to compute the optimal granularity based on generalized multigranulation rough sets. Finally, an illustrative example is given to show the effectiveness of the proposed approach. The main contribution of this paper is to construct the model of the optimal particle size selection on account of the generalized multi granularity, and overcome the limitation of the classical multi granularity.

Keywords Multigranulation rough set · Optimal granularity selection · Quality function · Support characteristic function

✉ Weihua Xu
chxuwh@gmail.com

1 School of Sciences, Chongqing University of Technology, Chongqing 400054, People's Republic of China

2 Department of Mathematics, Harbin Institute of Technology, Harbin 150001, People's Republic of China

1 Introduction

Rough set theory proposed by Pawlak (1982), is an extension of the classical set theory. This theory could be regarded as a mathematical and soft computing tool to handle vagueness, imprecision and uncertainty in data analysis. The methodology has received great attention in recent years, and it has been successfully applied in many science and engineering fields, such as pattern recognition, data mining, image processing, medical diagnosis and so on. Rough set theory was developed by practical needs to characterize, interpret, present, process indiscernibility of individuals (objects). The discernibility is typically characterized by an equivalence relation. A rough set is the result of two approximating crisp sets, which are lower and upper approximation sets, using equivalence classes. The key idea of rough set theory is the use of some known knowledge to approximate the inaccurate and uncertain knowledge in information systems.

However, partition or indiscernibility relation in Pawlak's original rough set theory, is still restrictive for many applications. To overcome such unreasonableness, the dominance-based rough set approach has been proposed by Greco et al. (1999). On the other hand, the generalization of rough sets is an interesting topic not only in mathematical point of view but also in practical point of view. Along this direction, rough sets have been generalized under similarity relations (Inuiguchi and Tanino 2001), covers (Bonikowski et al. 1998) and general relations (Inuiguchi and Tanino 2002; Shen and Jensen 2007; Yao 1996, 1998; Yao and Lin 1996). Those results demonstrate a diversity of generalizations. Moreover, the introduction of fuzziness into rough set approaches (Lu et al. 2016) has attracted researchers to obtain more realistic and useful tools.

In 1985, Zadeh first explored the concept of granular computing (Zadeh 1997) between 1996 and 1997. He thought that information granules refer to pieces, classes, and groups into which complex information are divided in accordance with the characteristics and processes of the understanding and decision-making. Currently, granular computing has been viewed as an emerging computing paradigm of information processing. It concerns the processing of complex information entities called information granules (Li et al. 2015; Pedrycz 2013; Pedrycz and Bargiela 2002, 2012; Xu and Li 2014). Information granules, as encountered in natural language, are implicit in their nature. To make them fully operational so that they become effectively used in the analysis and design of intelligent systems (Pedrycz 2013), we need to make information granules explicit. This is possible through a prudent formalization available within the realm of granular computing. In a general sense, by information granule, one regards a collection of elements drawn together by their closeness (resemblance, proximity, functionality, etc.) articulated in terms of some useful spatial, temporal or functional relationships. Subsequently, Granular Computing is about representing, constructing and processing information granules. Results of computing completed in the setting of Granular Computing come in the form of information granules. Information granules are building blocks reflective of domain knowledge about a problem.

From the perspective of granular computing (Liu et al. 2016), an equivalence relation on the universe can be regarded as a granularity, and the corresponding partition can be regarded as a granular structure (Qian et al. 2009). Hence, the classic rough set theory is based on a single granularity (only one equivalence relation). However, the rough set may be associated with multiple granular structures (Apolloni et al. 2016), which can be divided into two cases as follows:

Case 1 If there exists at least one granular structure such that elements surely belongs to a given concept, then we say that an element surely belong to the concept.

Case 2 If there exists at least one granular structure such that elements possibly belongs to a given concept, then we say that an element possibly belong to the concept.

Currently, Yao and Deng (2014) proposes a framework of quantitative rough sets based on subsethood measures. A specific quantitative rough set model is defined by a particular class of subsethood measures satisfying a set of axioms. Consequently, the framework enables us to classify and unify existing generalized rough set models [e.g., decision-theoretic rough sets (Xu and Wang 2016), probabilistic rough sets, and variable precision rough sets], to investigate limitations of existing models, and to develop new models. Actually, an attribute subset (Guo and Zheng 2014) induces an equivalence relation, and a partition formed by the equivalence relation can be regard as a granularity. Using a finer granular structure formed through combining two known granularities induced from two attribute subsets to describe a target concept, this combination destroys the original granular structure. Qian and Liang extended Pawlak's single granulation rough set model to a multiple granulation rough set model (Qian and Liang 2006; Qian et al. 2010), where the set approximations were defined using multi equivalence relations on the universe. Moreover, many researchers have extended the multigranulation rough sets (Dou et al. 2012). Xu et al. (2014) developed a multigranulation fuzzy rough set model, multigranulation rough sets based on tolerance relations (Xu et al. 2013), a multigranulation rough set model in ordered information systems (Xu et al. 2012) and a multigranulation fuzzy rough set in a fuzzy tolerance approximation space (Xu et al. 2011). Yang et al. proposed the hierarchical structure properties of the multigranulation rough sets (Yang et al. 2012), multigranulation rough set in incomplete information system (Yang et al. 2012a, b), and a test cost sensitive multigranulation rough set model (Yang and Qi 2013). Lin et al. (2012) presented a neighborhood-based multigranulation rough set. She et al. (2012) explored the topological structures and the properties of multigranulation rough sets. Qian et al. (2014) introduced three kinds of multigranulation decision-theoretic rough set models. Li et al. (2014) extended multigranulation decision-theoretic rough sets by considering dominance relations in ordered information system (Li and Xu 2015), and investigated relationships between multigranulation and classical T-fuzzy rough sets. Yao et al. proposed a unified framework to classify and compare existing studies. And an underlying principle is to explain rough sets in a multigranulation space through rough sets derived using individual equivalence relations (Yao and She 2016). Feng and Mi (2015) studied variable precision multigranulation fuzzy decision-theoretic rough sets in an information system. A novel membership degree based on single granulation rough sets and two operators based on this membership degree were investigated in their study. Zhang et al. (2015) established four kinds of constructive methods of rough approximation operators from existing rough sets and studied the non-dual multigranulation rough sets and hybrid multigranulation rough sets. Tan et al. (2016) employed the belief and plausibility functions from evidence theory to characterize the set approximations and attribute reductions in multigranulation rough set theory in incomplete information systems, and an attribute reduction algorithm for multigranulation rough sets was proposed based on evidence theory. Lin et al. (2015) proposed a two-grade fusion approach involved in the evidence theory and multigranulation rough set theory based on a well-defined distance function among granulation structures, and presented three types of covering based multigranulation rough sets whose set approximations were defined by different covering approximation operators (Lin et al. 2013). Li et al. investigated the relationship between multigranulation rough sets and concept lattices via rule acquisition (Li et al. 2016; Yang et al. 2009). Kumar and Inbarani applied rough set based data mining techniques for medical data to discover locally frequent diseases (Senthil Kumar and Hannah Inbarani 2015). Huang et al. (2014) developed a new multigranulation rough set model that was called intuitionistic fuzzy multigranulation rough set (IFMGRS) and three types of IFMGRSs that are generalizations of three existing intuitionistic fuzzy rough set models built. Liu et al. proposed four types of multi-granulation covering rough set (MGCRS) models under covering approximation space (Wang et al. 2017a), where a target concept was approximated by employing the maximal or minimal descriptors of objects in a given universe of discourse (Liu et al. 2014). Lin et al. (2014) presented a new feature selection method that selects distinguishing features by fusing neighborhood multi-granulation, and first used neighborhood rough sets as an effective granular computing tool. Yang et al. (2014) first explored the updating of the multigranulation rough approximations. Qian et al. (2014b) develop a new multigranulation rough set model based on "Seeking common ground while eliminating differences" (SCED) strategy, called pessimistic multigranulation rough sets based decision. Liang et al. (2012) proposed an efficient rough feature selection algorithm for large-scale data sets, which was stimulated from multi-granulation rough sets.

In multigranulation rough set theory, optimistic multigranulation (Wang et al. 2017b) and pessimistic multigranulation are two basic ways of research. For the lower approximation of a multigranulation rough set, the view of optimistic multigranulation reflects that there exists at least one granular structure such that elements surely belong to a given concept, and the view of pessimistic multigranulation shows that elements surely belong to a given concept in each granular structure. It is easy to notice that both optimistic and pessimistic conditions are too strict to a wide-range of conditions. So, we will introduce a parameter, namely an information level to propose a generalized multigranulation rough set model. The lower approximation of a concept is the set that all of the elements support the concept based on the information level is not less than the given parameter in the multigranulation perspective.

The motivation of this paper is as follows: three aspects. (1) How to generalize the classical multigranularity rough sets to the generalized multigranularity rough sets. (2) how to select the proper granularity is an important issue. It offers a systematic and theoretic framework for feature selection. (3) How to discover knowledge in hierarchically organized information tables is of particular importance in real life data mining. In this paper, to describe a novel granulation perspective, we will establish a special multigranulation rough set model, and discuss the methods of optimal granularity selection in this generalized multigranulation rough set model. This paper is organized as follows. In Sect. 2, some preliminary concepts of optimistic and pessimistic multigranulation rough set theories are briefly reviewed. In Sect. 3, we introduce a support characteristic function and propose the generalized multigranulation rough set model. In Sect. 4, Moreover, measures and properties of the generalized multigranulation rough set are carefully investigated. How to select the optimal granularity is discussed in generalized multigranulation rough set. In Sect. 5, we consider algorithms of the optimal granularity selection in the new rough set model. In Sect. 6, an illustrative example is given to show the effectiveness of the proposed approach. Finally, we conclude our contribution with a summary and an outlook for the further research.

2 Classic multigranulation rough sets

In an information system, the equivalence class of an object with respect to an attribute subset of A is a granularity from the viewpoint of granular computing. A partition of the universe is a granular structure. Rough set proposed by Pawlak is a single granulation rough set model, and the granular structure in this model is induced by the indiscernibility relation of the attribute set. In general, the above cases cannot always be satisfied or required in practical problems. In the three cases referred in reference (Qian et al. 2010), there are limitations in single granulation rough set for addressing practical problems with multiple partitions, and multigranulation rough set can now be used to effectively solve these problems. Under those circumstances, we must describe a target concept through multiple binary relations on the universe according to a user's requirements or targets of problem solving. In the literature (Qian et al. 2009, 2010; Dang and Qian 2009; Yao 2000), to apply rough set theory to practical problems widely, multigranulation rough set model has been studied based on multiple equivalence relations.

Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A$. Then P_i or U/P_i is referred to as a granulation. The equivalence class of an object x with respect to P_i is defined as $[x]_{P_i} = \{y \in U | f(x, a) = f(y, a)\}$ $(a \in P_i)$. The lower and upper approximation sets of X with respect to single P_i are defined as follows:

$$\underline{P_i}(X) = \{x \in U | [x]_{P_i} \subseteq X\},$$

$$\overline{P_i}(X) = \{x \in U | [x]_{P_i} \cap X \neq \emptyset\}.$$

Considering further studies on multigranulation rough set, we now review the two basic forms of multigranulation rough set model.

Definition 2.1 (Qian et al. 2010) Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \ldots, l)$. The optimistic multigranulation lower and upper approximation sets of X with respect to single P are defined as follows:

$$\underline{P}(X)_{\mathrm{OM}} = \{x \in U | \vee ([x]_{P_i} \subseteq X),\ i \leq l\};$$

$$\overline{P}(X)_{\mathrm{OM}} = \{x \in U | \wedge ([x]_{P_i} \cap X \neq \emptyset),\ i \leq l\},$$

where "$\vee$" means the logical operator "or", which represents that the alternative conditions are satisfied, and "$\wedge$" means the logical operator "and", which represents that all of the conditions are satisfied.

The set X is definable if and only if $\underline{P}(X)_{\mathrm{OM}} = \overline{P}(X)_{\mathrm{OM}}$. Otherwise, X is rough. $\underline{P}(X)_{\mathrm{OM}}$ and $\overline{P}(X)_{\mathrm{OM}}$ are referred to as optimistic lower and upper approximation sets, respectively.

From the above definition, the operators "$\vee$" and "$\wedge$" can be exchanged between the optimistic lower approximation set and the optimistic upper approximation set. Corresponding to optimistic multigranulation rough set, pessimistic multigranulation rough set model can be defined in the following.

Definition 2.2 (Qian et al. 2010) Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \ldots, l)$. The pessimistic multigranulation lower and upper approximation sets of X with respect to single P are defined as follows:

$$\underline{P}(X)_{\mathrm{PM}} = \{x \in U | \wedge ([x]_{P_i} \subseteq X),\ i \leq l\},$$

$$\overline{P}(X)_{\mathrm{PM}} = \{x \in U | \vee ([x]_{P_i} \cap X \neq \emptyset),\ i \leq l\}.$$

The set X is definable if and only if $\underline{P}(X)_{\mathrm{PM}} = \overline{P}(X)_{\mathrm{PM}}$. Otherwise, X is rough. $\underline{P}(X)_{\mathrm{PM}}$ and $\overline{P}(X)_{\mathrm{PM}}$ are referred to as pessimistic lower and upper approximation sets, respectively.

The uncertainty of a concept in a multigranulation rough set model is also due to the existence of a boundary region. The greater the boundary of a concept is, the lower its accuracy is, and the coarser the concept is. Similar to the measures in the Pawlak rough set model, the accuracy and roughness measures in optimistic multigranulation rough set and pessimistic multigranulation rough set were defined in the same way (Qian et al. 2010). As generalizations of the Pawlak rough set model, we only show the relations among optimistic multigranulation rough set, pessimistic multigranulation rough set and single granulation rough set in the following.

Proposition 2.1 (Qian et al. 2010) *Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \ldots, l)$. The following properties hold*:

(1) $\underline{P}(X)_{OM} = \bigcup_{i=1}^{l} \underline{P_i}(X)$;

(2) $\overline{P}(X)_{OM} = \bigcap_{i=1}^{l} \overline{P_i}(X)$;

(3) $\underline{P}(X)_{PM} = \bigcap_{i=1}^{l} \underline{P_i}(X)$;

(4) $\overline{P}(X)_{PM} = \bigcup_{i=1}^{l} \overline{P_i}(X)$;

(5) $\underline{P}(X)_{PM} \subseteq \underline{P}(X)_{OM}$;

(6) $\overline{P}(X)_{OM} \subseteq \overline{P}(X)_{PM}$.

In addition, there are many related properties as well as proof please refer to (Qian and Liang 2006; Qian et al. 2010).

3 Generalized multigranulation rough sets

To present and illustrate the generalized multigranulation rough set model, we first define the support characteristic function.

Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \ldots, l)$. A characteristic function $S_X^{P_i}(x)$ describes the inclusion relation between the class $[x]_{P_i}$, which is defined in the following:

$$S_X^{P_i}(x) = \begin{cases} 1, & [x]_{P_i} \subseteq X; \\ 0, & \text{else.} \end{cases} \quad (i \leq l).$$

We then call $S_X^{P_i}(x)$ the support characteristic function of x. It shows whether x supports the concept X or not precisely with respect to P_i.

From the support characteristic function for any $P_i \subseteq A$ $(i = 1, 2, \ldots, l)$ and $X \subseteq U$, the number of equivalence classes $[x]_{P_i}$ that satisfies $[x]_{P_i} \subseteq X$ can be computed by

$$\sum_{i=1}^{l} S_X^{P_i}(x).$$

At the same time, for any $x \in U$, the number of equivalence classes $[x]_{P_i}$ that satisfies $[x]_{P_i} \cap X \neq \emptyset$ can be represented by

$$\sum_{i=1}^{l}(1 - S^{P_i}_{\sim X}(x)).$$

By the support characteristic function, the lower and upper approximation sets in optimistic multigranulation rough set and pessimistic multigranulation rough set can be represented, respectively, with the following formulas:

$$\underline{P}(X)_{\mathrm{OM}} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l} S^{P_i}_{X}(x)}{l} > 0 \right\};$$

$$\overline{P}(X)_{\mathrm{OM}} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l}(1 - S^{P_i}_{\sim X}(x)).}{l} = 1 \right\}.$$

$$\underline{P}(X)_{\mathrm{PM}} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l} S^{P_i}_{X}(x)}{l} = 1 \right\};$$

$$\overline{P}(X)_{\mathrm{PM}} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l}(1 - S^{P_i}_{\sim X}(x)).}{l} > 0 \right\}.$$

From the view of granular computing, one can note that classical multigranulation rough set might not always be effective in practice. Optimistic multigranulation rough set might be so loose that the approximation sets cannot describe the concepts as precisely as possible. Additionally, pessimistic multigranulation rough set might be too strict to depict concepts on the universe.

In optimistic multigranulation rough set, we consider the case that an object x supports the concept X precisely if there exists at least one $P_i \subseteq P$ such that $[x]_{P_i} \subseteq X$. This model could bring in a large amount of useless information for the concept described. The descriptions and information in optimistic multigranulation rough set could be redundant and cannot show the nature of the concept. Furthermore, some useful information will be lost because this model demands that any object x can possibly describe a concept X in terms of multigranulation satisfying $[x]_{P_i} \cap X \neq \emptyset$ for all P_i. In practical applications, the object can possibly describe a concept by most of the granulations.

Conversely, an object x supporting the concept X precisely means that $[x]_{P_i} \subseteq X$ must hold for all P_i with respect to multigranulation in pessimistic multigranulation rough set. This approach also causes disadvantages in practice. If x supports the concept, then all granulations must be considered. This approach is so strict that some information and descriptions which are not very effective can be ignored. Thus, we can introduce the parameter β, i.e., the information level, to realize that the objects support a concept with respect to the majority granulations. The higher the information level β is, the stricter our requirements are. Our requirements can be employed to depict the concept better. Next, we will propose a novel multigranulation rough set model with a parameter $\beta \in (0.5, 1]$.

Definition 3.1 Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l, \}$ $P_i \subseteq A (i = 1, 2, \ldots, l)$. A characteristic function $S^{P_i}_{X}(x)$ describes the inclusion relation between the class $[x]_{P_i}$. For any $\beta \in (0.5, 1]$, generalized lower and upper approximation sets of X with respect to P are defined as follows:

$$\underline{P}(X)_{\beta} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l} S^{P_i}_{X}(x)}{l} \geq \beta \right\};$$

$$\overline{P}(X)_{\beta} = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l}(1 - S^{P_i}_{\sim X}(x)).}{l} > 1 - \beta \right\}.$$

The set X is definable if and only if $\underline{P}(X)_{\beta} = \overline{P}(X)_{\beta}$. Otherwise, X is rough. We call β the information level with respect to P.

For classic rough set model, the roughness or uncertainty in an information system is also due to the existence of a boundary region of concepts in generalized multigranulation rough set. The boundary region of a concept with respect to P in generalized multigranulation rough set is defined by

$$Bn(X)_{\mathrm{GM}} = \overline{P}(X)_{\beta} - \underline{P}(X)_{\beta}.$$

Objects in approximation sets and boundary regions are changed corresponding to the information level β. Additionally, we can have the following interpretations to approximation sets and boundary regions in generalized multigranulation rough set.

- The lower approximation set of a concept X is the set of all of the elements that can surely support the concept X on the basis of an information level not less than β in terms of the multigranulation.
- The upper approximation set of a concept X is the set of all of the elements that can possibly support the concept X on the basis of an information level not less than $1 - \beta$ in terms of the multigranulation.
- The boundary region of a concept X with respect to P is the set of all of the elements that cannot surely support either X or $\sim X$ on the basis of an information level β.

That is, $Bn(X)_{\mathrm{GM}}$ is the set of all of the elements that cause the uncertainty of X in an information system with respect to P on the basis of the information level β.

From the above, we can easily compare the approximation sets in optimistic multigranulation rough set, pessimistic multigranulation rough set and generalized multigranulation rough set. Next, we will study the relations between generalized multigranulation rough set, optimistic multigranulation rough set and pessimistic multigranulation rough set, and discuss some important properties of the approximation operators in generalized multigranulation rough set.

Figure 1 shows the relationships of the lower and upper approximations in generalized multigranulation rough set, optimistic multigranulation rough set and pessimistic multigranulation rough set, respectively.

Proposition 3.1 *Let $I=(U,A,V,f)$ be an information system, $X \subseteq U$ and $P=\{P_1,P_2,\ldots,P_l\}$, $P_i \subseteq A (i=1,2,\ldots,l)$. For any $\beta \in (0.5,1]$, the lower and upper approximation operators in generalized multigranulation rough set have the following relations with those in optimistic multigranulation rough set and pessimistic multigranulation rough set* (*Fig.* 1):

(1) $\underline{\underline{P}}(X)_{PM} \subseteq \underline{\underline{P}}(X)_{\beta} \subseteq \underline{\underline{P}}(X)_{OM}$;
(2) $\overline{\overline{P}}(X)_{OM} \subseteq \overline{\overline{P}}(X)_{\beta} \subseteq \overline{\overline{P}}(X)_{PM}$.

Proof

(1) For any x and β, one can prove $\underline{P}(X)_{\mathrm{PM}} \subseteq \underline{P}(X)_{\beta}$ through $x \in \underline{P}(X)_{\mathrm{PM}} \Leftrightarrow \frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq 1$. As $\beta \in (0.5,1]$, so $\frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq 1 \geq \beta$, that is to say $x \in \underline{P}(X)_{\beta}$. Similarly, for any $\beta \in (0.5,1]$, $x \in \underline{P}(X)_{\beta} \Leftrightarrow \frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta > 0 \Rightarrow x \in \underline{P}(X)_{\mathrm{OM}}$. Then $\underline{P}(X)_{\beta} \subseteq \underline{P}(X)_{\mathrm{OM}}$. This item is proved.
(2) This item can be obtained similarly.

□

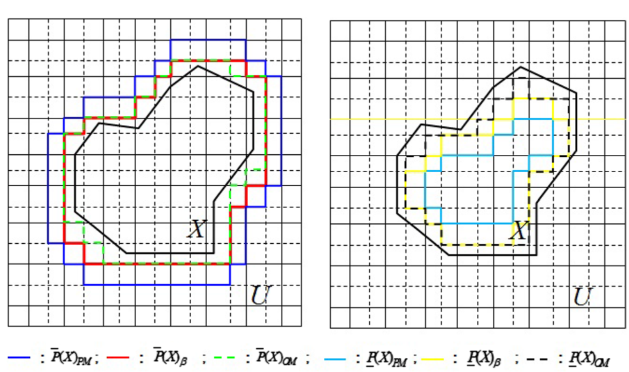

Fig. 1 Relationships of upper and lower approximations

Lemma 3.1 *For any $a_1,a_2,b_1,b_2 \in [0,1]$, the following inequalities hold*:

(1) $a_1 \wedge b_1 + a_2 \wedge b_2 \leq (a_1+a_2) \wedge (b_1+b_2)$;
(2) $a_1 \vee b_1 + a_2 \vee b_2 \geq (a_1+a_2) \vee (b_1+b_2)$,

where "$\wedge$" *and* "$\vee$" *represent the operators* "minimum" and "maximum", *respectively*.

Proof

(1) Because $a_1 \wedge b_1 \leq a_1$, $a_1 \wedge b_1 \leq b_1$ and $a_2 \wedge b_2 \leq a_2$, $a_2 \wedge b_2 \leq b_2$, we have that $a_1 \wedge b_1 + a_2 \wedge b_2 \leq (a_1+a_2)$. $a_1 \wedge b_1 + a_2 \wedge b_2 \leq b_1 + b_2$.Then we have $a_1 \wedge b_1 + a_2 \wedge b_2 \leq (a_1+a_2) \wedge (b_1+b_2)$.
(2) This item can be obtained similarly.

Proposition 3.2 *Let $I=(U,A,V,f)$ be an information system, $X \subseteq U$ and $P=\{P_1,P_2,\ldots,P_l\}$, $P_i \subseteq A$ $(i=1,2,\ldots,l)$. For any $\beta \in (0.5,1]$. The following properties are true.*

(1a) $\underline{\underline{P}}(\sim X)_{\beta} = \sim \overline{P}(X)_{\beta}$;
(1b) $\overline{P}(\sim X)_{\beta} = \sim \underline{P}(X)_{\beta}$.
(2a) $\underline{\underline{P}}(X)_{\beta} \subseteq X$;
(2b) $X \subseteq \overline{P}(X)_{\beta}$.
(3a) $\underline{\underline{P}}(\emptyset)_{\beta} = \overline{P}(\emptyset)_{\beta} = \emptyset$;
(3b) $\underline{P}(U)_{\beta} = \overline{P}(U)_{\beta} = U$.
(4a) $\underline{\underline{P}}(X \cap Y)_{\beta} \subseteq \underline{\underline{P}}(X)_{\beta} \cap \underline{\underline{P}}(Y)_{\beta}$;
(4b) $\overline{P}(X \cup Y)_{\beta} \supseteq \overline{P}(X)_{\beta} \cup \overline{P}(Y)_{\beta}$.
(5a) $\underline{\underline{P}}(X \cup Y)_{\beta} \supseteq \underline{\underline{P}}(X)_{\beta} \cup \underline{\underline{P}}(Y)_{\beta}$;
(5b) $\overline{P}(X \cap Y)_{\beta} \subseteq \overline{P}(X)_{\beta} \cap \overline{P}(Y)_{\beta}$.
(6a) $X \subseteq Y \Rightarrow \underline{\underline{P}}(X)_{\beta} \subseteq \underline{\underline{P}}(Y)_{\beta}$;
(6b) $X \subseteq Y \Rightarrow \overline{P}(X)_{\beta} \subseteq \overline{P}(Y)_{\beta}$.
(7a) $X \cap Y = \emptyset \Rightarrow \underline{\underline{P}}(X)_{\beta} \cap \underline{\underline{P}}(Y)_{\beta} = \emptyset$;
(7b) $X \cap Y = \emptyset \Rightarrow \overline{P}(X)_{\beta} \cap \overline{P}(Y)_{\beta} = \emptyset$.

Proof

(1a) For any $x \in U$,

$$x \in \overline{P}(X)_{\beta} \Leftrightarrow \frac{\sum_{i=1}^{l}(1 - S_{\sim X}^{P_i}(x)).}{l} > 1 - \beta,$$

we have that

$$x \in \sim \overline{P}(X)_\beta \Leftrightarrow \frac{\sum_{i=1}^{l}(1 - S_{\sim X}^{P_i}(x)).}{l} \leq 1 - \beta$$
$$\Leftrightarrow \frac{\sum_{i=1}^{l} S_{\sim X}^{P_i}(x).}{l} \geq \beta \Leftrightarrow \underline{P}(\sim X)_\beta.$$

Then this item is proved. Item (1b) can be proved similarly.

(2a) For any $x \in \underline{P}(X)_\beta$, we have

$$\frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta > 0.$$

There exists $i \leq l$ such that $[x]_{P_i} \subseteq X$. Thus, we can get $x \in X$.

(2b) By the duality and item (2a), we have

$$\sim \overline{P}(X)_\beta = \underline{P}(\sim X)_\beta \subseteq \sim X.$$

Thus, $X \subseteq \overline{P}(X)_\beta$.

(3a) (3b)From item (2) in Proposition 3.1, $S_\emptyset^{P_i}(x) = 0$ and $S_U^{P_i}(x) = 1$ ($\forall X \in U$), then we can have

$$\underline{P}(\emptyset)_\beta = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l} S_\emptyset^{P_i}(x)}{l} = \frac{\sum_{i=1}^{l} 0}{l} = 0 \geq \beta \right\} = \emptyset,$$
$$\underline{P}(U)_\beta = \left\{ x \in U \middle| \frac{\sum_{i=1}^{l} S_U^{P_i}(x)}{l} = \frac{\sum_{i=1}^{l} 1}{l} = 1 \geq \beta \right\} = U.$$

From the duality, we can easily have

$$\overline{P}(\emptyset)_\beta = \overline{P}(\sim U)_\beta = \sim \underline{P}(U)_\beta = \sim U = \emptyset.$$

$$\overline{P}(U)_\beta = \overline{P}(\sim \emptyset)_\beta = \sim \underline{P}(\emptyset)_\beta = \sim \emptyset = U.$$

(4a) For any $x \in \underline{P}(X \cap Y)_\beta$, we can get

$$\frac{\sum_{i=1}^{l} S_{X\cap Y}^{P_i}(x)}{l} = \frac{\sum_{i=1}^{l} S_X^{P_i}(x) \wedge S_Y^{P_i}(x)}{l} \geq \beta.$$

By Lemma 3.1, we have

$$\frac{\sum_{i=1}^{l} S_X^{P_i}(x) \wedge \sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \frac{\sum_{i=1}^{l} S_X^{P_i}(x) \wedge S_Y^{P_i}(x)}{l}.$$

Then, we have the following

$$\frac{\sum_{i=1}^{l} S_X^{P_i}(x) \wedge \sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \beta$$
$$\Leftrightarrow \left(\frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta \right) \wedge \left(\frac{\sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \beta \right)$$
$$\Leftrightarrow x \in \underline{P}(X)_\beta \wedge x \in \underline{P}(Y)_\beta.$$

Therefore $x \in \underline{P}(X)_\beta \cap \underline{P}(Y)_\beta$.

(4b) From the duality, this item can be proved by item (4a) in this proposition.

(5a) From Proposition 3.1, for any $x \in U$, $x \in \underline{P}(X)_\beta \cup \underline{P}(Y)_\beta$, it means that $x \in \underline{P}(X)_\beta \vee \underline{P}(Y)_\beta$,

$$\left(\frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta \right) \vee \left(\frac{\sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \beta \right)$$
$$\Leftrightarrow \frac{\sum_{i=1}^{l} S_X^{P_i}(x) \vee \sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \beta.$$

By Lemma 3.1, one can have

$$\frac{\sum_{i=1}^{l} S_X^{P_i}(x) \vee S_Y^{P_i}(x)}{l} \geq \frac{\sum_{i=1}^{l} S_X^{P_i}(x) \vee \sum_{i=1}^{l} S_Y^{P_i}(x)}{l}.$$

Then,

$$\frac{\sum_{i=1}^{l} S_{X\cup Y}^{P_i}(x)}{l} \geq \frac{\sum_{i=1}^{l} S_X^{P_i}(x) \vee S_Y^{P_i}(x)}{l} \geq \beta.$$

It means that $x \in \underline{P}(X \cup Y)_\beta$.

(5b) From duality, we can get it easily.

(6a) For any $x \in \underline{P}(X)_\beta$, we have $\frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta$. As $X \subseteq Y$, we have $S_X^{P_i}(x) \leq S_Y^{P_i}(x)$. Then we can get

$$\frac{\sum_{i=1}^{l} S_Y^{P_i}(x)}{l} \geq \frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta.$$

Thus, $x \in \underline{P}(Y)_\beta$ is obtained. Then this item is proved.

(6b) Similarly, $x \in \overline{P}(X)_\beta$, we have $\frac{\sum_{i=1}^{l}(1-S_{\sim X}^{P_i}(x))}{l} > 1 - \beta$. As $X \subseteq Y$, we have $S_{\sim X}^{P_i}(x) \geq S_{\sim Y}^{P_i}(x)$. Then

$$\frac{\sum_{i=1}^{l}(1 - S_{\sim Y}^{P_i}(x))}{l} \geq \frac{\sum_{i=1}^{l}(1 - S_{\sim X}^{P_i}(x))}{l} > 1 - \beta.$$

Additionally, $x \in \overline{P}(Y)_\beta$ is obtained.

(7a) From $X \cap Y = \emptyset$, we can directly obtain that $X \subseteq \sim Y$. We then have $\underline{P}(X)_\beta \subseteq \underline{P}(\sim Y)_\beta$. Moreover, from the duality and items (2a) (2b) in this proposition, we have $\underline{P}(\sim Y)_\beta = \sim \overline{P}(Y)_\beta \subseteq \underline{P}(Y)_\beta$. Thus, we can develop that $\underline{P}(X)_\beta \subseteq \sim \underline{P}(Y)_\beta$. That is to say $\underline{P}(X)_\beta \cap \underline{P}(Y)_\beta = \emptyset$. This item is proved.

(7b)It is easy to get this item.

□

Remark 1 The properties $\underline{P}(\underline{P}(X)_\beta)_\beta = \underline{P}(X)_\beta = \overline{P}(\underline{P}(X)_\beta)_\beta$ and $\overline{P}(\overline{P}(X)_\beta)_\beta = \overline{P}(X)_\beta = \underline{P}(\overline{P}(X)_\beta)_\beta$ do not hold in generalized multigranulation rough set.

For different information levels α and β, the following properties can be obtained.

Proposition 3.3 *Let* $I = (U, A, V, f)$ *be an information system,* $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A(i = 1, 2, \ldots, l)$. *For any* $\alpha \leq \beta$ and $\alpha, \beta \in (0.5, 1]$, *the following properties are true*.

(1) $\underline{P}(X)_\beta \subseteq \underline{P}(X)_\alpha$;
(2) $\overline{P}(X)_\alpha \subseteq \overline{P}(X)_\beta$.

Proof

(1) From Definition 3.1, on can have that for any $x \in \underline{P}(X)_\beta \Rightarrow \frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \beta \geq \alpha \Rightarrow x \in \underline{P}(X)_\alpha$.
(2) From the duality in Proposition 3.2 and item (1), this item can be proved easily.

□

Proposition 3.4 *Let* $I = (U, A, V, f)$ *be an information system,* $X \subseteq U$ *and* $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A(i = 1, 2, \ldots, l)$. *Additionally,* $Q = \{Q_i | Q_i \subseteq P_i, i = 1, 2, \ldots, l\}$ *is a set for which some attributes have been removed from the corresponding granulations. For any* $\beta \in (0.5, 1]$, *the following properties hold*:

(1) $\underline{Q}(X)_\beta \subseteq \underline{P}(X)_\beta$;
(2) $\overline{P}(X)_\beta \subseteq \overline{Q}(X)_\beta$.

Proof As $Q = \{Q_i | Q_i \subseteq P_i, i = 1, 2, \ldots, l\}$, so Q_i may be the empty set. If $Q_i = \emptyset$, then we denote $[x]_\emptyset = U$. This finding is logical and reasonable because all of the objects are indistinguishable with no attributes being considered.

(1) From the above assumptions, one can obtain that $[x]_{P_i} \subseteq [x]_{Q_i}$ holds for any $x \in U$ and $i \leq l$ is obvious. Furthermore, $S_X^{P_i}(x) \leq S_X^{Q_i}(x)$. Thus, $S_X^{P_i}(x) = 1 \Rightarrow S_X^{Q_i}(x) = 1$. Moreover, for any $x \in U$, we can obtain that

$$\begin{aligned} x \in \underline{Q}(X)_\beta &\Leftrightarrow \frac{\sum_{i=1}^{l} S_X^{Q_i}(x)}{l} \geq \beta \\ &\Rightarrow \frac{\sum_{i=1}^{l} S_X^{P_i}(x)}{l} \geq \frac{\sum_{i=1}^{l} S_X^{Q_i}(x)}{l} \geq \beta \\ &\Rightarrow x \in \underline{P}(X)_\beta. \end{aligned}$$

(2) This item can be proved using the duality. For any $\beta \in (0.5, 1]$, $\overline{P}(X)_\beta = \sim \underline{P}(X)_\beta \subseteq \sim \underline{Q}(X)_\beta = \overline{Q}(X)_\beta$.

□

Next, we will discuss some measures in generalized multigranulation rough sets to further study the new model.

Definition 3.2 Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A(i = 1, 2, \ldots, l)$. The accuracy and roughness of X are defined as

$$\alpha(X)_\beta^P = \frac{\underline{P}(X)_\beta}{\overline{P}(X)_\beta},$$

$$\rho(X)_\beta^P = 1 - \frac{\underline{P}(X)_\beta}{\overline{P}(X)_\beta}.$$

The relationships of accuracies and roughness among generalized, optimistic and pessimistic multigranulation rough sets are described as follows:

(1) $\alpha(X)_{\mathrm{PM}} \leq \alpha(X)_\beta^P \leq \alpha(X)_{\mathrm{OM}}$;
(2) $\rho(X)_{\mathrm{PM}} \geq \rho(X)_\beta^P \geq \rho(X)_{\mathrm{OM}}$.

Where $\alpha(X)_{\mathrm{OM}} = \frac{\underline{P}(X)_{\mathrm{OM}}}{\overline{P}(X)_{\mathrm{OM}}}$, $\alpha(X)_{\mathrm{PM}} = \frac{\underline{P}(X)_{\mathrm{PM}}}{\overline{P}(X)_{\mathrm{PM}}}$, $\rho(X)_{\mathrm{OM}} = 1 - \frac{\underline{P}(X)_{\mathrm{OM}}}{\overline{P}(X)_{\mathrm{OM}}}$, $\rho(X)_{\mathrm{PM}} = 1 - \frac{\underline{P}(X)_{\mathrm{PM}}}{\overline{P}(X)_{\mathrm{PM}}}$.

Proposition 3.5 *Let* $I = (U, A, V, f)$ *be an information system,* $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A(i = 1, 2, \ldots, l)$. *For any* $X, Y \subseteq U$ *and* $\beta \in (0.5, 1]$, *the following properties hold*:

(1) *X is more accurate than Y with respect to P under* β, *if and only if* $\alpha(X)_\beta^P \geq \alpha(Y)_\beta^P$;
(2) *X is more rough than Y with respect to P under* β, *if and only if* $\rho(X)_\beta^P \geq \rho(Y)_\beta^P$.

For different information levels $\alpha, \beta \in (0.5, 1]$. If $\alpha \leq \beta$, $\alpha(X)_\alpha^P \geq \alpha(X)_\beta^P$ and $\rho(X)_\alpha^P \geq \rho(X)_\beta^P$ hold.

In Pawlak rough sets model, a parameter called the dependent degree is used to illustrate the importance of a condition attribute subset with respect to the decision attributes in a target information system. This parameter can be defined similarly in generalized multigranulation rough set, as follows:

Let $I = (U, A, V, f)$ be an information system, $X \subseteq U$ and $P = \{P_1, P_2, \ldots, P_l\}$, $P_i \subseteq A(i = 1, 2, \ldots, l)$, $U/d = \{D_1, D_2, \ldots, D_r\}$. The dependent degree of P with respect to d under the information level β is defined by

$$\gamma(P, d)_\beta = \frac{\left|\bigcup_{k=1}^{r} \underline{P}(D_k)_\beta\right|}{|U|} = \frac{\sum_{k=1}^{r} \left|\underline{P}(D_k)_\beta\right|}{|U|}.$$

From Definition 3.3, the dependent degrees in pessimistic multigranulation rough set, generalized multigranulation rough set and optimistic multigranulation rough set have the following relations.

$$\gamma(P,d)_{\mathrm{PM}} \le \gamma(P,d)_\beta \le \gamma(P,d)_{\mathrm{OM}},$$

where

$$\gamma(P,d)_{\mathrm{PM}} = \frac{\sum_{k=1}^{r} |\underline{P}(D_k)_{\mathrm{PM}}|}{|U|}, \gamma(P,d)_{\mathrm{OM}} = \frac{\sum_{k=1}^{r} |\underline{P}(D_k)_{\mathrm{OM}}|}{|U|}.$$

4 Optimal granulation selection in generalized multigranulation rough sets

It is important to select the optimal granulation in detail corresponding a suitable information system. In this section, we investigate optimal granulation selection with different requirements for the generalized multigranulation rough sets.

Let $I = (U, A \cup \{d\}, V, f)$ be a target information system, $P = \{P_1, P_2, \dots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \dots, l)$. We say that I is granular consistent if for any $P_i \in P$, $[x]_{P_i} \subseteq [x]_d$ holds. Otherwise, I is a granular inconsistent information system.

For a granular consistent information system $I = (U, A \cup \{d\}, V, f)$ and $Q \subseteq P$, if the dependent degree of Q with respect to d under the level β: $\gamma(Q,d)_\beta = \gamma(P,d)_\beta$, we call Q is a granularity selection. $\forall R \in Q$, if the dependent degree of R with respect to d under the level β: $\gamma(Q - \{R\}, d)_\beta \ne \gamma(P,d)_\beta$, then Q is an optimal granulation selection of I.

Let $I = (U, A \cup \{d\}, V, f)$ be a granular inconsistent information system, and $P = \{P_1, P_2, \dots, P_l\}$, $P_i \subseteq A$ $(i = 1, 2, \dots, l)$. The lower and upper approximation granular distribution functions of I are denoted as follows:

$$f(\underline{P})_\beta = (\underline{P}(D_1)_\beta, \underline{P}(D_2)_\beta, \dots, \underline{P}(D_r)_\beta),$$
$$f(\overline{P})_\beta = (\overline{P}(D_1)_\beta, \overline{P}(D_2)_\beta, \dots, \overline{P}(D_r)_\beta).$$

The lower approximation granular distribution function figures all of the certain knowledge representations, and the upper approximation granular distribution function shows all of the possible knowledge representations in the sense of multigranulation. Moreover, optimal granularity selection can be acquired by considering these representations in terms of multigranulation.

Definition 4.1 Let $I = (U, A \cup \{d\}, V, f)$ be a granular inconsistent information system, and $P = \{P_1, P_2, \dots, P_l\}$. $P_i \subseteq A(i = 1, 2, \dots, l)$, $Q = \{P_i | i \le l\}$, and $R = \{P_i | i < l\}$.

(1) If $f(\underline{P})_\beta = f(\underline{Q})_\beta$, we say that Q is a lower distribution granulation selection of I. Moreover, if Q is a lower distribution granulation selection and no proper R of Q is a lower distribution granulation selection, then Q is called a lower distribution optimal granulation selection of I.

(2) If $f(\overline{P})_\beta = f(\overline{Q})_\beta$, we say that Q is a upper distribution granulation selection of I. Moreover, if Q is a upper distribution granulation selection and no proper R of Q is a upper distribution granulation selection, then Q is called a upper distribution optimal granulation selection of I.

Let $I = (U, A \cup \{d\}, V, f)$ be a granular inconsistent information system, and $P = \{P_1, P_2, \dots, P_l\}$, $P_i \subseteq A(i = 1, 2, \dots, l)$. The lower and upper approximation granular quality functions of I are denoted as follows:

$$\sigma_\beta^P = \frac{\sum_{k=1}^{r} |\underline{P}(D_k)_\beta|}{|U|},$$
$$\lambda_\beta^P = \frac{\sum_{k=1}^{r} |\overline{P}(D_k)_\beta|}{|U|}.$$

The lower approximation granular quality function lays out the number of objects in all of the certain knowledge representations, and the upper approximation granular quality function delivers the number of objects in all of the possible knowledge representations in the sense of multigranulation.

Definition 4.2 Let $I = (U, A \cup \{d\}, V, f)$ be a granular inconsistent information system, and $P = \{P_1, P_2, \dots, P_l\}$. $P_i \subseteq A(i = 1, 2, \dots, l)$, $Q = \{P_i | i \le l\}$, and $R = \{P_i | i < l\}$.

(1) If $\sigma_\beta^P = \sigma_\beta^Q$, then we say that Q is a lower quality granulation selection of I. Moreover, if Q is a lower quality granular and no proper R of Q is a lower quality granulation selection, then Q is referred to as a lower quality optimal granulation selection of I.

(2) If $\lambda_\beta^P = \lambda_\beta^Q$, we say that Q is an upper quality granulation selection of I. Moreover, if Q is an upper quality granulation selection and no proper R of Q is an upper quality granulation selection, then Q is referred to as an upper quality optimal granulation selection of I.

The optimal granulation selections defined in Definitions 4.1 and 4.2 are entirely the same when the considered target information system is granular consistent. Additionally, the method of computing the significance of every granulation and every condition attribute in significant granulations is the same as considering $\sigma_\beta^P = \sigma_\beta^Q$. Thus, this method can also be used in granular inconsistent information system, and the results obtained are lower approximation quality granulation selection.

Proposition 4.1 *Let* $I = (U, A \cup \{d\}, V, f)$ *be a granular inconsistent information system, and* $P = \{P_1, P_2, \dots, P_l\}$, $P_i \subseteq A (i = 1, 2, \dots, l)$.

1. *The lower approximation distribution optimal granulation selection is equivalent to the lower quality optimal granulation selection in generalized multigranulation rough set.*
2. *The upper approximation distribution optimal granulation selection is equivalent to the upper quality optimal granulation selection in generalized multigranulation rough set.*

Proof Assume that $Q \subseteq P$.

(1) $\Rightarrow f(\underline{Q})_\beta = f(\underline{P})_\beta \Rightarrow \sigma_\beta^Q = \sigma_\beta^P$ is obvious.

$\Leftarrow$ For any $D_i \in U/d$, we can have that $\underline{Q}(D_i)_\beta \subseteq \underline{P}(D_i)_\beta$. Then we can get $\sigma_\beta^Q \le \sigma_\beta^P$. If $\sigma_\beta^Q = \sigma_\beta^P$, for any $D_i \in U/d$, we can have $\underline{Q}(D_i)_\beta = \underline{P}(D_i)_\beta$. Otherwise, if there exists $D_{i_0} \in U/d$ such that $\underline{Q}(D_{i_0})_\beta \subsetneq \underline{P}(D_{i_0})_\beta$, then $\sigma_\beta^Q < \sigma_\beta^P$.

Therefore, $\sigma_\beta^Q = \sigma_\beta^P \Rightarrow f(\underline{Q})_\beta = f(\underline{P})_\beta$.

(2) $\Rightarrow f(\overline{Q})_\beta = f(\overline{P})_\beta \Rightarrow \lambda_\beta^Q = \lambda_\beta^P$ is obvious.

$\Leftarrow$ For any $D_i \in U/d$, we can have that $\overline{P}(D_i)_\beta \subseteq \overline{Q}(D_i)_\beta$. Then we can get $\lambda_\beta^Q \le \lambda_\beta^P$. If there exists $D_{i_0} \in U/d$ such that $\overline{P}(D_{i_0})_\beta \subsetneq \underline{P}(D_{i_0})_\beta$, $\lambda_\beta^P < \lambda_\beta^Q$.

Therefore, $\lambda_\beta^Q = \lambda_\beta^P \Rightarrow f(\overline{Q})_\beta = f(\overline{P})_\beta$. □

When $\beta = 1$, generalized multigranulation rough set can be degenerated into pessimistic multigranulation rough set. That is to say, pessimistic multigranulation rough set is a special case of generalized multigranulation rough set while $\beta = 1$. Thus, the above proposition holds for pessimistic multigranulation rough set. From the proof of the proposition, we can easily obtain that this proposition also holds for optimistic multigranulation rough set. However, it would not hold for other rough set models such as variable precision being considered in the sense of multigranulation. The lower and upper quality consistent reductions provide an easy and quick way to check reductions in computing by programs on computers.

From the above, we can know that the higher the information level β is, the stricter our requirements are. The reason is that when $\beta = 1$, generalized multigranulation rough set degenerated into pessimistic multigranulation rough set. Optimistic multigranulation rough set needs only at least one granule that supports the concept, while pessimistic multigranulation rough set needs all granule to support the concept.

Information level β must be subordinated to the majority principle, so β need to satisfy $\beta \in [0.5, 1]$. The bigger the β the finer the information granular, the results will be more accurate. The smaller the β the thicker the information granular, the results will be more macroscopic. Therefore, β can be artificially adjusted according to our needs. For a granular consistent information system, from the perspective of dependence, optimal granularity selection level: to find the smallest subset Q of P and Q meets $\gamma(Q, d)_\beta = \gamma(P, d)_\beta$. Thus, Q is the optimal granularity of the information system. The dependent degree of P with regard to d can be observed in an overall and systematic way. Furthermore, it also improves the computational efficiency. For a granular inconsistent information system, from the perspective of the lower distribution functions, the criterion for optimal granularity selection is to find the smallest subset Q of P and make Q satisfy $f(\underline{P})_\beta = f(\underline{Q})_\beta$. So Q becomes the optimal granularity of the information system. Each decision class lower approximate distribution can be observed under the attribute set P and information level β. The upper approximation distribution is similarly analyzed. From the perspective of lower approximation granular quality functions, the level of optimal granularity selection is to find the smallest subset Q of P and such that $\sigma_\beta^P = \sigma_\beta^Q$. Subsequently, Q is the optimal granularity of the information system. The proportion of lower approximation of decision classes in the total objects can be considered from the macroscopic point of view. This approach can effectively promote the performance of computing works. For the upper approximation granular quality function, the situation is analogous to the lower approximation.

5 Algorithm

According to the above theory, one can select the optimal granulation in generalized multigranulation rough sets. In this section, we present the algorithm of optimal granulation selection. And a real-life case study is given to show effectiveness of the proposed approach.

We outline the optimal granulation selection process of the lower and upper distribution in Algorithm 1. And the optimal granulation selection process of lower and upper quality are shown in Algorithm 2.

Algorithm 1: Lower and upper distribution optimal granularity selection in generalized multigranulation rough sets

```
Input  : I = (U, A ⋃{d}, V, f), the information level β, the number
         of objects N;
Output : (1) Lower distribution optimal granularity selection; (2)
         Upper distribution optimal granularity selection.
begin
    for j = 1 : N do
        for each x_j, U/d = (D_1, D_2, ···, D_r) do
            for i = 1 : l do
                for each P_i, U/P_i = {[x_j]_{P_i} | x_j ∈ U} do
                    if [x_j]_{P_i} ⊆ [x_j]_d then
                        End
                    else
                        for k = 1 : r do
                            if [x_j]_{P_i} ⊆ D_k then
                                S^{P_i}_{D_k}(x_j) = 1
                            else
                                S^{P_i}_{D_k}(x_j) = 0
                            end
                            if [x_j]_{P_i} ⊆∼ D_k then
                                S^{P_i}_{∼D_k}(x_j) = 1
                            else
                                S^{P_i}_{∼D_k}(x_j) = 0
                            end
                        end
                    end
                end
            end
        end
    end
    compute: Σ_{i=1}^{l} S^{P_i}_{D_k}(x_j) (j = 1, 2, ···, N; k = 1, 2, ···, r),
             Σ_{i=1}^{l} (1 − S^{P_i}_{∼D_k}(x_j)) (j = 1, 2, ···, N; k = 1, 2, ···, r).
    for j = 1 : N do
        for k = 1 : r do
            P̲(D_k)_β = {x_j ∈ U | (Σ_{i=1}^{l} S^{P_i}_{D_k}(x_j)) / l ≥
            β} and P̄(D_k)_β = {x_j ∈ U | (Σ_{i=1}^{l} (1−S^{P_i}_{∼D_k}(x_j))) / l >
            1 − β}.
        end
    end
    compute: f(P̲)_β = (P̲(D_1)_β, P̲(D_2)_β, ···, P̲(D_r)_β),
             f(P̄)_β = (P̄(D_1)_β, P̄(D_2)_β, ···, P̄(D_r)_β).
    for granular set Q_1 ⊂ P do
        if f(Q̲_1)_β ≠ f(P̲)_β then
            P is lower distribution optimal granularity selection;
        else
            ∀R_1 ⊂ Q_1 if f(R̲_1)_β ≠ f(Q̲_1)_β then
                Q_1 is lower distribution optimal granularity
                selection;
            else
                Q_1 is not lower distribution optimal granularity
                selection;
            end
        end
    end
    for granular set Q_2 ⊂ P do
        if f(Q̄_2)_β ≠ f(P̄)_β then
            P is upper distribution optimal granularity selection;
        else
            ∀R_2 ⊂ Q_2
            if f(R̄_2)_β ≠ f(Q̄_2)_β then
                Q_2 is upper distribution optimal granularity
                selection;
            else
                Q_2 is not upper distribution optimal granularity
                selection.
            end
        end
    end
end
```

Algorithm 2: Lower and upper quality optimal granularity selection in generalized multigranulation rough sets

```
Input  : I = (U, A ⋃{d}, V, f), the information level β, the number
         of objects N;
Output : (1) Lower quality optimal granularity selection; (2) Upper
         quality optimal granularity selection.
begin
    for j = 1 : N do
        for each x_j, U/d = (D_1, D_2, ···, D_r). do
            for i = 1 : l do
                for each P_i, U/P_i = {[x_j]_{P_i} | x_j ∈ U} do
                    for k = 1 : r do
                        if [x_j]_{P_i} ⊆ D_k then
                            S^{P_i}_{D_k}(x_j) = 1
                        else
                            S^{P_i}_{D_k}(x_j) = 0.
                        end
                        if [x_j]_{P_i} ⊆∼ D_k then
                            S^{P_i}_{∼D_k}(x_j) = 1
                        else
                            S^{P_i}_{∼D_k}(x_j) = 0.
                        end
                    end
                end
            end
        end
    end
    compute: Σ_{i=1}^{l} S^{P_i}_{D_k}(x_j) (j = 1, 2, ···, N; k = 1, 2, ···, r),
             Σ_{i=1}^{l} (1 − S^{P_i}_{∼D_k}(x_j)) (j = 1, 2, ···, N; k = 1, 2, ···, r).
    for j = 1 : N do
        for k = 1 : r do
            P̲(D_k)_β = {x_j ∈ U | (Σ_{i=1}^{l} S^{P_i}_{D_k}(x_j)) / l ≥
            β} and P̄(D_k)_β = {x_j ∈ U | (Σ_{i=1}^{l} (1−S^{P_i}_{∼D_k}(x_j))) / l >
            1 − β}.
        end
    end
    compute: σ^P_β = (Σ_{k=1}^{l} |P̲(D_k)_β|) / |U|, λ^P_β = (Σ_{k=1}^{l} |P̄(D_k)_β|) / |U|.
    for granular set Q_1 ⊂ P do
        if f(Q̲_1)_β ≠ f(P̲)_β then
            P is (lower quality) optimal granularity selection;
        else
            ∀R_1 ⊂ Q_1 if f(R̲_1)_β ≠ f(Q̲_1)_β then
                Q_1 is (lower quality) optimal granularity selection;
            else
                Q_1 is not (lower quality) optimal granularity
                selection;
            end
        end
    end
    for i = 1 : l do
        if [x_j]_{P_i} ⊆ [x_j]_d then
            end
        else
            for granular set Q_2 ⊂ P do
                if f(Q̄_2)_β ≠ f(P̄)_β then
                    P is upper quality optimal granularity selection;
                else
                    ∀R_2 ⊂ Q_2
                    if f(R̄_2)_β ≠ f(Q̄_2)_β then
                        Q_2 is upper quality optimal granularity
                        selection;
                    else
                        Q_2 is not upper quality optimal
                        granularity selection.
                    end
                end
            end
        end
    end
end
```

In Algorithm 1, it should be noted that Q_1 and Q_2 take all of the subsets of P.

Time complexity analysis of Algorithm 1 Let $I=(U,A\cup\{d\},V,f))$ be a target information system. The number of objects and attributes are denoted by N and $|A|$. The number of objects and attributes in the i-th granule P_i are denoted by N_i and K_i $(i\in\{1,2,\ldots,l\}$, $\Sigma N_i=N$ and $\Sigma K_i=|A|)$, respectively. $U/d=\{D_1,D_2,\ldots,D_r\}$ is the decision classes. We take a variable t_i to stand for the time complexity in an implementation. In the next, we can analyze the time complexity of Algorithm 1 step by step.

The time complexity to do initialized setting and input the information table, the information level, and the granule set is 0, then the analysis to do the initial settings is finished. The time complexity to calculate U/d is denoted by $t_1=(N-1)+(N-2)+\cdots+1=N\times(N-1)/2$, the time complexity to calculate $U/P_i(i=1,2,\ldots,l)$ is $t_2=N\times|A|/2$. So the time is $t_3=t_1+t_2=N\times(N-1)/2+N\times|A|/2$. The next to judge whether the information system is granular consistent is $t_4=N\times l$. The time to obtain lower and upper approximations in generalized multigranulation rough sets is $t_5=2\times N\times l\times r+2\times N$.

The first four steps calculate the lower and upper approximation granular distribution functions of I, lower and upper distribution optimal granulation selection. The time complexity of methods are denoted by $t_6=2^{l-1}\times(N/2+N\times|A|/2+2\times N\times l\times r+2\times N)\times 2$ $=2^{l-1}\times(N\times|A|+4\times N\times l\times r+5N/2)$.

From the above analysis, we can know that the maximum time complexity of the main part in the Algorithm 1 is

$$\begin{aligned}t^1_{\text{main}}&=t_3+t_4+t_5+t_6\\&=N^2/2+(2^{l+1}+2)\times N\times l\times r\\&\quad+(2^l\times 5/4+3/2)\times N+(2^{l-1}+1/2)\times N\times|A|\\&\quad+N\times l\end{aligned}$$

As $r\,(1\le r\le N)$ is the number of decision classes, and $l\,(1\le l\le|A|)$ is the number of granular structure, the maximum complexity of the main algorithm is approximately $O((2^{|A|+1}+2)\times N^2\times|A|)$.

The explanation of the Fig. 2: In Algorithm 1, the lower distribution optimal granularity selection and upper distribution optimal granularity selection are accordance with this flow graph. Input a target information system (an information table $I=(U,A\cup\{d\},V,f))$, an information level β, and the granule set $P=\{P_1,P_2,\ldots,P_l\}$, $P_i\subseteq A(i=1,2,\ldots,l)$. We first calculate U/d and U/P_i to make the judgement of whether this information system is granular consistent or not. If it is granular inconsistent, we cannot do the lower and upper distribution optimal

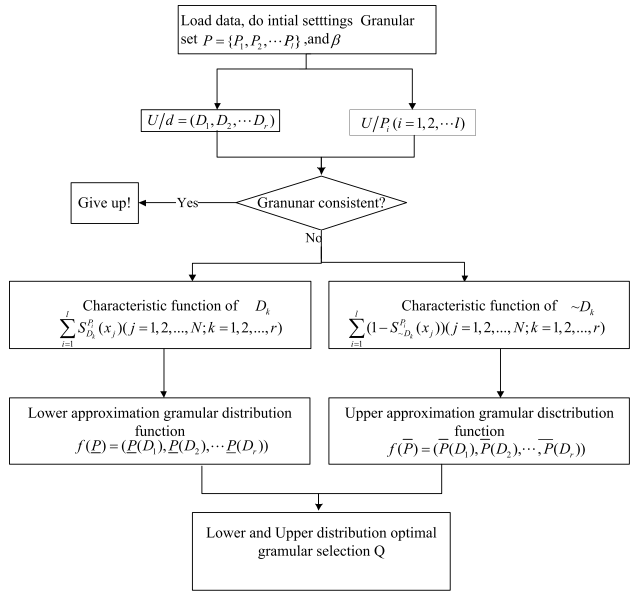

Fig. 2 The program flow graph of lower and upper distribution optimal granulation selection

granularity selection. If it is granular inconsistent, we then compute the lower and upper approximations of each decision class and its complement in generalized multigranulation rough set. After that, we get the lower and upper approximation granular distribution functions of I which are $f(\underline{P})_\beta$ and $f(\overline{P})_\beta$, respectively. According to the optimal granularity selection of lower distribution and upper distribution in Definition 4.1, we can obtain the lower and upper distribution granularity selections, respectively. Among the obtained lower and upper distribution granularity selections, we can further get the lower distribution optimal granularity selection and upper distribution optimal granularity selection.

In Algorithm 2, it should be noted that Q_1 and Q_2 take all of the subsets of P.

Time complexity analysis of Algorithm 2 Let $I=(U,A\cup\{d\},V,f))$ be a target information system. The number of objects and attributes are denoted by N and $|A|$. The number of objects and attributes in the i-th granule P_i are denoted by N_i and K_i $(i\in\{1,2,\ldots,l\}$, $\Sigma N_i=N$ and $\Sigma K_i=|A|)$, respectively. $U/d=\{D_1,D_2,\ldots,D_r\}$ is the decision classes. We take a variable t_i to stand for the time complexity in an implementation. Next, we can analyze the time complexity of Algorithm 2. The time complexity to do initialized setting and input the information table, the information level, and the granule set is 0, then the analysis to do initial settings is finished. The time complexity to calculate $U\,/\,d$ is denoted by $t_1=(N-1)+(N-2)+\cdots+1=N\times(N-1)/2$, the time complexity to calculate $U/P_i(i=1,2,\ldots,l)$ is $t_2=N\times|A|/2$. So the time to finish them is

$t_3 = t_1 + t_2 = N \times (N-1)/2 + N \times |A|/2$. The time to obtain lower and upper approximations in generalized multigranulation rough sets is $t_4 = 2 \times N \times l \times r + 2 \times N$.

The first three steps calculate the lower and upper approximation granular distribution functions of I, lower and upper distribution optimal granularity selection. To judge whether the information system is granular consistent is $t_5 = N \times l$. The time complexity of methods are denoted by $t_6 = 2^{l-1} \times (N \times |A|/2 + 2 \times N \times l \times r + 5N/2) \times 3/2$.

From the above analysis, we can know that the maximum time complexity of the main part in Algorithm 2 is

$$\begin{aligned} t^2_{\text{main}} &= t_3 + t_4 + t_5 + t_6 \\ &= (3 \times 2^{l-1} + 2) \times N \times l \times r + (15/4 \times 2^{l-1} + 2) \\ &\quad \times N + (3/8 \times 2^{l-1} + 1/2) \times N \times |A| \\ &\quad + 1/2 \times N \times (N-1) + N \times l \end{aligned}$$

The maximum complexity of the main algorithm is approximately $O((3 \times 2^{|A|-1} + 2) \times N^2 \times |A|)$.

The explanation of Fig. 3: In Algorithm 2, the lower quality optimal granularity selection and upper quality optimal granularity selection are in accordance with this flow graph. Unlike Algorithm 1, we can get the optimal granularity selection when it is granular consistent, so is it inconsistent, we can get the lower quality optimal granularity selection. After calculating U/d and U/P_i $(i = 1, 2, \ldots, l)$, we obtain the characteristic functions of D_k and $\sim D_k$. Then we further get the lower and upper

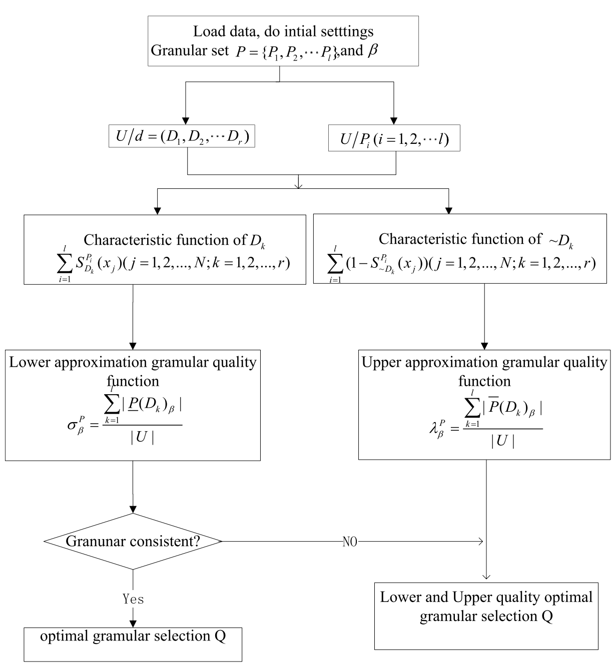

Fig. 3 The program flow graph of lower and upper quality optimal granularity selection

approximation granular quality functions, respectively. After getting the lower approximation granular quality function, we need to make the judgement of whether this information system is granular consistent. If it is granular consistent, we can do the optimal granularity selection. If it is granular inconsistent, we then compute the lower and upper approximations of each decision class and its complement in generalized multigranulation rough set. After that, we get the lower and upper approximation granular distribution functions of I which are $\sigma^P_\beta = \frac{\sum_{k=1}^{r} |\underline{P}(D_k)_\beta|}{|U|}$ and $\lambda^P_\beta = \frac{\sum_{k=1}^{r} |\overline{P}(D_k)_\beta|}{|U|}$, respectively. According to Definition 4.2, we can obtain the lower and upper quality granularity selections, respectively. Among the obtained lower and upper quality granularity selections, we can further get the lower quality optimal granularity selection and upper quality optimal granularity selection. Multigranulation rough set can be very useful in many case, especially in handing problem in information system.

6 Case study

Suppose that Table 1 is an information system $I = (U, C \cup \{d\}, V, f)$ which concerns the achievements of some students and $U = \{x_1, x_2, \ldots, x_{20}\}$ is a universe including twenty students in a school, a_1(*Chinese*), a_2(*Mathematics*), a_3(*English*), a_4(*History*), a_5(*Geography*), a_6(*Politics*), a_7(*Physics*), a_8(*Chemistry*), a_9(*Biology*) are the conditional attributes of the system, and d(*Decision*) is the decision attribute given by the experts according to the achievements of these students. We use "1" to express that the student is excellent and "0" to express that the student is not excellent.

However, in the college entrance examination, junior colleges only select excellent students from the view of three major subjects (*Chinese*, *Mathematics*, *English*). Universities of undergraduate level select excellent students from the view of arts (*History*, *Geography*, *Politics*) and from the view of science (*Physics*, *Chemistry*, *Biology*).

From the point of view of arts, selecting the attribute set $\{a_4, a_5, a_6\}$ is better than the attribute set $\{a_7, a_8, a_9\}$. And from the point of view of science, selecting the attribute set $\{a_7, a_8, a_9\}$ is better than the attribute set $\{a_4, a_5, a_6\}$. So we can get the following three granulations.

$$P_1 = \{a_1, a_2, a_3\},$$
$$P_2 = \{a_4, a_5, a_6\},$$
$$P_3 = \{a_7, a_8, a_9\}.$$

If we consider only one of these conditions, we can obtain that

Table 1 A target information system

U	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	d
x_1	1	0	2	1	1	2	3	1	2	0
x_2	2	1	4	2	4	0	1	3	0	1
x_3	0	2	4	3	1	2	3	0	2	1
x_4	2	0	3	3	1	2	2	3	0	1
x_5	0	2	4	4	0	0	3	0	2	1
x_6	1	0	2	0	2	3	3	0	2	1
x_7	1	1	1	2	4	0	4	4	3	1
x_8	0	3	4	1	1	2	4	4	3	0
x_9	0	3	4	3	3	4	2	3	0	0
x_{10}	1	1	1	0	2	3	3	0	1	0
x_{11}	3	2	2	2	4	0	0	2	2	1
x_{12}	2	1	4	2	3	4	1	3	0	1
x_{13}	3	2	2	2	3	4	0	2	2	1
x_{14}	0	3	4	4	0	0	3	0	1	0
x_{15}	1	0	2	3	3	4	2	3	0	0
x_{16}	1	1	1	1	1	2	3	4	4	0
x_{17}	3	2	2	2	3	4	3	4	4	1
x_{18}	2	0	3	2	3	4	2	3	0	1
x_{19}	0	3	4	1	1	2	4	4	3	0
x_{20}	1	0	2	3	3	4	1	3	0	1

$$
\begin{aligned}
U/P_1 = \{&\{x_1,x_6,x_{15},x_{20}\},\{x_2,x_{12}\},\{x_3,x_5\},\\
&\{x_8,x_9,x_{14},x_{19}\},\{x_4,x_{18}\},\{x_7,x_{10},x_{16}\},\\
&\{x_{11},x_{13},x_{17}\}\};\\
U/P_2 = \{&\{x_1,x_8,x_{16},x_{19}\},\{x_2,x_7,x_{11}\},\{x_3,x_4\},\\
&\{x_5,x_{14}\},\{x_6,x_{10}\},\{x_{12},x_{13},x_{17},x_{18}\},\\
&\{x_9,x_{15},x_{20}\}\};\\
U/P_3 = \{&\{x_1\},\{x_3,x_5,x_6\},\{x_{10},x_{14}\},\{x_2,x_{12},x_{20}\},\\
&\{x_4,x_9,x_{15},x_{18}\},\{x_7,x_8,x_{19}\},\{x_{11},x_{13}\},\{x_{16},x_{17}\}\};\\
U/d = \{D_1,D_2\} = \{&\{x_1,x_8,x_9,x_{10},x_{14},x_{15},x_{16},x_{19}\},\\
&\{x_2,x_3,x_4,x_5,x_6,x_7,x_{11},x_{12},x_{13},x_{17},x_{18},x_{20}\}\}.
\end{aligned}
$$

First, according to existence $[x]_P \not\subset [x]_d$,it is easy to check that Table 1 is a granular inconsistent information system.

Next, on the basis of generalized multi granularity, how to choose the best particle size of the two algorithms is explained.

The nonempty subsets of P are denoted by

$$P = \{P_1,P_2,P_3\}, Q_1 = \{P_1,P_2\}, Q_2 = \{P_1,P_3\},$$
$$Q_3 = \{P_2,P_3\}, Q_4 = \{P_1\}, Q_5 = \{P_2\}, Q_6 = \{P_3\}.$$

We set the information level $\beta = 2/3$.

The calculation steps of the algorithm 1 are as follows:

Step 1. From Definition 3.1, we can calculate the generalized lower and upper approximations of the decision class D_1 as follows:

$$\underline{P}(D_1)_{2/3} = \{x_1,x_8,x_{14},x_{19}\};$$
$$\overline{P}(D_1)_{2/3} = \{x_1,x_8,x_9,x_{10},x_{14},x_{15},x_{16},x_{19}\}.$$

The generalized lower and upper approximations of the decision class D_2 are

$$
\begin{aligned}
\underline{P}(D_2)_{2/3} = \{&x_2,x_3,x_4,x_5,x_6,x_7,x_{11},x_{12},x_{13},x_{17},x_{18},\\
&x_{20}\};\\
\overline{P}(D_2)_{2/3} = \{&x_2,x_3,x_4,x_5,x_6,x_7,x_9,x_{10},x_{11},x_{12},x_{13},x_{15},\\
&x_{16},x_{17},x_{18},x_{20}\}.
\end{aligned}
$$

Step 2. According to the definition of distribution function , we can obtain the lower and upper approximation granular distribution functions as follows:

$$
\begin{aligned}
f(\underline{P})_{2/3} &= (\underline{P}(D_1)_{2/3},\underline{P}(D_2)_{2/3})\\
&= (\{x_1,x_8,x_{14},x_{19}\},\{x_2,x_3,x_4,x_5,x_6,x_7,x_{11},x_{12},\\
&\quad x_{13},x_{17},x_{18},x_{20}\}).\\
f(\bar{P})_{2/3} &= (\bar{P}(D_1)_{2/3},\bar{P}(D_2)_{2/3})\\
&= (\{x_1,x_8,x_9,x_{10},x_{14},x_{15},x_{16},x_{19}\},\{x_2,x_3,x_4,x_5,x_6,\\
&\quad x_7,x_9,x_{10},x_{11},x_{12},x_{13},x_{15},x_{16}, x_{17},x_{18},x_{20}\}).
\end{aligned}
$$

Step 3. For Q_1,Q_2,Q_3,Q_4,Q_5,Q_6, we can get the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

For $Q_1 = \{P_1, P_2\}$, we can get the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_1}(D_1)_{2/3} = \{x_8, x_{19}\}$;

$\overline{Q_1}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$\underline{Q_1}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}$;

$\overline{Q_1}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{10}, x_{11}, x_{12}, x_{13}, x_{15}, x_{17}, x_{18}, x_{20}\}$.

For $Q_2 = \{P_1, P_3\}$, we can calculate the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_2}(D_1)_{2/3} = \{x_{14}\}$;

$\overline{Q_2}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$\underline{Q_2}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}$;

$\overline{Q_2}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{20}\}$.

For $Q_3 = \{P_2, P_3\}$, we can compute the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_3}(D_1)_{2/3} = \{x_1\}$;

$\overline{Q_3}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$\underline{Q_3}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}$;

$\overline{Q_3}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{11}, x_{12}, x_{13}, x_{15}, x_{17}, x_{18}, x_{20}\}$.

For $Q_4 = \{P_1\}$, we can get the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_4}(D_1)_{2/3} = \{x_8, x_9, x_{14}, x_{19}\}$;

$\overline{Q_4}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$\underline{Q_4}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}$;

$$\overline{Q_4}(D_2)_{2/3} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_{10}, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{20}\}.$$

For $Q_5 = \{P_2\}$, we can receive the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_5}(D_1)_{2/3} = \{x_1, x_8, x_{16}, x_{19}\}$;

$\overline{Q_5}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$$\underline{Q_5}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\};$$

$$\overline{Q_5}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{17}, x_{18}, x_{20}\}.$$

For $Q_6 = \{P_3\}$, we can obtain the generalized multigranulation lower and upper approximations of the decision classes D_1 and D_2:

$\underline{Q_6}(D_1)_{2/3} = \{x_1, x_{10}, x_{14}\}$;

$\overline{Q_6}(D_1)_{2/3} = \{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}$.

$\underline{Q_6}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}$;

$$\overline{Q_6}(D_2)_{2/3} = \{x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}\}.$$

Step 4. For $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6$, the lower approximation granular distribution functions are computed as follows:

$$\begin{aligned}
f(\underline{Q_1})_{2/3} &= (\underline{Q_1}(D_1)_{2/3}, \underline{Q_1}(D_2)_{2/3}) \\
&= (\{x_8, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}). \\
f(\underline{Q_2})_{2/3} &= (\underline{Q_2}(D_1)_{2/3}, \underline{Q_2}(D_2)_{2/3}) \\
&= (\{x_{14}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}). \\
f(\underline{Q_3})_{2/3} &= (\underline{Q_3}(D_1)_{2/3}, \underline{Q_3}(D_2)_{2/3}) \\
&= (\{x_1\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}). \\
f(\underline{Q_4})_{2/3} &= (\underline{Q_4}(D_1)_{2/3}, \underline{Q_4}(D_2)_{2/3}) \\
&= (\{x_8, x_9, x_{14}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}). \\
f(\underline{Q_5})_{2/3} &= (\underline{Q_5}(D_1)_{2/3}, \underline{Q_5}(D_2)_{2/3}) \\
&= (\{x_1, x_8, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}). \\
f(\underline{Q_6})_{2/3} &= (\underline{Q_6}(D_1)_{2/3}, \underline{Q_6}(D_2)_{2/3}) \\
&= (\{x_1, x_{10}, x_{14}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{17}, x_{18}, x_{20}\}).
\end{aligned}$$

Step 5. For $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6$, the upper approximation granular distribution functions are computed as follows:

$$\begin{aligned}
f(\overline{Q_1})_{2/3} &= (\overline{Q_1}(D_1)_{2/3}, \overline{Q_1}(D_2)_{2/3}) \\
&= (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{10}, x_{11}, x_{12}, x_{13}, x_{15}, x_{17}, x_{18}, x_{20}\}). \\
f(\overline{Q_2})_{2/3} &= (\overline{Q_2}(D_1)_{2/3}, \overline{Q_2}(D_2)_{2/3}) \\
&= (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{20}\}).
\end{aligned}$$

$$f(\overline{Q_3})_{2/3} = (\overline{Q_3}(D_1)_{2/3}, \overline{Q_3}(D_2)_{2/3}) = (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{11}, x_{12}, x_{13}, x_{15}, x_{17}, x_{18}, x_{20}\}).$$

$$f(\overline{Q_4})_{2/3} = (\overline{Q_4}(D_1)_{2/3}, \overline{Q_4}(D_2)_{2/3}) = (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_{10}, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{20}\}).$$

$$f(\overline{Q_5})_{2/3} = (\overline{Q_5}(D_1)_{2/3}, \overline{Q_5}(D_2)_{2/3}) = (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{17}, x_{18}, x_{20}\}).$$

$$f(\overline{Q_6})_{2/3} = (\overline{Q_6}(D_1)_{2/3}, \overline{Q_6}(D_2)_{2/3}) = (\{x_1, x_8, x_9, x_{10}, x_{14}, x_{15}, x_{16}, x_{19}\}, \{x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{11}, x_{12}, x_{13}, x_{15}, x_{16}, x_{17}, x_{18}, x_{19}, x_{20}\}).$$

Step 6. According to Definition 4.1, $P = \{P_1, P_2, P_3\}$ is the lower distribution optimal granularity selection of the system. And $P = \{P_1, P_2, P_3\}$ is also the upper distribution optimal granularity selection of I.

The calculation steps of the algorithm 2 are as follows:

Step 1. Same as the step 1 in algorithm 1

Step 2. According to the definition of quality function, the lower and upper approximation granular quality functions can be also computed as

$$\sigma^P_{2/3} = \frac{\sum_{k=1}^{2} |\underline{P}(D_k)_{2/3}|}{|U|} = \frac{16}{20}, \quad \lambda^P_{2/3} = \frac{\sum_{k=1}^{2} |\overline{P}(D_k)_{2/3}|}{|U|} = \frac{24}{20}.$$

Step 3. Same as the step 3 in algorithm 1

Step 4. For $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6$, The lower approximation granular quality functions can be also computed as:

$$\sigma^{Q_1}_{2/3} = \frac{14}{20}, \sigma^{Q_2}_{2/3} = \frac{13}{20}, \sigma^{Q_3}_{2/3} = \frac{13}{20},$$
$$\sigma^{Q_4}_{2/3} = \frac{16}{20}, \sigma^{Q_5}_{2/3} = \frac{16}{20}, \sigma^{Q_6}_{2/3} = \frac{15}{20}.$$

Table 2 Results of optimal granularity selection

Type	Optimal granularity selection	Number
Lower distribution	P	1
Upper distribution	P	1
Lower quality	Q_4 or Q_5	2
Upper quality	Q_4 or Q_5	2

Step 5. For $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6$, the upper approximation granular quality functions can be also computed as:

$$\lambda^{Q_1}_{2/3} = \frac{22}{20}, \lambda^{Q_2}_{2/3} = \frac{22}{20}, \lambda^{Q_3}_{2/3} = \frac{22}{20},$$
$$\lambda^{Q_4}_{2/3} = \frac{24}{20}, \lambda^{Q_5}_{2/3} = \frac{24}{20}, \lambda^{Q_6}_{2/3} = \frac{25}{20}.$$

Step 6. According to Definition 4.2, $Q_4 = \{P_1\}$ or $Q_5 = \{P_2\}$ are the lower quality optimal granularity selection of I. And $Q_4 = \{P_1\}$ or $Q_5 = \{P_2\}$ are also the upper quality optimal granularity selection of I

Finally, we get results as shown in the following Table 2.

7 Conclusions

The multigranulation rough set proposed by Qian et al. is an important development of Pawlak's rough set theory. By considering the strict optimistic and pessimistic conditions in classic multigranulation rough set model, the support characteristic function has been presented in our work. We also introduced the information level β to propose novel multigranulation rough set, called generalized multigranulation rough set model. The main contributions of this paper are as follows. First, the generalized multi granularity rough set model is discussed, and the corresponding measures and performance are discussed. Second, to deal with the appropriate size of the selection, we also study the four best particle size selection method, and design a selection algorithm. Finally, we have constructed a real life example to explain and illustrate the best particle size selection method. In the future, we will study other new particle size selection methods and the corresponding attribute reduction.

Acknowledgements The author would like to thank the valuable suggestions from the anonymous referees, and Professor Witold Pedrycz, the editor in chief, for improving the quality of the paper. This work is supported by Natural Science Foundation of China (No. 61105041, No. 61472463, No. 61402064), National Natural Science Foundation of CQ CSTC (No. cstc 2013jcyjA40051), Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information (Nanjing University of Science and Technology), Ministry of Education (No. 30920140122006), and Key Laboratory of Oceanographic Big Data Mining and Application of Zhejiang Province (No. OBDMA201503).

References

Apolloni B, Bassis S, Rota J, Galliani GL, Gioia M, Ferrari L (2016) A neurofuzzy algorithm for learning from complex granules. Granul Comput 1(4):225–246

Bonikowski Z, Bryniarski E, Wybraniec-Skardowska U (1998) Extensions and intensions in the rough set theory. Inf Sci 107:149–167

Dou H, Yang X, Fan J, Xu S (2012) The models of variable precision multigranulation rough sets. In: Lecture notes in computer science, vol 7414, No 8, pp 465–473

Feng T, Mi J (2015) Variable precision multigranulation decisiontheoretic fuzzy rough sets. Knowl Based Syst 91(4):93–101

Greco S, Matarazzo B, Słowiński R (1999) The use of rough sets and fuzzy sets in MCDM. In: Gal T, Stewart TJ, Hanne T (eds) Multicriteria decision making: advances in MCDM models, algorithms, theory, and applications. Kluwer Academic Publishers, Boston, MA, pp 14-1–14-59

Guo C, Zheng X (2014) Feature subset selection approach based on fuzzy rough set for high-dimensional data[C]. In: 2014 IEEE international conference on granular computing (GrC), vol 49, No 12, pp 72–75

Huang B, Guo C, Zhuang Y, Li H, Zhou X (2014) Intuitionistic fuzzy multigranulation rough sets. Inf Sci 277(2):299–320

Inuiguchi M, Tanino T (2001) On rough sets under generalized equivalence relations. Bull Int Rough Set Soc 5(1/2):167–171

Inuiguchi M, Tanino T (2002) Generalized rough sets and rule extraction. In: Rough sets and current trends in computing. Springer, Berlin, pp 105–112

Li J, Mei C, Xu W, Qian Y (2015) Concept learning via granular computing: a cognitive viewpoint. Inf Sci 298(15):447–467

Li J, Ren Y, Mei C, Qian Y, Yang X (2016) A comparative study of multigranulation rough sets and concept lattices via rule acquisition. Knowl Based Syst 91(1):152–164

Li W, Xu W (2015) Multigranulation decision-theoretic rough set inordered information system. Fundam Inf 139(1):1–23

Li W, Zhang X, Sun W (2014) Further study of multigranulation T-fuzzy rough sets. Sci World J 14(2):1–18

Liang J, Wang F, Dang C, Qian Y (2012) An efficient rough feature selection algorithm with a multigranulation view. Int J Approx Reason 53(6):912–926

Lin Y, Li J, Lin P, Lin G, Chen J (2014) Feature selection via neighborhood multi-granulation fusion. Knowl Based Syst 67:162–168

Lin G, Liang J, Qian Y (2013) Multigranulation rough sets: from partition to covering. Inf Sci 241:101–118

Lin G, Liang J, Qian Y (2015) An information fusion approach by combining multigranulation rough sets and evidence theory. Inf Sci 314(14):184–199

Lin G, Qian Y, Li J (2012) NMGRS: neighborhoodbased multigranulation rough sets. Int J Approx Reason 53(7):1080–1093

Liu H, Gegov A, Cocea M (2016) Rule-based systems: a granular computing perspective. Granul Comput 1(4):259–274

Liu C, Miao D, Qian J (2014) On multi-granulation covering rough sets. Int J Approx Reason 55(6):1404–1418

Lu J, Li D, Zhai Y, Li H, Bai H (2016) A model for type-2 fuzzy rough sets. Inf Sci 328:359–377

Pawlak Z (1982) Rough sets. Int J Comput Inf Sci 11:341–356

Pedrycz W, Bargiela A (2002) Granular computing: a granular signature of data. IEEE Trans Syst 32(2):212–224

Pedrycz W, Bargiela A (2012) An optimization of allocation of information granularity in the interpretation of data structures: toward granular fuzzy clustering. IEEE Trans Syst 42(3):582–590

Pedrycz W (2013) Granular computing analysis and design of intelligent systems.CRC Press Taylor & Francis Group, Boca Raton

Qian Y, Li S, Liang J, Shi Z, Wang F (2014) Pessimistic rough set based decisions: a multigranulation fusion strategy. Inf Sci 264:196–210

Qian Y, Liang J, Dang C (2009) Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. Int J Approx Reason 50:174–188

Qian Y, Liang J, Dang C (2009) Incomplete multigranulation rough set. IEEE Trans Syst 40(2):420–432

Qian Y, Liang J, Yao Y, Dang C (2010) A multigranulation rough set. Inf Sci 180:949–970

Qian Y, Zhang H, Sang Y, Liang J (2014) Multigranulation decision-theoretic rough sets. Int J Approx Reason 55(1):225–237

Qian Y, Liang J (2006) Rough set method based on multigranulations. In: IEEE conference on cognitive Informations vol 1, pp 297–304

Senthil Kumar S, Hannah Inbarani H (2015) Optimistic multi-granulation rough set based classification for medical diagnosis. Proced Comput Sci 47:374–382

She Y, He X (2012) On the structure of the multigranulation rough set model. Knowl Based Syst 36:81–92

Shen Q, Jensen R (2007) Rough sets, their extensions and applications. Int J Autom Comput 4:100–106

Tan A, Wu W, Li J, Lin G (2016) Evidence-theory-based numerical characterization of multigranulation rough sets in incomplete information systems. Fuzzy Sets Syst 294(7):18–35

Wang G, Li Y, Li X (2017a) Approximation performance of the nonlinear hybrid fuzzy system based on variable universe. Granul Comput 2:1–12

Wang G, Yang J, Xu J (2017b) Granular computing: from granularity optimization to multi-granularity joint problem solving. Granul Comput 2:1–16

Xu W, Sun W, Zhang X, Zhang W (2012) Multiple granula tion rough set approach to ordered information systems. Int J Gen Syst 41(5):475–501

Xu Z, Wang H (2016) Managing multi-granularity linguistic information in qualitative group decision making: an overview. Granul Comput 1(1):21–35

Xu W, Wang Q, Luo S (2014) Multi-granularion fuzzy rough sets. J Intell Fuzzy Syst 26(3):1323–1340

Xu W, Wang Q, Zhang X (2011) Multi-granulation fuzzy rough sets in a fuzzy tolerance approximation space. Int J Fuzzy Syst 3(4):246–259

Xu W, Wang Q, Zhang X (2013) Multi-granulation rough sets based on tolerance relations. Soft Comput 17:1241–1252

Xu W, Li W (2014) Granular computing approach to two way learning based on formal concept analysis in fuzzy datasets. IEEE Trans Cybern. doi:10.1109/TCYB.2014.2361772

Yang X, Qi Y, Song X, Yang J (2013) Test cost sensitive multi granulation rough set: model and minimal cost selection. Inf Sci 250:184–199

Yang X, Qi Y, Yu H, Song X, Yang J (2014) Updating multi granulation rough approximations with increasing of granular structures. Knowl Based Syst 64:59–69

Yang X, Qian Y, Yang J (2012) Hierarchical structures on multigranulation spaces. J Comput Sci Technol 27(6):1169–1183

Yang X, Song X, Chen Z, Yang J (2012) On multigranulation rough sets in incomplete information system. Int J Mach Learn Cybern 3:223–232

Yang X, Song X, Chen Z, Yang J (2012) Multigranulation rough sets in incomplete information system. In: Incomplete information system and rough set theory. Springer, Berlin, pp 195–222

Yang X, Yang J, Hu X (2009) A new rough set model for knowledge acquisition in incomplete information system. In: IEEE International conference on granular computing, vol 259, No 4, pp 696–701

Yao Y (1996) Two views of the theory of rough sets infinite universes. Int J Approx Reason 15:291–317

Yao Y (1998) Relational interpretations of neighborhood op erators and rough set approximation operators. Inf Sci 111:239–259

Yao Y (2000) Information granulation and rough set approximation. Int J Intell Syst 16:87–104

Yao Y, Deng X (2014) Quantitative rough sets based on sub sethood measures. Inf Sci 265(5):306–322

Yao Y, Lin T (1996) Generalization of rough sets using modal logics. Intell Autom Soft Comput 2(2):103–120

Yao Y, She Y (2016) Rough set models in multigranulation spaces. Inf Sci 327:40–56

Zadeh LA (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 19(2):111–127

Zhang X, Miao D, Liu C, Le M (2015) Constructive methods of rough approximation operators and multigranulation rough sets. Knowl Based Syst. doi:10.1016/j.knosys.09.036 (Available online 8 Oct 2015)