Information Sciences 616 (2022) 66–87

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Generalized multigranulation sequential three-way decision models for hierarchical classification

Jin Qian [a,d,*], Chengxin Hong [a], Ying Yu [a], Caihui Liu [b], Duoqian Miao [a,c]

[a] School of Software, East China Jiaotong University, Nanchang 330013, Jiangxi, China
[b] Department of Mathematics and Computer Science, Gannan Normal University, Ganzhou, 341000, China
[c] Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
[d] School of Computer, Jiangsu University of Science and Technology, Zhenjiang 212003, Jiangsu, China

ARTICLE INFO

Article history:
Received 4 July 2022
Received in revised form 22 August 2022
Accepted 3 October 2022
Available online 17 October 2022

Keywords:
Hierarchical rough set model
Multigranulation
Sequential three-way decisions
Hierarchical classification

ABSTRACT

Hierarchical classification is an important research hotspot in machine learning due to the widespread existence of data with hierarchical class structures. The existing sequential three-way decision models mainly constructed the hierarchical condition information granules via concept hierarchy tree to discuss the three probabilistic regions for flat classification. However, in real-world applications, one may face not only the tree-structured data with hierarchical condition attributes but also more often the multi-level data with hierarchical decision attribute (hierarchical class labels). How to obtain acceptable decisions under different levels of granularity is the most important issue within the multi-level and multi-view data. To this end, we construct a generalized hierarchical decision table and propose a generalized hierarchical multigranulation sequential three-way decision model by combining multi-granularity and sequential three-way decisions. Specifically, we first design a generalized hierarchical decision table using concept hierarchy trees of all conditional attributes and decision attribute, and explore some basic properties. Then we decompose and aggregate condition and decision granules under different levels of granularity, propose the optimistic and pessimistic generalized hierarchical multigranulation three-way decision models to update the three probabilistic regions for flat and hierarchical classification, and discuss the relationships between these two models. Finally, the experimental results demonstrate that the proposed models are more suitable for different applications. These models will provide a novel insight and enrich the development of multigranulation three-way decisions from the perspective of multi-level and multi-view.

1. Introduction

Three-way decision (3WD) [1–4] is an effective tool in solving complex problem. The basic idea of the model is to divide the objects of the universe into three pair-wise disjoint parts which are labelled as acceptance, rejection, and uncertainty, and to act on the three regions of objects with different strategies. For a cost-sensitive decision-making problem under multiple levels of granularity, sequential three-way decisions (S3WD) [5] mainly transform delayed decisions into definite (ac-

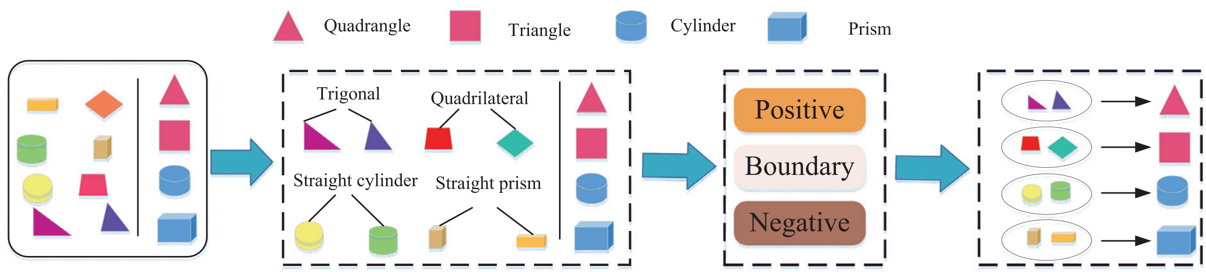
* Corresponding author.
E-mail addresses: qjqjlqyf@163.com (J. Qian), hongchengxinaq@163.com (C. Hong), yuyingjx@163.com (Y. Yu), liu_caihui@163.com (C. Liu), miaoduoqian@163.com (D. Miao).

https://doi.org/10.1016/j.ins.2022.10.014
0020-0255/

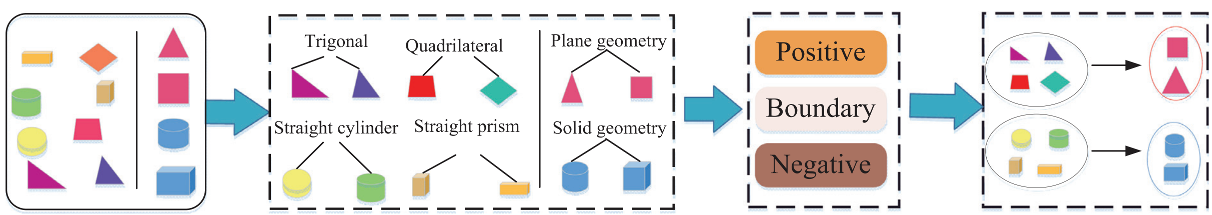
cept and reject) decisions by adding additional information. Nowadays, three-way decision model has been widely used in many fields, including attribute reduction [7,8], face recognition [6], project investment [9,10], image processing [11], text sentiment classification [12], three-way recommendation [13,14] and so on.

Granular computing (GrC) [15–19] is a paradigm of information processing, which simulates human thinking and solve the granule-oriented complex problem under different levels of granularity [20,21]. It has become a hotspot in the fields of artificial intelligence and knowledge discovery [22–26]. Many researchers mainly proposed multigranulation rough set model, hierarchical rough set model and multi-scale rough set model through constructing the granules and granular structures from multi-view and multi-level [27] for complex problem solving. Qian et al. [28] proposed multigranulation rough set to extend the classical single-granulation rough set by defining an approximation of the set using multiple equivalence relations on the universe. Feng and Miao [29] presented a hierarchical rough set model to transform one-dimensional data into multi-dimensional data by constructing a concept hierarchy tree from the ontology perspective of data. Further, Qian et al. [30] combined the hierarchical rough set model and MapReduce to propose a hierarchical attribute reduction algorithm. From the perspective of mathematics, Wu and Leung [31,32] proposed the multi-scale information tables using multi-scale granular labelled partition to describe the information granules of the scale. Huang et al. [33] proposed the generalized multi-scale decision tables by introducing multi-scale to decision classes and defined the optimal scale in the generalized multi-scale decision table, but neglected to consider the problem from a multi-granularity perspective. Although hierarchical rough set model and multi-scale rough set model have different starting points for hierarchical structured data, they can have the same effect for mining generalized decision rules.

In recent years, a unified model of sequential three-way decision and granular computing [5] is introduced to deal with the uncertainty and the cost of decision process and decision result. Hao et al. [34] introduced sequential three-way decisions into the multi-scale decision tables to study the optimal scale selection problem of dynamic sequential update information. Qian et al. [35] combined multi-granularity and three-way decision to implement five multigranulation sequential three-way decisions models with typical aggregation strategies. Qian et al. [36] proposed a hierarchical sequential three-way decision model by selecting sets of attributes with nested relationships and making sequential three-way decisions at multiple levels of granularity to obtain more refined rules from a single perspective. Up to until now, sequential three-way decisions have usually assume that the class labels are flat. In many practical classification problems such as web categorization, image recognition and gene classification, there are complex classification structures, where the class labels to be predicted are hierarchically organized [37–40]. Indeed, the knowledge representation and data analysis methods for flat classification are far from sufficient to meet the user needs of real-world applications for hierarchical classification. For example, given a hierarchy containing a path *Root* → *Sciences* → *Biology*, an article discussing about *Sciences* in general should only be categorized as *Root* → *Sciences* and not *Biology*. Another example is that in evaluating the appraisals of the company, the board of directors simply know the overall performance, while the general manager must have deep insights into the detailed ones of individual items. How to get a certain level of decision-making from multi-level and multi-view data is an important topic. It has been reported that hierarchical methods produce better performance than flat classification techniques on tree-structured hierarchies. Fig. 1 illustrates the advantage of the generalized hierarchical three-way decisions through the concept hierarchy tree for hierarchical structured data. One can generate four decision rules for flat classification in Fig. 1(a), and



(a) HS3WD model for flat classification



(b) GHS3WD model for hierarchical classification

Fig. 1. The advantage of the generalized hierarchical three-way decision models.

more easily acquire two generalized decision rules for hierarchical classification in Fig. 1(b). Unfortunately, the existing sequential three-way decision models have paid little attentions to the classification of data with hierarchical class labels from the perspective of multi-level and multi-view. To this end, inspired by the ideas of multi-view, multi-level and sequential three-way decision, we design a novel generalized hierarchical multi-granulation sequential three-way decision model. More specifically, we firstly construct information granules under different levels of granularity based on indiscernibility relations. Secondly, we propose a generalized hierarchical decision tables via the construction of the concept hierarchy trees for conditional and decision attributes. Finally, we design the optimistic and pessimistic multigranulation sequential three-way decision models by combining multi-granularity and sequential three-way decisions to discuss the relationships and properties of the three probabilistic regions.

The rest of the paper is organized as follows. Section 2 briefly reviews the Pawlak rough set model, hierarchical decision table, multi-granulation rough set and sequential three-way decisions. In Section 3, we construct the generalized hierarchical decision tables. Sections 4 and 5 propose a generalized hierarchical multi-granulation sequential three-way decision model by combining multi-granularity and sequential three-way decisions, strategies of condition granule level ascension and decision granule level descension are designed, then designs the corresponding algorithms and explore some properties of the proposed model. Section 6 gives the relevant experiments and conclusions. Finally, the paper ends with conclusions and further work in Section 7.

2. Preliminaries

In this section, we will review some basic concepts of Pawlak rough set model, multi-granulation rough set model, hierarchical decision table and sequential three-way decisions. For a detailed description, please refer to there classical papers [28,29,41,42,5].

2.1. Pawlak rough set model

In general, we use a four-tuple $S = (U, AT = C \cup D, \{V_a | a \in AT\}, \{f_a | a \in AT\})$ to represent the information system, where $U = \{x_1, x_2, \cdots, x_n\}$ denotes a non-empty finite set; $C = \{a_1, a_2, \cdots, a_s\}$denotes a non-empty finite attribute set, and $D = \{d\}$is a decision attribute; $V = \cup_{a \in At} V_a$, where V_a denotes the value of attribute a;f is an information function that maps an object x in U to exactly one value V in V_a. We focus on knowledge of the partition of the universe, since knowledge and equivalence relations can be mutually determined. We can define an equivalence relation for $A \subseteq AT$ as $IND(A) = \{(x, y) \in U \times U | \forall a \in A, f_a(x) = f_a(y)\}$, and the partition generated by $IND(A)$ is denoted as $U/IND(A)$, simply as π_A. For simplicity, we generally denote $[x]_A$ instead of $[x]_{IND(A)}$. Consider a partition $\pi_D = \{D_1, D_2, \ldots, D_k\}$on universe U with respect to the decision attribute and a partition $\pi_A = \{A_1, A_2, \ldots, A_q\}$on universe U with respect to conditional attributes.

Definition 1. For a decision table S, given a concept X and a decision class $D_j \in \pi_D$, the upper and lower approximations of D_j with respect to A_i are defined as follows:

$$\begin{aligned} \underline{apr}_{A_i}(X) &= \left\{ x \in U | [x]_{A_i} \subseteq D_j \right\}, \\ \overline{apr}_{A_i}(X) &= \left\{ x \in U | [x]_{A_i} \cap D_j \neq \phi \right\}. \end{aligned} \tag{1}$$

where $|\cdot|$ is the cardinality of the set.

Therefore, the universe is divided into three pair-wise disjoint regions based on the upper and lower approximations, namely the positive region $POS_{A_i}(\pi_D)$, boundary region $BND_{A_i}(\pi_D)$ and negative region $NEG_{A_i}(\pi_D)$:

$$\begin{aligned} POS_{A_i}(\pi_D) &= \underset{1 \leqslant j \leqslant k}{\cup} \underline{apr}_{A_i}(D_j); \\ BND_{A_i}(\pi_D) &= \underset{1 \leqslant j \leqslant k}{\cup} BND_{A_i}(D_j) \\ &= \underset{1 \leqslant j \leqslant k}{\cup} \left(\overline{apr}_{A_i}(D_j) - \underline{apr}_{A_i}(D_j) \right); \\ NEG_{A_i}(\pi_D) &= U - POS_{A_i}(\pi_D) \cup BND_{A_i}(\pi_D). \end{aligned} \tag{2}$$

2.2. Hierarchical decision table

Feng and Miao [29] proposed a hierarchical rough set model by combining concept hierarchy tree and rough set to describe multi-dimensional data. In what follows, we briefly review the hierarchical decision table.

Definition 2. Let $HT = \left(U, \{a_i^l | i = 1, 2, \ldots, s; l = 1, 2, \ldots, m\}, V, f\right)$ be a hierarchical table, $\{a_i | i = 1, 2, \ldots, s\}$denotes a non-empty finite attribute set, and the conditional attribute a_i has m levels $C = \{a_i^l | i = 1, 2, \ldots, s; l = 1, 2, \ldots, m\}$.

Definition 3. Let $HT = (U, AT = \{a_i^l | i = 1,2,\ldots,s; l = 1,2,\ldots,m\} \cup \{d\}, V, f)$ be a hierarchical decision table. The index set $L = (l_1, l_2, \ldots, l_s)$ is called a level combination of conditional attributes, which denotes the combination of the conditional attribute a_i at the l_i-th level ($i = 1,2,\ldots,s$). Each level combination $L = (l_1, l_2, \ldots, l_s)$ can form a single-level information table $C^L = \left\{a_1^{l_1}, a_2^{l_2}, \ldots, a_s^{l_s}\right\}$.

Definition 4. Given the $L_1 = \left(l_1^1, l_2^1, \ldots, l_s^1\right)$-th decision table and $L_2 = \left(l_1^2, l_2^2, \ldots, l_s^2\right)$-th decision table, if $l_i^1 \leqslant l_i^2$ ($i = 1,2,\ldots,s$), then L_1 is said to be coarser than L_2 or L_2 is the finer than L_1, which can be denoted as $L_1 \succcurlyeq L_2$. Furthermore, if there exists $i \in \{1,2,\ldots,s\}$ such that $l_i^1 < l_i^2$, then L_1 is said to be strictly coarser than L_2 or L_2 is the strictly finer than L_1, namely $L_1 \succ L_2$.

2.3. QIAN's MGRS

In the Pawlak rough set model, the concepts are often expressed in terms of a single equivalence relation, but in some situations, we describe a relatively complex target concept in terms of multiple equivalence relations depending on the user demand and the choice of goals to solve problems [43–45]. Qian et al. [28] proposed a multi-granulation rough set model. In what follows, the optimistic and pessimistic multi-granulation rough set models are briefly reviewed below.

Definition 5. Given a granular structure $GS = \{A_1, A_2, \cdots, A_q\}$ and $\forall X \subseteq U$, the optimistic multigranulation lower and upper approximations $\underline{\sum_{i=1}^q A_i^O}(X)$ and $\overline{\sum_{i=1}^q A_i^O}(X)$ are defined as follows:

$$
\begin{aligned}
\underline{\sum_{i=1}^q A_i^O}(X) &= \left\{x \in U | [x]_{A_1} \subseteq X \vee [x]_{A_2} \subseteq X \vee \cdots \vee [x]_{A_q} \subseteq X\right\}, \\
\overline{\sum_{i=1}^q A_i^O}(X) &= \sim \underline{\sum_{i=1}^q A_i^O}(\sim X).
\end{aligned} \tag{3}
$$

where$\sim X$ is the complement of set X.

The pair $< \underline{\sum_{i=1}^q A_i^O}(X), \overline{\sum_{i=1}^q A_i^O}(X) >$is called the optimistic multigranulation rough sets of X.

Definition 6. Given a granular structure $GS = \{A_1, A_2, \cdots, A_q\}$and $\forall X \subseteq U$, the pessimistic multi-granulation lower and upper approximations $\underline{\sum_{i=1}^q A_i^P}(X)$ and $\overline{\sum_{i=1}^q A_i^P}(X)$ are defined as follows:

$$
\begin{aligned}
\underline{\sum_{i=1}^q A_i^P}(X) &= \left\{x \in U | [x]_{A_1} \subseteq X \wedge [x]_{A_2} \subseteq X \wedge \cdots \wedge [x]_{A_q} \subseteq X\right\}, \\
\overline{\sum_{i=1}^q A_i^P}(X) &= \sim \underline{\sum_{i=1}^q A_i^P}(\sim X).
\end{aligned} \tag{4}
$$

where$\sim X$ is the complement of set X.

The pair $< \underline{\sum_{i=1}^q A_i^P}(X), \overline{\sum_{i=1}^q A_i^P}(X) >$is called the pessimistic multi-granulation rough sets of X.

2.4. Sequential three-way decisions

As we all know, the sequential three-way decision model is the evolution of the sequential multi-step classical three-way decision [5,46]. In what follows, we briefly review the sequential three-way decisions.

Definition 7. Given a l-th level of the granular structure $GS^l = \left\{A_1^l, A_2^l, \cdots, A_q^l\right\}$and a decision class D_j, the lower approximation $\underline{apr}_{A_i^l}(D_j)$ and the upper approximation $\overline{apr}_{A_i^l}(D_j)$ are defined by

$$
\begin{aligned}
\underline{apr}_{A_i^l}(D_j) &= \left\{x \in U^l | [x]_{A_i^l} \subseteq D_j\right\}, \\
\overline{apr}_{A_i^l}(D_j) &= \left\{x \in U^l | [x]_{A_i^l} \cap D_j \neq \varnothing\right\}.
\end{aligned} \tag{5}
$$

where $U^1 = U, U^{l+1} = BND_{A_i^l}(D_j) = \overline{apr}_{A_i^l}(D_j) - \underline{apr}_{A_i^l}(D_j)$, $[x]_{A_i^l}$ denotes the equivalence class containing x in the partition U^l / A_i^l.

This pair $< \underline{apr}_{A_i^l}(D_j), \overline{apr}_{A_i^l}(D_j) >$is called the lower and upper approximations induced by A_i^l with respect to D_j. We can obtain the following three disjoint regions.

$$POS_{A_i^l}(D_j) = \underline{apr}_{A_i^l}(D_j); \tag{6}$$

$$BND_{A_i^l}(D_j) = \overline{apr}_{A_i^l}(D_j) - \underline{apr}_{A_i^l}(D_j); \tag{7}$$

$$NEG_{A_i^l}(D_j) = U^l - POS_{A_i^l}(D_j) - NEG_{A_i^l}(D_j). \tag{8}$$

3. Construction of the generalized hierarchical decision tables

In this section, we first briefly illustrate the construction of information granules and concept hierarchy tree, then define the generalized hierarchical decision tables, and finally explore some related properties.

3.1. Construction of concept hierarchy and information granules

The conditional attribute a_i is formed along the $l(a_i)+1$ hierarchy level: $0, 1, \cdots, l(a_i)$. Level 0 is the special value of Any (*). Similarly, the depth of the concept hierarchy tree of the decision attribute is $l(d)+1$. Fig. 2 shows the concept hierarchy tree constructed by each attribute, raw data is represented as a lattice. In order to simplify this study, we extend the attributes of the insufficient hierarchy. If the highest level l of attribute a_i satisfies $0 < l < m$, extend it to $a_i^1, a_i^2, \ldots, a_i^l, a_i^{(l+1)*}, \ldots, a_i^{(m)*}$, fill all attributes of levels $l+1$ to m with a_i^l. The method of decision attributes extension is similar to that of the conditional attributes extension. For readability, the description of the symbols throughout our paper are shown in Table 1.

As we all know, information granules are constructed by the objects based on indiscernibility. However, constructing an ideal information granule is a complex and interesting issue. In Fig. 3, we use aggregation and decomposition operations to construct information granules based on an indistinguishable relation. Red arrows and blue arrows indicate refinement paths and coarsened paths, respectively. For convenience, $cg_i^{l,t}$ denotes a condition granule, where l represents the level and t denotes the number of the ascension attributes in a condition granule, respectively. In other words, a multi-level granular structure $GS^{l,t}$ is a parallel set of condition granules $\left\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\right\}$.

In Fig. 3, the top node $cg_1^{1,0}$ represents the most generalized condition granule, while the bottom node $cg_1^{2,3}$ denotes the detailed condition granule. Note that the node $cg_1^{1,3}$ is the largest granule at the first level, and node $cg_1^{2,0}$ is the smallest granule at the second level, but these two granules are the same granule under different levels of granularity.

Then the positive region and the classification accuracy of D_j^h with respect to $cg_i^{l,t}$ are defined as

$$Pos_{cg_i^{l,t}}\left(\pi_D^h\right) = \cup_{D_j^h \in U/E_{dg^h}} \underline{E}_{cg_i^{l,t}}\left(D_j^h\right); \tag{9}$$

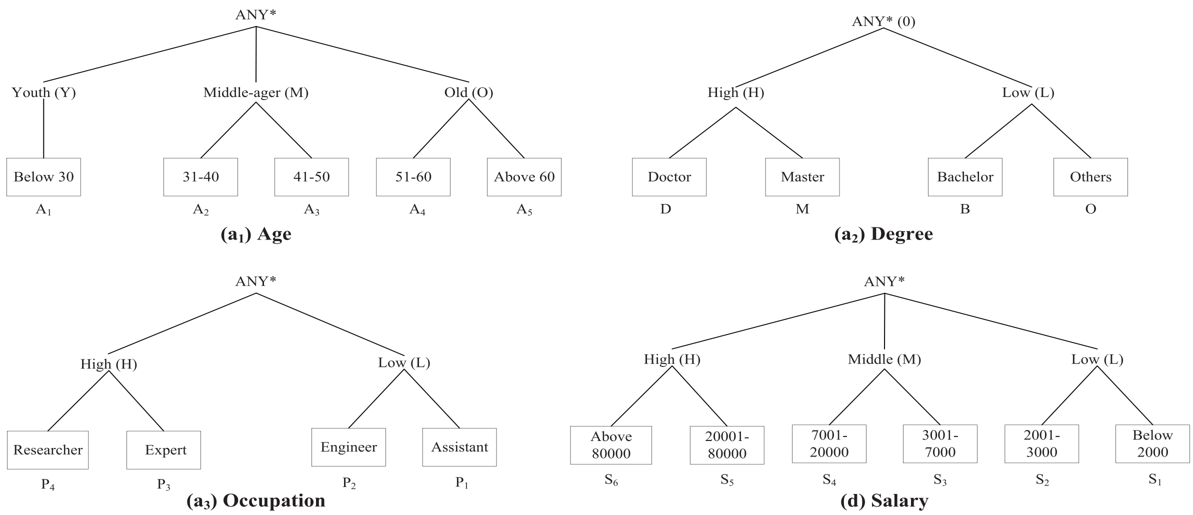


Fig. 2. Concept hierarchy tree for each attribute.

Table 1
Description of the symbols.

Symbol	Meaning
l	Condition granule (conditional attribute) level
t	Number of attributes ascending in a condition granule
m	Maximal level of the condition granule
q	Maximal serial number of the condition granule
n	Maximal level of the decision granule
h	Decision granule (decision attribute) level
j	Serial number of decision class
k	Maximal serial number of decision class
cg_i	i-th condition granule with serial number i where $i \in \{1, 2, \ldots q\}$
$cg_i^{l,t}$	i-th condition granule with the number of ascending attribute number t at level l
GS	Granular structure $GS = \left\{cg_1, cg_2, \ldots, cg_q\right\}$
$GS^{l,t}$	Granular structure with the number of ascending attributes t at level l
d^h	Decision granule (decision attribute) at level h
D_j	j-th decision class where $j \in \{1, 2, \ldots k\}$
D_j^h	j-th decision class at level h

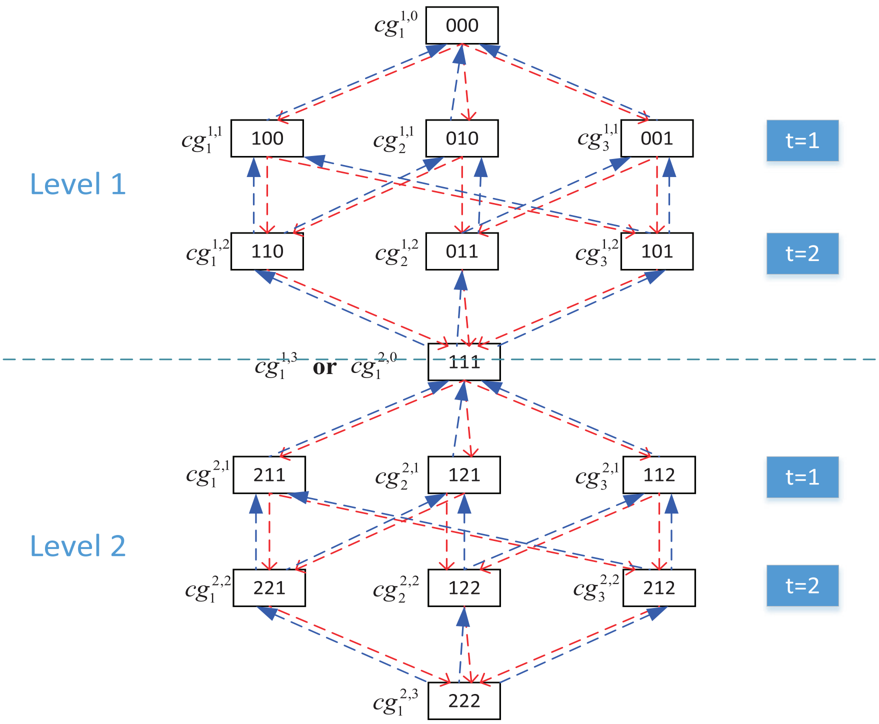


Fig. 3. Condition granules under different levels of granularity.

$$Acc\left(cg_i^{l,t}, \pi_D^h\right) = \frac{|Pos_{cg_i^{l,t}}(\pi_D^h)|}{|U|}. \tag{10}$$

3.2. Generalized hierarchical decision table

In this section, we define the generalized hierarchical decision table and explore the related properties.

Definition 8. Let $GH = \left(U, \{a_i^l | i = 1, 2, \ldots, q; l = 1, 2, \ldots, m\} \cup \left\{d^h | h = 1, 2, \ldots, n\right\}, V, f\right)$ be a generalized hierarchical decision table, where q denotes the maximal number of conditional attributes, m denotes the maximal level of conditional attributes, and n denotes the maximal level of decision attributes, $D = \{d\}$ is a non-empty finite set of decision attribute, and d has n levels $D = \left\{d^h | h = 1, 2, \ldots, n\right\}$.

Through Definition 8, we know that the generalized hierarchical decision tables are composed of $\left(\prod_{j=1}^{m} I_j\right) \times n$ single decision table, where I_j denotes the level of conditional attribute a_j.

If d is strictly confined to the h-th level, the generalized hierarchical decision table $GH = \left(U, \{a_i^l | i = 1, 2, \ldots, q; l = 1, 2, \ldots, m\} \cup \left\{d^h | h = 1, 2, \ldots, n\right\}, V, f\right)$ can be denoted as $GH^h = \left(U, \{a_i^l | i = 1, 2, \ldots, q; l = 1, 2, \ldots, n\} \cup \left\{d^h\right\}, V, f\right)$.

Definition 9. Let $GH = \left(U, \{a_i^l | i = 1, 2, \ldots, q; l = 1, 2, \ldots, m\} \cup \left\{d^h | h = 1, 2, \ldots, n\right\}, V, f\right)$ be a generalized hierarchical decision table, the indiscernible relationship on $cg_i^{l,t}$ is defined as

$$IND\left(cg_i^{l,t}\right) = \left\{(x, y) \in U \times U | \forall a^l \in cg_i^{l,t}, f(x, a^l) = f(y, a^l)\right\}. \tag{11}$$

For an arbitrary $x \in U$, the equivalence relation $E_{cg_i^{l,t}}$ is derived from the indiscernible relation $IND\left(cg_i^{l,t}\right)$. Thus, the equivalence class of x according to the indiscernible relation $IND\left(cg_i^{l,t}\right)$ can be notated as $[x]_{cg_i^{l,t}}$ as follows:

$$[x]_{cg_i^{l,t}} = \left\{y \in U | (x, y) \in E_{cg_i^{l,t}}\right\}. \tag{12}$$

Then, we can obtain a family of the equivalence classes $U/IND\left(cg_i^{l,t}\right)$ as

$$U/IND\left(cg_i^{l,t}\right) = \left\{[x]_{cg_i^{l,t}} | x \in U\right\}. \tag{13}$$

The lower approximation of X with respect to $cg_i^{l,t}$ are defined as

$$\underline{E}_{cg_i^{l,t}}(X) = \cup\left\{x | [x]_{cg_i^{l,t}} \subseteq X\right\}. \tag{14}$$

Proposition 1. Let $GH = \left(U, \{a_i^l | i = 1, 2, \ldots, s; l = 1, 2, \ldots, m\} \cup \left\{d^t | t = 1, 2, \ldots, n\right\}, V, f\right)$ be a generalized hierarchical decision table, l and l' represent the level of condition granules, h and h' denote the level of decision granules, the following properties hold true:

(1) $l < l' \rightarrow E_{cg_i^{l,t}} \precsim E_{cg_i^{l',t}}$;
(2) $l < l' \rightarrow U/E_{cg_i^{l,t}} \precsim U/E_{cg_i^{l',t}}$;
(3) $l < l' \rightarrow \underline{E}_{cg_i^{l,t}}(X) \subseteq \underline{E}_{cg_i^{l',t}}(X)$;
(4) $h < h' \rightarrow E_{dg^h} \succsim E_{dg^{h'}}$;
(5) $h < h' \rightarrow U/E_{dg^h} \succsim U/E_{dg^{h'}}$;
(6) $h < h' \rightarrow [x]_{dg^h} \supseteq [x]_{dg^{h'}}$.

In Proposition 1, the properties (1) - (3) show that the equivalence relation, the partition and the lower approximation of X induced by the corresponding attributes become coarser with the level ascension of condition granules while the properties (4)-(6) illustrate that the equivalence relation, the partition induced and the equivalence classes of x decrease when the level of decision granules increases.

It should be noted that this paper mainly discusses the changes of condition granules at the l-th level and decision granules at the h-th level. The changes of the number of attributes ascending t is similar to those of condition granules at the l-th level, which will not be repeatedly listed here.

Proposition 2. Let $GH = \left(U, \{a_i^l | i = 1, 2, \ldots, s; l = 1, 2, \ldots, m\} \cup \left\{d^h | h = 1, 2, \ldots, n\right\}, V, f\right)$ be a generalized hierarchical decision table, l and l' represent the level of condition granules, and h and h' denote the level of decision granules, then

(1) $l < l' \rightarrow Pos_{cg_i^{l,t}}(\pi_D^h) \subseteq Pos_{cg_i^{l',t}}(\pi_D^h)$;
(2) $l < l' \rightarrow Acc\left(cg_i^{l,t}, \pi_D^h\right) \leqslant Acc\left(cg_i^{l',t}, \pi_D^h\right)$;
(3) $h < h' \rightarrow Pos_{cg_i^{l,t}}(\pi_D^h) \supseteq Pos_{cg_i^{l,t}}(\pi_D^{h'})$;
(4) $h < h' \rightarrow Acc\left(cg_i^{l,t}, \pi_D^h\right) \geqslant Acc\left(cg_i^{l,t}, \pi_D^{h'}\right)$.

In Proposition 2, the properties (1) and (2) show that the size of positive region and the classification accuracy increase as the condition granules become finer while (3) and (4) illustrate that the size of positive region and the classification accuracy decrease with the increasing level of decision granules.

4. Multigranulation sequential three-way decision model under condition granule level ascension for hierarchical classification

In what follows, we first combine sequential three-way decisions and multi-granularity to design a generalized hierarchical multigranulation sequential three-way decision model (GHMS3WD) at the fixed decision granule level, then propose a corresponding algorithm for computing the three probabilistic regions, and finally discuss the relationships and some properties. In GHMS3WD, we refer to the decision attribute as decision granule. Fig. 4 illustrates the sequential three-way decisions at different levels of condition granules and decision granules. The sequential three-way decision process for condition granules from level 1 to level m is shown in Fig. 4(a).

Definition 10. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \left\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\right\}$, the lower and upper approximations $\underline{apr}_{cg_i^{l,t}}\left(D_j^h\right)$ and $\overline{apr}_{cg_i^{l,t}}\left(D_j^h\right)$ are defined by

$$\begin{aligned} \underline{apr}_{cg_i^{l,t}}\left(D_j^h\right) &= \left\{x \in U^{l,t} | [x]_{cg_i^{l,t}} \subseteq D_j^h\right\}, \\ \overline{apr}_{cg_i^{l,t}}\left(D_j^h\right) &= \left\{x \in U^{l,t} | [x]_{cg_i^{l,t}} \cap D_j^h \neq \varnothing\right\}. \end{aligned} \tag{15}$$

where $U^{1,t} = U, U^{l+1,t} = \overline{apr}_{cg_i^{l,t}}\left(D_j^h\right) - \underline{apr}_{cg_i^{l,t}}\left(D_j^h\right)$ is the gradually reduced universe.

Proposition 3. Given a generalized hierarchical decision table GH, the following updated three probabilistic regions of the condition granule $cg_i^{l,t}$ with respect to D_j^h hold true:

(1)

$$\begin{aligned} POS_{cg_i^{l,t}}\left(D_j^h\right) &= \underline{apr}_{cg_i^{l,t}}\left(D_j^h\right) \\ &= POS_{cg_i^{l-1,t}}\left(D_j^h\right) \cup \left\{x | [x]_{cg_i^{l,t}} \subseteq D_j^h, x \in BND_{cg_i^{l-1,t}}\left(D_j^h\right)\right\}; \end{aligned} \tag{16}$$

(2)

$$\begin{aligned} BND_{cg_i^{l,t}}\left(D_j^h\right) &= \overline{apr}_{cg_i^{l,t}}\left(D_j^h\right) - \underline{apr}_{cg_i^{l,t}}\left(D_j^h\right) \\ &= U^{l,t} - POS_{cg_i^{l,t}}\left(D_j^h\right) - NEG_{cg_i^{l,t}}\left(D_j^h\right); \end{aligned} \tag{17}$$

(3)

$$\begin{aligned} NEG_{cg_i^{l,t}}\left(D_j^h\right) &= U^{l,t} - \overline{apr}_{cg_i^{l,t}}\left(D_j^h\right) \\ &= NEG_{cg_i^{l-1,t}}\left(D_j^h\right) \cup \left\{x | [x]_{cg_i^{l,t}} \cap D_j^h \neq \varnothing, x \in BND_{cg_i^{l-1,t}}\left(D_j^h\right)\right\}. \end{aligned} \tag{18}$$

Proof. The equivalence class $[x]_{cg_i^{l,t}}$ of the l-th level will be further divided into the equivalence class $[x]_{cg_i^{l+1,t}}$ of the *(l + 1)*-th level. We easily get $[x]_{cg_i^{l,t}} \subseteq [x]_{cg_i^{l-1,t}}$.

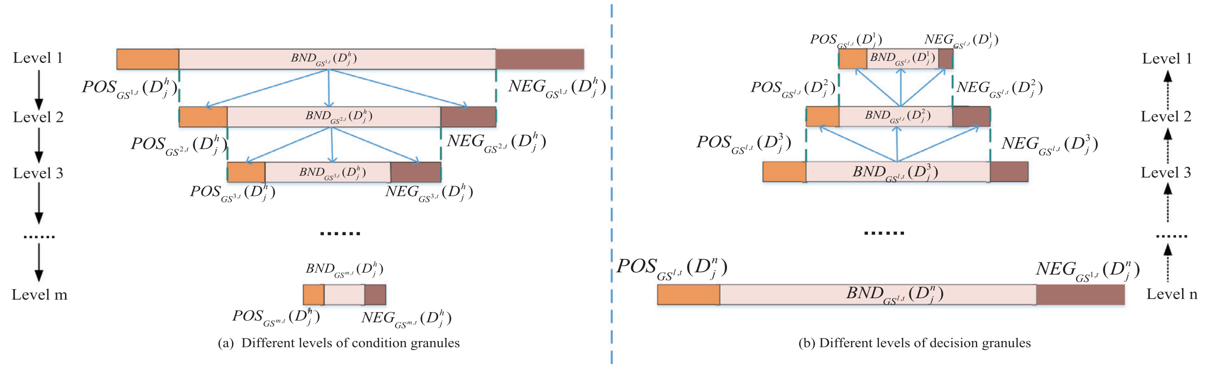


Fig. 4. The sequential process between different condition granules and decision granules.

(1) For any $x \in POS_{cg_i^{l-1,t}}\left(D_j^h\right)$, we get $[x]_{cg_i^{l-1,t}} \subseteq D_j^h$. Then, $[x]_{cg_i^{l,t}} \subseteq D_j^h$ holds when $[x]_{cg_i^{l,t}} \subseteq [x]_{cg_i^{l-1,t}}$. So, we obtain the result that $x \in POS_{cg_i^{l,t}}\left(D_j^h\right)$.

(2) For any $x \in BND_{cg_i^{l,t}}\left(D_j^h\right)$, it is easy to obtain $[x]_{cg_i^{l,t}} \cap D_j^h \neq \emptyset$and $[x]_{cg_i^{l,t}} \neg\subseteq D_j^h$, which implies $x \notin NEG_{cg_i^{l,t}}\left(D_j^h\right)$ and $x \notin POS_{cg_i^{l,t}}\left(D_j^h\right)$ because of $[x]_{cg_i^{l,t}} \subseteq [x]_{cg_i^{l-1,t}}$. As a result, $x \in BND_{cg_i^{l-1,t}}\left(D_j^h\right)$ holds.

(3) For any $x \in NEG_{cg_i^{l-1,t}}\left(D_j^h\right)$, we have $[x]_{cg_i^{l-1,t}} \subseteq D_j^h$. It is easy to know that $[x]_{cg_i^{l,t}} \subseteq D_j^h$ is true when $[x]_{cg_i^{l,t}} \subseteq [x]_{cg_i^{l-1,t}}$. So, $x \in NEG_{cg_i^{l,t}}\left(D_j^h\right)$. □

Proposition 4. Given a generalized hierarchical decision table $GH, POS_{cg_i^{l,t}}\left(D_j^h\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^h\right)$ monotonically become bigger while $BND_{cg_i^{l,t}}\left(D_j^h\right)$ turns monotonically smaller when the level of condition granule increases.

Proof. From Proposition 3, we know that $POS_{cg_i^{l,t}}\left(D_j^h\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^h\right)$ are parts of the $POS_{cg_i^{l+1,t}}\left(D_j^h\right)$ and $NEG_{cg_i^{l+1,t}}\left(D_j^h\right)$, and the added ones are selected from $BND_{cg_i^{l,t}}\left(D_j^h\right)$. Thus, $POS_{cg_i^{l,t}}\left(D_j^h\right) \subseteq POS_{cg_i^{l+1,t}}\left(D_j^h\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^h\right) \subseteq NEG_{cg_i^{l+1,t}}\left(D_j^h\right)$. Since $BND_{cg_i^{l,t}}\left(D_j^h\right) = U^{l,t} - POS_{cg_i^{l,t}}\left(D_j^h\right) - NEG_{cg_i^{l,t}}\left(D_j^h\right)$ and $BND_{cg_i^{l+1,t}}\left(D_j^h\right) = BND_{cg_i^{l,t}}\left(D_j^h\right) - POS_{cg_i^{l+1,t}}\left(D_j^h\right) - NEG_{cg_i^{l+1,t}}\left(D_j^h\right)$, then $BND_{cg_i^{l,t}}\left(D_j^h\right) \supseteq BND_{cg_i^{l+1,t}}\left(D_j^h\right)$.

In what follows, we construct an algorithm to compute the regions of sequential three-way decisions under a condition granule as shown in Algorithm 1. The main idea of Algorithm 1 is to first deletes the objects belonging to the probabilistic positive region or negative region under the first level of granularity, and then obtain the updated region $U^{2,t} = BND_{cg_i^{l,t}}\left(D_j^h\right)$. For the updated region $U^{2,t}$, delete the objects belonging to the positive region and negative region at the next level of granularity, and repeat these steps until the updated universe becomes an empty set or no level of granularity can be computed. It is easy to observe that the time complexity of Algorithm 1 is $O\left(m|D_j^h||U|^2\right)$.

Algorithm 1: Computing the regions of sequential three-way decisions under a granular structure.

input : An universal set of object, U; the number of attributes ascending, t; a condition granule, $cg_i^{l,t}, l \in \{1, 2, ..., m\}$
output: Three regions, $POS_{cg_i}(D_j^h)$, $BND_{cg_i}(D_j^h)$ and $NEG_{cg_i}(D_j^h)$

```
1  POS_cgi(D_j^h) = ∅, BND_cgi(D_j^h) = U and NEG_cgi(D_j^h) = ∅;
2  l = 1, U^{1,t} = U;
3  for l ← 1 to m do
4      if U^{l,t} = ∅ or l > m turn to 9;
5      Compute POS_{cgi^{l,t}}(D_j^h) and NEG_{cgi^{l,t}}(D_j^h) according to Definition 10;
6      POS_cgi(D_j^h) = POS_cgi(D_j^h) ∪ POS_{cgi^{l,t}}(D_j^h) ; NEG_cgi(D_j^h) = NEG_cgi(D_j^h) ∪ NEG_{cgi^{l,t}}(D_j^h);
7      BND_cgi(D_j^h) = U^{l,t} − POS_{cgi^{l,t}}(D_j^h) − NEG_{cgi^{l,t}}(D_j^h);
8      U^{l+1,t} = BND_cgi(D_j^h) ;
9  end
10 Output POS_cgi(D_j^h), BND_cgi(D_j^h) and NEG_cgi(D_j^h);
```

To understand this algorithm, we illustrate the idea through Example 1. Table 2 is a generalized hierarchical decision table that describes the income of the Developers, $U = \{x_1, x_2, \ldots, x_{12}\}$ is composed of Developers, {Age,Degree,Occupation} are conditional attributes and Salary is a decision attribute.

•;For the conditional attribute a_1, the first values of the concept hierarchy are $\{Y, M, O\}$ which represent 'Youth', 'Middle-ager' and 'Old'. The second values of the concept hierarchy are $\{A_1, A_2, A_3, A_4, A_5\}$ which represent 'Below 30', '31–40', '41–50', '51–60' and 'Above 60'.

•;For the conditional attribute a_2, the first values of the concept hierarchy are $\{H, L\}$ which represent 'High' and 'Low'. The second values of concept hierarchy are $\{D, M, B, O\}$ which represent 'Doctor', 'Master', 'Bachelor' and 'Others'.

•;For the conditional attribute a_3, the first values of the concept hierarchy are $\{H, M, L\}$ which represent 'High', and 'Low'. The second values of concept hierarchy are $\{P_4, P_3, P_2, P_1\}$ which represent 'Researcher', 'Expert', 'Engineer' and 'Assistant'.

•;For the decision attribute d, the first values of the concept hierarchy are $\{H, M, L\}$ which represent 'High', 'Middle' and 'Low'. The second values of the concept hierarchy are $\{S_6, S_5, S_4, S_3, S_2, S_1\}$ which represent 'Above 80000', '20001–80000', '7001–20000', '3001–7000', '2001–3000' and 'Below 2000'.

Table 2
Generalized hierarchical decision tables.

U	Age		Degree		Occupation		Salary	
	a_1^1	a_1^2	a_2^1	a_2^2	a_3^1	a_3^2	d^1	d^2
x_1	Y	A_1	H	M	H	P_3	M	S_3
x_2	O	A_4	L	B	H	P_3	M	S_4
x_3	O	A_4	H	M	H	P_4	M	S_4
x_4	M	A_2	L	B	L	P_2	M	S_3
x_5	O	A_5	L	B	H	P_4	H	S_6
x_6	Y	A_1	H	M	H	P_4	M	S_4
x_7	M	A_3	H	M	H	P_4	M	S_4
x_8	Y	A_1	L	O	L	P_2	M	S_3
x_9	Y	A_1	L	B	L	P_1	L	S_2
x_{10}	O	A_4	L	B	H	P_4	H	S_5
x_{11}	M	A_3	L	O	H	P_3	M	S_4
x_{12}	M	A_3	H	D	H	P_4	H	S_5

Example 1. As shown in Table 2, consider $U^{1,t} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}, t = 2$, a decision class $D_1^1 = \{x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_{11}\}$, the condition granules $cg_1^{1,2} = \{a_1^1, a_2^1, a_3^0\}$ and $cg_1^{2,2} = \{a_1^2, a_2^2, a_3^1\}$.

(1) $U^{1,1} = U$. For the first level of granularity, we can compute
$U^{1,2}/cg_1^{1,2} = \{\{x_1, x_6\}, \{x_8, x_9\}, \{x_2, x_5, x_{10}\}, \{x_3\}, \{x_4, x_{11}\}, \{x_7, x_{12}\}\}$;
$POS_{cg_1^{1,2}}\left(D_1^1\right) = \{x_1, x_3, x_4, x_6, x_{11}\}$;
$BND_{cg_1^{1,2}}\left(D_1^1\right) = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$;
$NEG_{cg_1^{1,2}}\left(D_1^1\right) = \emptyset$.
Thus, we can have $POS_{cg_1}\left(D_1^1\right) = \{x_1, x_3, x_4, x_6, x_{11}\}$, $BND_{cg_1}\left(D_1^1\right) = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$, $NEG_{cg_1}\left(D_1^1\right) = \emptyset$, and $U^{2,1} = BND_{cg_1}\left(D_1^1\right) = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$.
(2) Updating the reduced universe $U^{2,1} = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$. For the second level of granularity, we can compute
$U^{2,1}/cg_1^{2,2} = \{\{x_2, x_{10}\}, \{x_5\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{12}\}\}$;
$POS_{cg_1^{2,2}}\left(D_1^1\right) = \{x_7, x_8\}$;
$BND_{cg_1^{2,2}}\left(D_1^1\right) = \{x_2, x_{10}\}$;
$NEG_{cg_1^{2,2}}\left(D_1^1\right) = \{x_5, x_9, x_{12}\}$.
Therefore, we can obtain $POS_{cg_1}\left(D_1^1\right) = POS_{cg_1}\left(D_1^1\right) \cup POS_{cg_1^{2,2}}\left(D_1^1\right) = \{x_1, x_3, x_4, x_6, x_7, x_8, x_{11}\}$, $BND_{cg_1}\left(D_1^1\right) = U^{2,1} - POS_{cg_1^{2,2}}\left(D_1^1\right) - NEG_{cg_1^{2,2}}\left(D_1^1\right) = \{x_2, x_{10}\}$, and $NEG_{cg_1}\left(D_1^1\right) = NEG_{cg_1}\left(D_1^1\right) \cup NEG_{cg_1^{2,2}}\left(D_1^1\right) = \{x_5, x_9, x_{12}\}$.

In what follows, we design the optimistic and pessimistic generalized hierarchical multigranulation sequential three-way decision models, and discuss the corresponding properties.

Definition 11. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, then the lower and upper approximations of the optimistic generalized hierarchical multigranulation sequential three-way decision with respect to D_j^h are defined as

$$\begin{aligned} \underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) &= \left\{x \in U^{l,t} \mid [x]_{cg_1^{l,t}} \subseteq D_j^h \vee [x]_{cg_2^{l,t}} \subseteq D_j^h \vee \ldots \vee [x]_{cg_q^{l,t}} \subseteq D_j^h\right\}, \\ \overline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) &= \sim \underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(\sim D_j^h\right). \end{aligned} \tag{19}$$

where $U^{1,t} = U, U^{l+1,t} = \overline{\sum_{i=1}^{m} cg_i^{l,t\,O}}\left(D_j^h\right) - \underline{\sum_{i=1}^{m} cg_i^{l,t\,O}}\left(D_j^h\right)$ is the gradually reduced universe, q represents the maximum number of parallel condition granule sets, and $[x]_{cg_i^{l,t}} (1 \leqslant i \leqslant q)$ denotes the equivalence class containing x in the partition $U^l/cg_i^{l,t}$.

The pair $< \underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right), \overline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) >$is called the optimistic generalized hierarchical rough set of D_j^h with respect to the set of condition granule cg in $U^{l,t}$.

Definition 12. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t}$ = $\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, then the lower and upper approximations of pessimistic generalized hierarchical multigranulation sequential three-way decision with respect to D_j^h are defined as

$$\begin{aligned} \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) &= \left\{x \in U^{l,t} | [x]_{cg_1^{l,t}} \subseteq D_j^h \wedge [x]_{cg_2^{l,t}} \subseteq D_j^h \wedge \ldots \wedge [x]_{cg_q^{l,t}} \subseteq D_j^h\right\}, \\ \overline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) &= \sim \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(\sim D_j^h\right). \end{aligned} \tag{20}$$

where $U^1 = U, U^{l+1} = \overline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) - \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right)$ is the gradually reduced universe, q represents the maximal number of parallel condition granule sets, and $[x]_{cg_i^{l,t}} (1 \leqslant i \leqslant q)$ denotes the equivalence class containing x in the partition $U^l / cg_i^{l,t}$.

The pair $< \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right), \overline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) >$is called the pessimistic generalized hierarchical rough set of D_j^h with respect to the condition granules cg in $U^{l,t}$.

Definition 13. Given a multilevel granular structure $GS^{l,t}$ = $\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, a parallel set of condition granules $cg_i^{l,t} \in GS^{l,t}$, and a decision partition $\pi_D^h = \left\{D_1^h, D_2^h, \ldots, D_k^h\right\}$, then the lower and upper approximations with respect to a partition π_D^h are defined as

$$\begin{aligned} \underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}(\pi_D^h) &= \left(\underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_1^h\right), \underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_2^h\right), \ldots, \underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_k^h\right)\right), \\ \overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}(\pi_D^h) &= \left(\overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_1^h\right), \overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_2^h\right), \ldots, \overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_k^h\right)\right). \end{aligned} \tag{21}$$

where Δdenotes a generalized aggregation strategy.

Based on the above definition, the positive, boundary and negative regions of π_D^h are defined as follows:

$$\begin{aligned} & POS_{GS^{l,t}}^{\Delta}(\pi_D^h) = \bigcup_{1 \leqslant j \leqslant k} \underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_j^h\right); \\ & BND_{GS^{l,t}}^{\Delta}(\pi_D^h) = \bigcup_{1 \leqslant j \leqslant k} \left(\overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_j^h\right) - \underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_j^h\right)\right); \\ & NEG_{GS^{l,t}}^{\Delta}(\pi_D^h) = U^{l,t} - POS_{GS^{l,t}}^{\Delta}(\pi_D^h) \cup BND_{GS^{l,t}}^{\Delta}(\pi_D^h). \end{aligned} \tag{22}$$

It should be pointed out that $NEG_{GS}^{\triangle}(\pi_D)$ is empty set. Thus, the negative regions of π_D are not considered. By observing these definitions, we can easily derive the relationships between the optimistic and pessimistic sequential three-way decision models as follows.

Proposition 5. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t}$ = $\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, if $0 < l < l'$, then the following properties hold.

(1) $\underline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_j^h\right) \subseteq \underline{\sum_{i=1}^{q} cg_i^{l',t\,\Delta}}\left(D_j^h\right)$;
(2) $\overline{\sum_{i=1}^{q} cg_i^{l,t\,\Delta}}\left(D_j^h\right) \supseteq \overline{\sum_{i=1}^{q} cg_i^{l',t\,\Delta}}\left(D_j^h\right)$.

Proof. According to Proposition 3, it is clear that Proposition 5 holds true. □.

Proposition 6. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t}$ = $\{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, we have the following properties.

(1) $POS^{\Delta}_{GS^{1,t}}\left(D^h_j\right) \subseteq POS^{\Delta}_{GS^{2,t}}\left(D^h_j\right) \subseteq \ldots \subseteq POS^{\Delta}_{GS^{m,t}}\left(D^h_j\right)$;
(2) $BND^{\Delta}_{GS^{1,t}}\left(D^h_j\right) \supseteq BND^{\Delta}_{GS^{2,t}}\left(D^h_j\right) \supseteq \ldots \supseteq BND^{\Delta}_{GS^{m,t}}\left(D^h_j\right)$;
(3) $NEG^{\Delta}_{GS^{1,t}}\left(D^h_j\right) \subseteq POS^{\Delta}_{GS^{2,t}}\left(D^h_j\right) \subseteq \ldots \subseteq POS^{\Delta}_{GS^{m,t}}\left(D^h_j\right)$;
(4) $POS^{\Delta}_{GS^{1,t}}\left(\pi^h_D\right) \subseteq POS^{\Delta}_{GS^{2,t}}\left(\pi^h_D\right) \subseteq \ldots \subseteq POS^{\Delta}_{GS^{m,t}}\left(\pi^h_D\right)$;
(5) $BND^{\Delta}_{GS^{1,t}}\left(\pi^h_D\right) \supseteq BND^{\Delta}_{GS^{2,t}}\left(\pi^h_D\right) \supseteq \ldots \supseteq BND^{\Delta}_{GS^{m,t}}\left(\pi^h_D\right)$.

Proof. According to Proposition 5, it is clear that Proposition 6 holds. □.

Proposition 5 shows that the higher the level of condition granules, the bigger the lower approximation and the smaller the upper approximation.

In what follows, we construct an algorithm to compute the three probabilistic regions of multigranulation sequential three-way decisions under different levels of multi-granularity as shown in Algorithm2. The main idea of Algorithm2 is to first delete the objects belonging to the positive region under the first level of granularity and then obtain the updated region $U^{2,t} = BND^{\triangle}_{GS^{1,t}}(\pi^h_D)$. For the updated region $U^{2,t}$, delete the objects belonging to the positive region at the next level of granularity, and repeat these steps until the updated universe becomes an empty set or no level of granularity can be computed. It is easy to observe that the time complexity of Algorithm2 is $O\left(m\sum_{i=1}^{q} cg_i|U|^2\right)$.

Algorithm2: Computing the regions under different multi-granulation sequential three-way decisions.

input : An universal set of objects, U; number of attributes ascending, t; a granule structures, $GS = \{cg^{l,t}_1, cg^{l,t}_2, ..., cg^{l,t}_q\}, l \in \{1, 2, ..., m\}$
output: Two regions, $POS^{\triangle}_{GS}(\pi^h_D)$ and $BND^{\triangle}_{GS}(\pi^h_D)$

1 $POS^{\triangle}_{GS}(\pi^h_D) = \emptyset$ and $BND^{\triangle}_{GS}(\pi^h_D) = U$;
2 $U^{1,t} = U$;
3 **for** $l \leftarrow 1$ **to** m **do**
4 if $U^{l,t} = \emptyset$ or $l > m$ turn to 11;
5 **for** $i \leftarrow 1$ **to** q **do**
6 Computing $POS^{\triangle}_{cg^{l,t}_i}(\pi^h_D)$;
7 $POS^{\triangle}_{GS}(\pi^h_D) = POS^{\triangle}_{GS}(\pi^h_D) \cup POS^{\triangle}_{cg^{l,t}_i}(\pi^h_D)$;
8 **end**
9 $BND^{\triangle}_{GS}(\pi^h_D) = U^{l,t} - POS^{\triangle}_{GS}(\pi^h_D)$;
10 $U^{l+1,t} = BND^{\triangle}_{GS}(\pi^h_D)$;
11 **end**
12 Output $POS^{\triangle}_{GS}(\pi^h_D)$, $BND^{\triangle}_{GS}(\pi^h_D)$;

In particular, the generalized hierarchical multigranulation model degenerates to the traditional multigranulation model when $t = 1$, and the generalized hierarchical multigranulation model has only one condition granule per level when t takes the maximum number of attributes.

Example 2. Consider the universe set of objects $U^{1,t} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$, $t = 2$, the decision partition $U/D^1 = \pi^1_D = \{\{x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_{11}\}, \{x_5, x_{10}, x_{12}\}, \{x_9\}\}$, and two granular structures $GS^{1,2} = \{cg^{1,2}_1, cg^{1,2}_2, cg^{1,2}_3\}$ and $GS^{2,2} = \{cg^{2,2}_1, cg^{2,2}_2, cg^{2,2}_3\}$.

(1) For the granular structure $GS^{1,2}$, the positive and boundary regions with respect to π^1_D are computed as
$POS^{O}_{GS^{1,2}}\left(\pi^1_D\right) = \{x_1, x_3, x_4, x_6, x_{11}\}$;
$BND^{O}_{GS^{1,2}}\left(\pi^1_D\right) = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$;
$POS^{P}_{GS^{1,2}}\left(\pi^1_D\right) = \varnothing$;
$BND^{P}_{GS^{1,2}}\left(\pi^1_D\right) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$..
(2) For the granular structure $GS^{2,2}$, the positive and boundary regions with respect to π^1_D are computed as
$POS^{O}_{GS^{2,2}}\left(\pi^1_D\right) = \{x_2, x_5, x_7, x_8, x_9, x_{10}, x_{12}\}$;
$BND^{O}_{GS^{l,2}}\left(\pi^1_D\right) = \varnothing$;
$POS^{P}_{GS^{2,2}}\left(\pi^1_D\right) = \{x_1, x_3, x_4, x_5, x_6, x_8, x_9, x_{11}\}$;
$BND^{P}_{GS^{l,2}}\left(\pi^1_D\right) = \{x_2, x_7, x_{10}, x_{12}\}$.

Therefore, we can have.
$POS^O_{GS}(\pi^1_D)$ = $POS^O_{GS^{1,2}}(\pi^1_D) \cup POS^O_{GS^{2,2}}(\pi^1_D)$ = $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$;
$BND^O_{GS}(\pi^1_D)$ = $BND^O_{GS^{1,2}}(\pi^1_D)$ - $POS^O_{GS^{2,2}}(\pi^1_D)$ = $\emptyset$;
$POS^P_{GS}(\pi^1_D)$ = $POS^P_{GS^{1,2}}(\pi^1_D) \cup POS^P_{GS^{2,2}}(\pi^1_D)$ = $\{x_1, x_3, x_4, x_5, x_6, x_8, x_9, x_{11}\}$;
$BND^P_{GS}(\pi^1_D)$ = $BND^P_{GS^{1,2}}(\pi^1_D)$ - $POS^P_{GS^{2,2}}(\pi^1_D)$ = $\{x_2, x_7, x_{10}, x_{12}\}$.

5. Multigranulation sequential three-way decision model under decision granule level descension for hierarchical classification

In this section, we consider the fixed levels of condition granules and dynamically change the levels of decision granules, and update the three probabilistic regions at different levels. The sequential decision-making process for hierarchical classification is shown in Fig. 4 (b) from level n to level 1. It is worth that the sequential process of decision granules is the reverse one of condition granules.

Definition 14. For a generalized hierarchical decision table GH, a decision class D^h_j and a multilevel granular structure $GS^{l,t}$ = $\{cg^{l,t}_1, cg^{l,t}_2, \ldots, cg^{l,t}_q\}$, then the lower and upper approximations $\underline{apr}_{cg^{l,t}_i}\left(D^h_j\right)$ and $\overline{apr}_{cg^{l,t}_i}\left(D^h_j\right)$ are defined by

$$\begin{aligned} \underline{apr}_{cg^{l,t}_i}\left(D^h_j\right) &= \left\{x \in U^{h,t} | [x]_{cg^{l,t}_i} \subseteq D^h_j\right\}, \\ \overline{apr}_{cg^{l,t}_i}\left(D^h_j\right) &= \left\{x \in U^{h,t} | [x]_{cg^{l,t}_i} \cap D^h_j \neq \emptyset\right\}. \end{aligned} \tag{23}$$

where $U^{n,t} = U, U^{h,t} = \overline{apr}_{cg^{l,t}_i}\left(D^{h+1}_j\right) - \underline{apr}_{cg^{l,t}_i}\left(D^{h+1}_j\right)$ is the gradually reduced universe, and n denotes the maximum level of D_j.

Proposition 7. Given a generalized hierarchical decision table GH, the updated three probabilistic regions for the condition granule cg with respect to D hold true:

(1)

$$\begin{aligned} POS_{cg^{l,t}_i}\left(D^h_j\right) &= \underline{apr}_{cg^{l,t}_i}\left(D^h_j\right) \\ &= POS_{cg^{l,t}_i}\left(D^{h+1}_j\right) \cup \left\{x | [x]_{cg^{l,t}_i} \subseteq D^h_j, x \in BND_{cg^{l,t}_i}\left(D^{h+1}_j\right)\right\}; \end{aligned} \tag{24}$$

(2)

$$\begin{aligned} BND_{cg^{l,t}_i}\left(D^h_j\right) &= \overline{apr}_{cg^{l,t}_i}\left(D^h_j\right) - \underline{apr}_{cg^{l,t}_i}\left(D^h_j\right) \\ &= U^{h,t} - POS_{cg^{l,t}_i}\left(D^h_j\right) - NEG_{cg^{l,t}_i}\left(D^h_j\right); \end{aligned} \tag{25}$$

(3)

$$\begin{aligned} NEG_{cg^{l,t}_i}\left(D^h_j\right) &= U^{h,t} - \overline{apr}_{cg^{l+1,t}_i}\left(D^h_j\right) \\ &= NEG_{cg^{l,t}_i}\left(D^{h+1}_j\right) \cup \left\{x | [x]_{cg^{l+1,t}_i} \cap D^h_j = \emptyset, x \in BND_{cg^{l,t}_i}\left(D^{h+1}_j\right)\right\}. \end{aligned} \tag{26}$$

Proof. The equivalence class D^h_j of the h-th level will be further divided into the equivalence class D^{h+1}_j of the $(h+1)$-th level. We easily get $D^{h+1}_j \subseteq D^h_j$.

(1) For any $x \in POS_{cg^{l,t}_i}\left(D^{h+1}_j\right)$, we have $[x]_{cg^{l,t}_i} \subseteq D^{h+1}_j$, then $[x]_{cg^{l,t}_i} \subseteq D^h_j$ since $D^{h+1}_j \subseteq D^h_j$. Thus, $x \in POS_{cg^{l+1,t}_i}\left(D^h_j\right)$.
(2) For any $x \in BND_{cg^{l,t}_i}\left(D^h_j\right)$, it is easy to obtain $[x]_{cg^{l,t}_i} \cap D^h_j \neq \emptyset$ and $[x]_{cg^{l,t}_i} \neg \subseteq D^h_j$, so, $x \notin NEG_{cg^{l,t}_i}\left(D^h_j\right)$ and $x \notin POS_{cg^{l,t}_i}\left(D^h_j\right)$, because of $D^{h+1}_j \subseteq D^h_j$. Therefore, $x \in BND_{cg^{l,t}_i}\left(D^{h+1}_j\right)$.
(3) For any $x \in NEG_{cg^{l,t}_i}\left(D^{h+1}_j\right)$, we have $[x]_{cg^{l,t}_i} \subseteq D^{h+1}_j$, then $[x]_{cg^{l+1,t}_i} \subseteq D^h_j$ because of $D^{h+1}_j \subseteq D^h_j$. Thus, $x \in NEG_{cg^{l,t}_i}\left(D^h_j\right)$. □

Proposition 8. Given a generalized hierarchical decision table $GH, POS_{cg^{l,t}_i}\left(D^h_j\right)$ and $NEG_{cg^{l,t}_i}\left(D^h_j\right)$ monotonically become larger and $BND_{cg^{l,t}_i}\left(D^h_j\right)$ monotonically turns smaller with the decreasing of the level of decision granule.

Proof. From Proposition 7, we obtain the $POS_{cg_i^{l,t}}\left(D_j^{h+1}\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^{h+1}\right)$ are parts of the $POS_{cg_i^{l,t}}\left(D_j^{h}\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^{h}\right)$, and the added ones are selected from $BND_{cg_i^{l,t}}\left(D_j^{h+1}\right)$. Thus, $POS_{cg_i^{l,t}}\left(D_j^{h}\right) \supseteq POS_{cg_i^{l,t}}\left(D_j^{h+1}\right)$ and $NEG_{cg_i^{l,t}}\left(D_j^{h}\right) \supseteq NEG_{cg_i^{l,t}}\left(D_j^{h+1}\right)$. Since $BND_{cg_i^{l,t}}\left(D_j^{h}\right) = D_j - POS_{cg_i^{l,t}}\left(D_j^{h}\right) - NEG_{cg_i^{l,t}}\left(D_j^{h}\right)$ and $BND_{cg_i^{l,t}}\left(D_j^{h+1}\right) = D_j - POS_{cg_i^{l,t}}\left(D_j^{h+1}\right) - NEG_{cg_i^{l,t}}\left(D_j^{h+1}\right)$, then $BND_{cg_i^{l,t}}\left(D_j^{h}\right) \subseteq BND_{cg_i^{l,t}}\left(D_j^{h+1}\right)$. □.

Definition 15. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, then the lower and upper approximations of the optimistic generalized hierarchical multi-granulation sequential three-way decision with respect to D_j^h are defined as

$$\begin{aligned}\underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) &= \left\{x \in U^{h,t} | [x]_{cg_1^{l,t}} \subseteq D_j^h \vee [x]_{cg_2^{l,t}} \subseteq D_j^h \vee \ldots \vee [x]_{cg_q^{l,t}} \subseteq D_j^h\right\},\\ \overline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) &= \sim \underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(\sim D_j^h\right).\end{aligned} \tag{27}$$

where $U^{n,t} = U, U^{h-1,t} = \overline{\sum_{i=1}^{m} cg_i^{l,t\,O}}\left(D_j^h\right) - \underline{\sum_{i=1}^{m} cg_i^{l,t\,O}}\left(D_j^h\right)$ is the gradually reduced universe, q represents the maximum number of parallel condition granule sets, and $[x]_{cg_i^{l,t}} (1 \leqslant i \leqslant q)$ represents the equivalence class containing x in the partition $U^{h,t}/cg_i^{l,t}$.

The pair $< \underline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right), \overline{\sum_{i=1}^{q} cg_i^{l,t\,O}}\left(D_j^h\right) >$is called the optimistic generalized hierarchical rough set of D_j^h with respect to the set of condition granules cg in $U^{h,t}$.

Definition 16. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, then the lower and upper approximations of the pessimistic multi-granulation generalized hierarchical decision with respect to D_j^h are defined as

$$\begin{aligned}\underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) &= \left\{x \in U^{h,t} | [x]_{cg_1^{l,t}} \subseteq D_j^h \wedge [x]_{cg_2^{l,t}} \subseteq D_j^h \wedge \ldots \wedge [x]_{cg_q^{l,t}} \subseteq D_j^h\right\},\\ \overline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) &= \sim \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(\sim D_j^h\right).\end{aligned} \tag{28}$$

where $U^{n,t} = U, U^{h-1} = \overline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right) - \underline{\sum_{i=1}^{q} cg_i^{l,t\,P}}\left(D_j^h\right)$ is the gradually reduced universe, q represents the maximal number of parallel condition granule sets, and $[x]_{cg_i^{l,t}} (1 \leqslant i \leqslant q)$ denotes the equivalence class containing x in the partition $U^{h,t}/cg_i^{l,t}$.

Proposition 9. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, if $\forall D_j^{h'} \subseteq D_j^h \subseteq U$, then

$$\underline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}\left(D_j^h\right) \supseteq \underline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}\left(D_j^{h'}\right), \overline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}\left(D_j^h\right) \subseteq \overline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}\left(D_j^{h'}\right); \tag{29}$$

$$\underline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}(\pi_D^h) \supseteq \underline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}(\pi_D^{h'}), \overline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}(\pi_D^h) \subseteq \overline{\sum_{i=1}^{q} cg_i^{l,t\Delta}}(\pi_D^{h'}). \tag{30}$$

Proof. According to Proposition 7, it obviously holds true. □.

Proposition 9 shows that the higher the levels of decision granules, the smaller the lower approximation while the bigger the upper approximation.

Proposition 10. Given a generalized hierarchical decision table GH, a decision class D_j^h and a multilevel granular structure $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, \ldots, cg_q^{l,t}\}$, then the following properties hold:

(1) $POS_{GS^{l,t}}^{\Delta}\left(D_j^n\right) \subseteq POS_{GS^{l,t}}^{\Delta}\left(D_j^{n-1}\right) \subseteq \ldots \subseteq POS_{GS^{l,t}}^{\Delta}\left(D_j^1\right)$;

(2) $BND^{\Delta}_{GS^{l,t}}\left(D_j^n\right) \supseteq BND^{\Delta}_{GS^{l,t}}\left(D_j^{n-1}\right) \supseteq \ldots \supseteq BND^{\Delta}_{GS^{l,t}}\left(D_j^1\right)$;
(3) $NEG^{\Delta}_{GS^{l,t}}\left(D_j^n\right) \subseteq NEG^{\Delta}_{GS^{l,t}}\left(D_j^{n-1}\right) \subseteq \ldots \subseteq NEG^{\Delta}_{GS^{l,t}}\left(D_j^1\right)$;
(4) $POS^{\Delta}_{GS^{l,t}}(\pi_D^n) \subseteq POS^{\Delta}_{GS^{l,t}}(\pi_D^{n-1}) \subseteq \ldots \subseteq POS^{\Delta}_{GS^{l,t}}(\pi_D^1)$;
(5) $BND^{\Delta}_{GS^{l,t}}(\pi_D^n) \supseteq BND^{\Delta}_{GS^{l,t}}(\pi_D^{n-1}) \supseteq \ldots \supseteq BND^{\Delta}_{GS^{l,t}}(\pi_D^1)$.

Proof. According to Proposition 9, it is clear that Proposition 10 holds true. □.

In what follows, we construct an algorithm to compute the three probabilistic regions of multigranulation sequential three-way decisions under different levels of multi-granularity as shown in Algorithm 3. The main idea of Algorithm 3 is to first delete the objects belonging to the positive region under the maximum level of granularity and then obtain the updated region $U^{n-1,t} = BND^{\triangle}_{GS^{l,t}}(\pi_D^n)$. For the updated region $U^{n-1,t}$, delete the objects belonging to the positive region at the next level of granularity, and repeat these steps until the updated universe becomes an empty set or no level of granularity can be computed. It is easy to observe that the time complexity of Algorithm 3 is $O\left(n\sum_{i=1}^{q} cg_i|U|^2\right)$.

Algorithm 3: Computing the regions of decision granule level descension under different multi-granulation sequential three-way decisions.

input : An universal set of objects, U; number of attributes ascending, t; a granule structures, $GS^{l,t} = \{cg_1^{l,t}, cg_2^{l,t}, ..., cg_q^{l,t}\}, l \in \{1, 2, ..., m\}$
output: Two regions, $POS^{\triangle}_{GS^{l,t}}(\pi_D)$ and $BND^{\triangle}_{GS^{l,t}}(\pi_D)$

```
1  POS△_{GS^{l,t}}(π_D) = ∅ and BND△_{GS^{l,t}}(π_D) = U;
2  U^{n,t} = U;
3  for h ← n to 1 do
4      if U^{h,t} = ∅ or h < n turn to 11;
5      for i ← 1 to q do
6          Computing POS△_{cg_i^{l,t}}(π_D^h);
7          POS△_{GS^{l,t}}(π_D) = POS△_{GS^{l,t}}(π_D) ∪ POS△_{cg_i^{l,t}}(π_D^h);
8      end
9      BND△_{GS^{l,t}}(π_D) = U^{h,t} − POS△_{GS^{l,t}}(π_D);
10     U^{h−1,t} = BND△_{GS^{l,t}}(π_D);
11 end
12 Output POS△_{GS^{l,t}}(π_D), BND△_{GS^{l,t}}(π_D);
```

Example 3. Consider a universe set of objects $U^{2,t} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$, t = 2, $U/D^1 = \pi_D^1 = \{\{x_1, x_2, x_3, x_4, x_6, x_7, x_8, x_{11}\},\{x_5, x_{10}, x_{12}\}, \{x_9\}\}$, $U/D^2 = \pi_D^2 = \{\{x_1, x_4, x_8\}, \{x_2, x_3, x_6, x_7, x_{11}\}, \{x_{10}, x_{12}\}, \{x_5\}, \{x_9\}\}$, and a granular structure $GS^{2,2} = \{cg_1^{2,2}, cg_2^{2,2}, cg_3^{2,2}\}$.

(1) For the granular structure $GS^{2,2}$, the positive and boundary regions with respect to π_D^2 are computed as
$POS^{O}_{GS^{2,2}}\left(\pi_D^2\right) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$, $BND^{O}_{GS^{2,2}}\left(\pi_D^2\right) = \varnothing$;
$POS^{P}_{GS^{2,2}}\left(\pi_D^2\right) = \{x_3, x_4, x_8, x_9, x_{11}\}$; $BND^{P}_{GS^{2,2}}\left(\pi_D^2\right) = \{x_1, x_2, x_5, x_6, x_7, x_{10}, x_{12}\}$.
(2) For the granular structure $GS^{2,2}$, the positive and boundary regions with respect to π_D^1 are computed as.
$POS^{O}_{GS^{2,2}}\left(\pi_D^1\right) = \varnothing$;
$BND^{O}_{GS^{2,2}}\left(\pi_D^1\right) = \varnothing$;
$POS^{P}_{GS^{2,2}}\left(\pi_D^1\right) = \{x_1, x_5, x_6\}$;
$BND^{P}_{GS^{2,2}}\left(\pi_D^1\right) = \{x_2, x_7, x_{10}, x_{12}\}$.
Therefore, we can have $POS^{O}_{GS^{2,2}}\left(\pi_D^h\right) = POS^{O}_{GS^{2,2}}\left(\pi_D^2\right) \cup POS^{O}_{GS^{2,2}}\left(\pi_D^1\right) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$, $BND^{O}_{GS^{2,2}}\left(\pi_D^h\right) = BND^{O}_{GS^{2,2}}\left(\pi_D^2\right) - POS^{O}_{GS^{2,2}}\left(\pi_D^1\right) = \varnothing$, $POS^{P}_{GS^{2,2}}\left(\pi_D^h\right) = POS^{P}_{GS^{2,2}}\left(\pi_D^2\right) \cup POS^{P}_{GS^{2,2}}\left(\pi_D^1\right) = \{x_1, x_3, x_4, x_5, x_6, x_8, x_9, x_{11}\}$, $BND^{P}_{GS^{2,2}}\left(\pi_D^h\right) = BND^{P}_{GS^{2,2}}\left(\pi_D^2\right) - POS^{P}_{GS^{2,2}}\left(\pi_D^1\right) = \{x_2, x_7, x_{10}, x_{12}\}$.

6. Experiments and analysis

6.1. Data sets

In order to evaluate our models, we perform some experiments on a personal computer with windows 10, 1.8 GHz CPU and 8 GB memory. The software is IntelliJ idea 2017.3. The following experiments mainly compare the size of regions under different levels of granularity. For convenience, we abbreviate the optimistic and pessimistic generalized hierarchical multi-granulation sequential three-way decisions as OGHMS3WD and PGHMS3WD, respectively. The characteristics of the six datasets are described in Table 3. The numbers of decision granules in the dataset are shown in Table 4.

It is worth mentioning that we need to preprocess the dataset. We delete the 1st, 18th to 24th and 26th to 28th attributes in Fars because these are also not relevant to the fatal accident results. Then, we use Rosetta software (http://www.lcb.uu.se/tools/rosetta/) to convert the continuous data to discrete values. Finally, we construct the concept hierarchy tree and stratify the experimental data by general social cognition (some information from Baidu Encyclopedia).

6.2. Comparison of the positive regions under different levels of granularity

In what follows, we compute the number of the positive regions and analyze the uncertainty of the boundary regions of the generalized hierarchical sequential three-way decisions. For convenience, we use *OClDh* and *PClDh* ($l = 1, 2, 3; h = 1, 2, 3$) to denote the optimistic and pessimistic strategies to select the condition granules at the l-th level and the decision granules at the h-th level, respectively. Due to space limitations, we only show the change of the number of the positive regions at the second level of condition granules and decision granules. Figs. 5 and 6 illustrate the changes of the numbers of the positive regions. By observing the numbers of the positive regions from different hierarchical sequential three-way decision models, we can conclude as follows:

•; The positive regions will become enlarge as the levels of condition granules increase, indicating that the probabilistic positive region is monotonic.

•; The positive regions grow with the decreasing level of decision granules, indicating that the generalization of decision granules are conducive to the division of positive regions.

•; The positive regions increase monotonically with the ascending number of attributes.

6.3. Comparison of uncertainty of the boundary regions under different levels of granularity

We employ the deferment rate to evaluate the quality of the probabilistic boundary regions under different sequential three-way decisions as follows:

$$DR^{l,h,t} = \frac{|BND^{\Delta}_{GS^{l,t}}(\pi^h_D)|}{|U|}. \tag{31}$$

Table 3
Description of the datasets.

No.	Dataset	$\|U\|$	$\|C\|$	$\|V_d\|$
1	Abalone	4177	8	28
2	Deal winequality red	1599	11	6
3	Fars	100968	14	8
4	Glass	214	9	6
5	Marketing	6876	6	9
6	Obesity	2111	16	7

Table 4
Description of the decision granules.

No.	Datasets	Attributes	Decision attributes under different levels		
			Level-1	Level-2	Level-3
1	Abalone	28	5	9	28
2	Deal winequality red	6	2	4	6
3	Fars	8	5	6	8
4	Glass	6	2	3	6
5	Marketing	9	2	3	9
6	Obesity	7	3	4	7

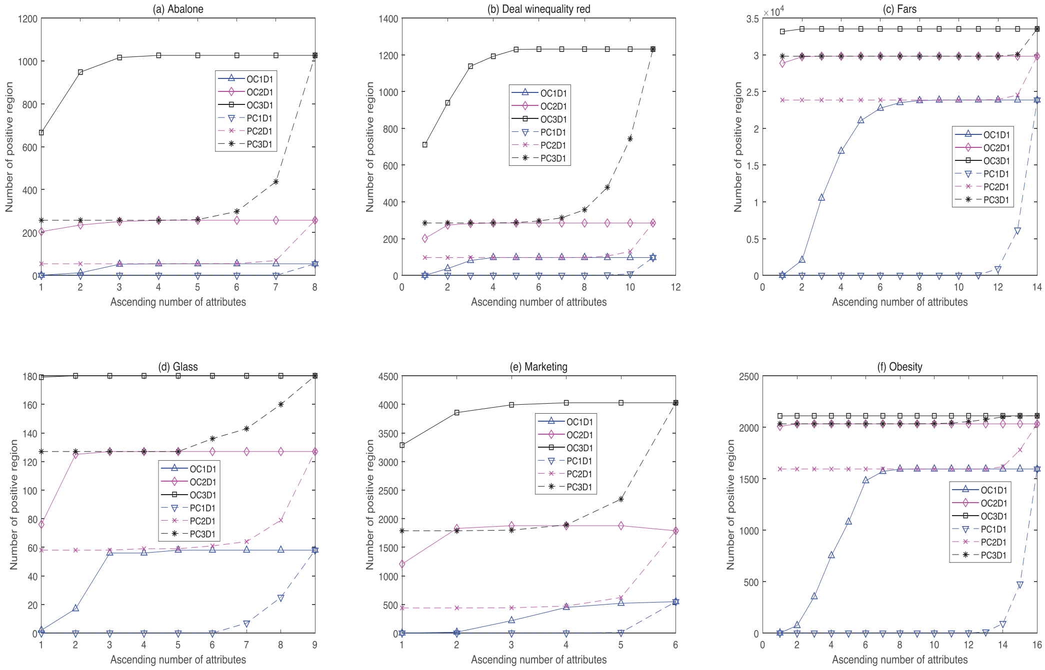


Fig. 5. Positive regions under different levels of condition granules and the second level of decision granules.

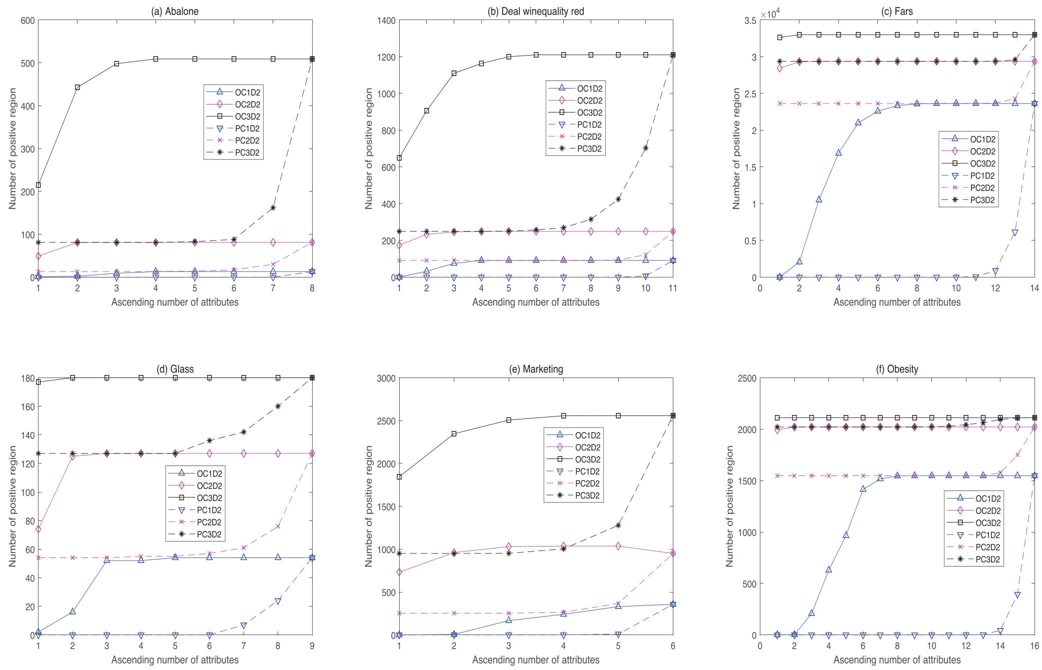


Fig. 6. Positive regions under different levels of decision granules and the second level of condition granules.

The experiment results of variation of deferment rate on six data sets under different levels of condition granules and decision granules are shown in Figs. 7 and 8. From the two figures, we can observe the changes of deferment rates as follows.

•; The boundary region reduces monotonically with the increasing levels of condition granules or decreasing levels of decision granules.

•; The boundary region reduces monotonically with the ascending number of attributes.

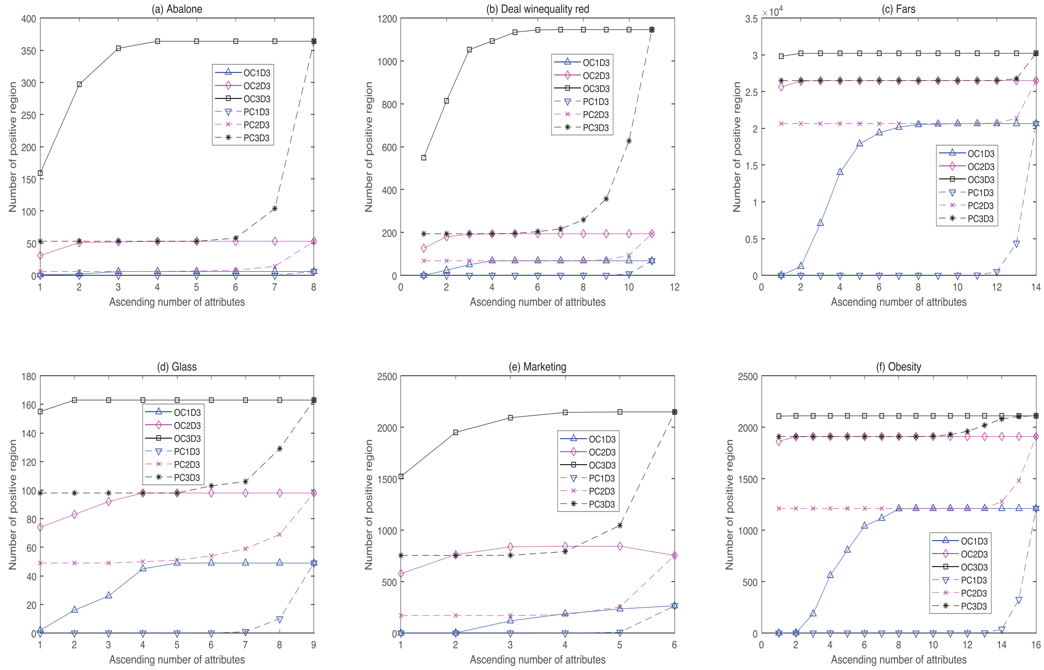


Fig. 7. Uncertainty of the boundary regions under different levels of condition granules and the second level of decision granules.

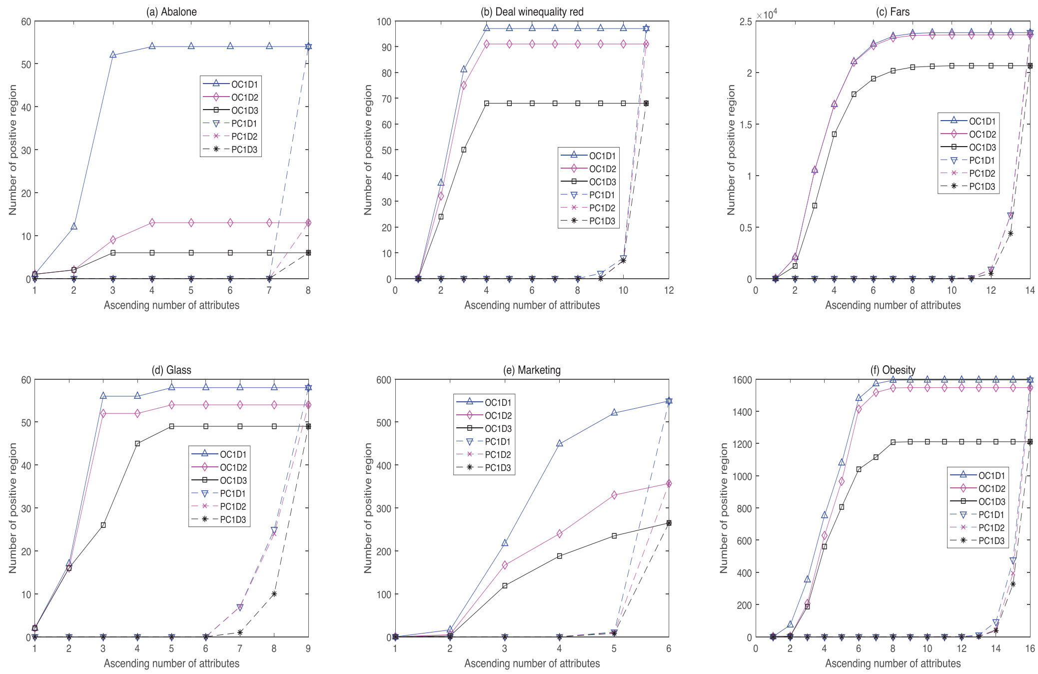


Fig. 8. Uncertainty of the boundary regions under different levels of decision granules and the second level of condition granules.

6.4. Comparison of the positive and boundary regions under different levels of granularity

In this subsection, we mainly compare the number of positive regions under different strategies and levels of granularity. For convenience, ODGl and PDGl denote the decision granules at level l under the optimistic strategy and the pessimistic strategy, respectively. Figs. 9 and 10 illustrate the changes of the positive and boundary regions under different levels of condition granules and decision granules using optimistic and pessimistic strategies, where the subgraphs (I) → (II) → (III)

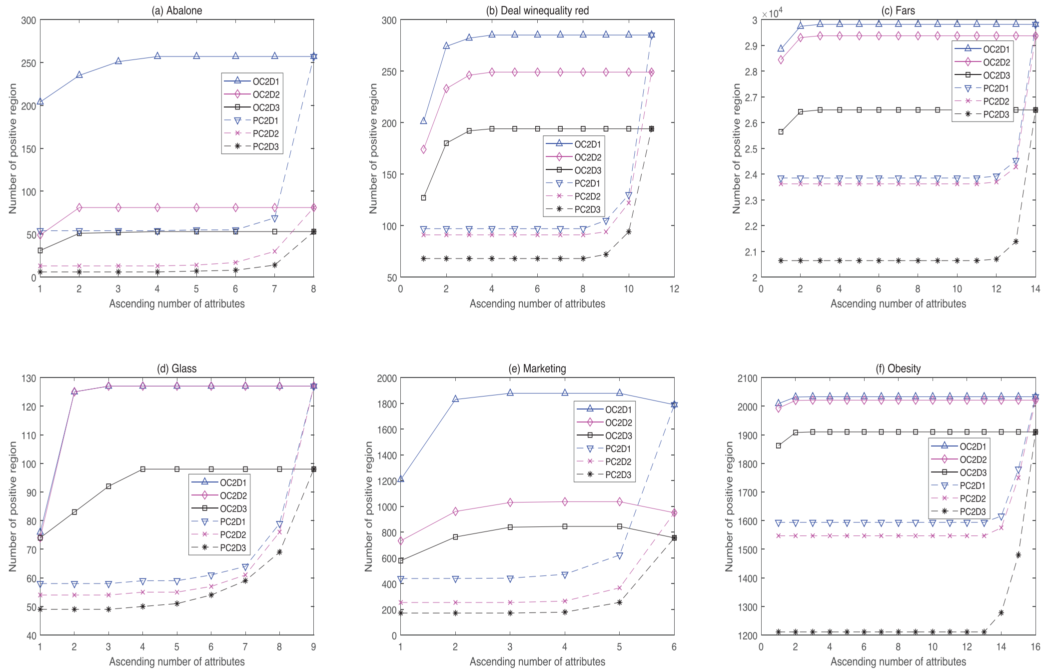


Fig. 9. OGHMS3WD under different levels of condition granules on Marketing.

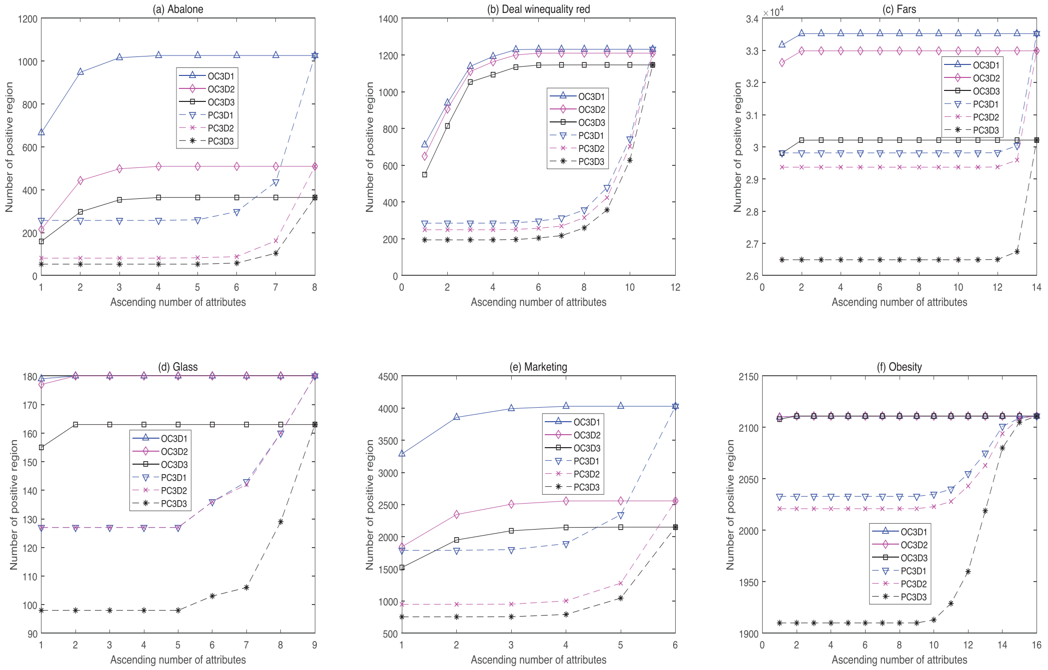


Fig. 10. PGHMS3WD under different levels of condition granules on Marketing.

denote the cases of the first, second and third levels of condition granules under the same level of decision granules, respectively. One can observe that the higher the condition granule levels the larger the size of the positive regions. Figs. 11 and 12 illustrate the changes of the positive and boundary regions under different levels of decision granules using different strategies as well, where the subgraphs (I) → (II) → (III) represent the cases of the third, second and first levels of decision granules under the same level of condition granules, respectively. Further, one can find that the positive regions increase as the levels

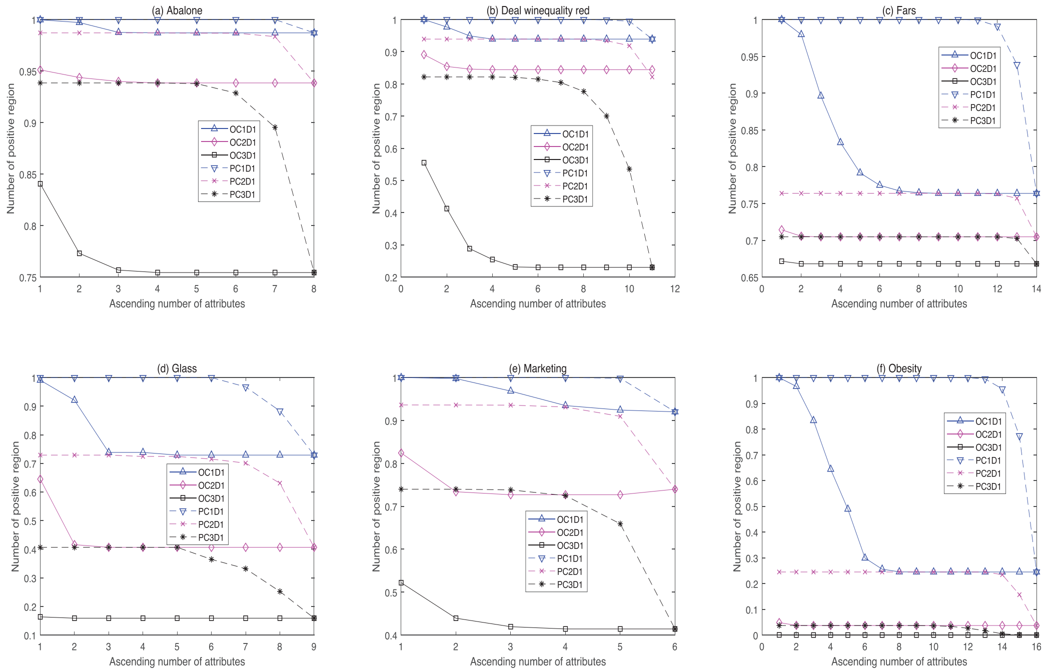


Fig. 11. OGHMS3WD under different levels of decision granules on Marketing.

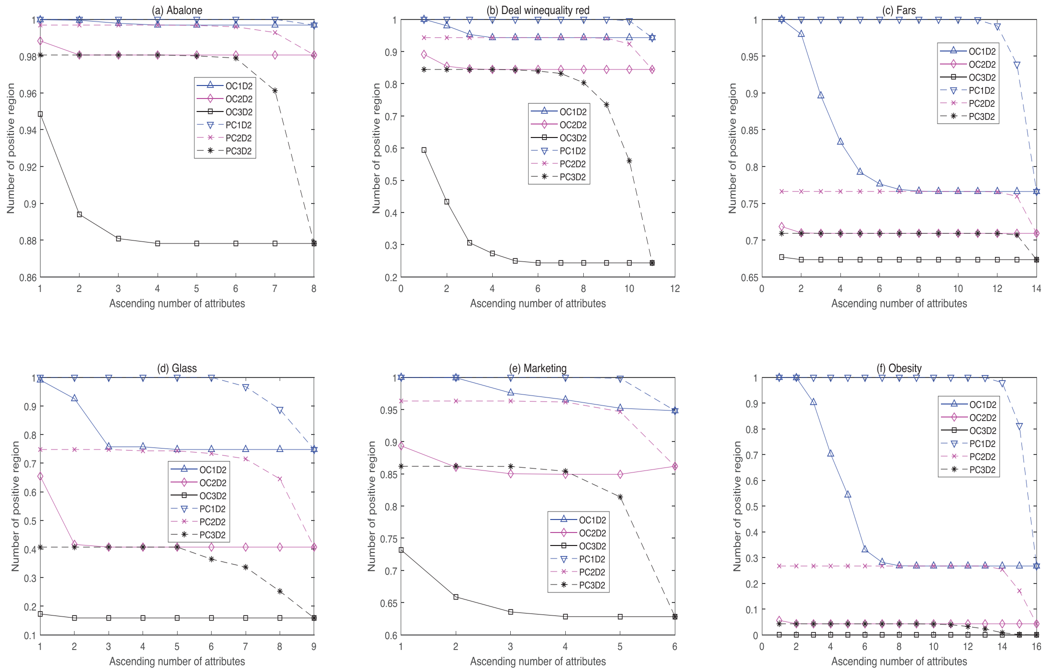


Fig. 12. PGHMS3WD under different levels of decision granules on Marketing.

of decision granules decreases. For Marketing, we construct six condition granules based on the number of conditional attributes and find that the positive region increases with the ascending number of attributes, which is because the more the number of attributes, and the more the number of objects induced by the positive regions.

7. Conclusions

In this paper, we propose a generalized hierarchical multi-granulation sequential three-way decision model for hierarchical classification. With this framework, we design the condition (attribute) granules and decision granules at different levels of granularity, integrate the generalized hierarchical decision tables and multigranulation rough set into sequential three-way decisions, and further analyze the properties and relationships of the three probabilistic regions under different multi-granulation sequential three-way decisions. It provides a novel insight from the perspective of multi-level and multi-view for the existing models.

In the future work, we will focus on the uncertainty and attribute reduction of this generalized hierarchical multigranulation sequential three-way decision model.

CRediT authorship contribution statement

Jin Qian: Conceptualization, Methodology, Writing - review & editing, Validation, Supervision. **Chengxin Hong:** Conceptualization, Methodology, Writing - original draft, Software, Validation. **Ying Yu:** Writing - review & editing. **Caihui Liu:** Writing - review & editing. **Duoqian Miao:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research is supported by the National Natural Science Foundation of China under Grant Nos. 62066014, 62163016, 62166001, 61976158, Double thousand plan of Jiangxi Province of China, Jiangxi Province Natural Science Foundation of China under Grant Nos.20202BABL202018, 20202BAB202010, 20212ACB202001.

References

[1] Y.Y. Yao, Three-way decisions with probabilitic rough sets, Inf. Sci. 180 (2010) 341–353.
[2] Y.Y. Yao, The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. 181 (2011) 1080–1096.
[3] Y.Y. Yao, Tri-level thinking: models of three-way decision, Int. J. Mach. Learn. Cybern. 11 (2020) 947–959.
[4] Y.Y. Yao, The geometry of three-way decision, Appl. Intell. 51 (2021) 6298–6325.
[5] Y.Y. Yao, Granular computing and sequential three-way decisions, in: Proceedings of International Conference on Rough Sets and Knowledge Technology, Springer, 2013.
[6] H.X. Li, L.B. Zhang, B. Huang, X.Z. Zhou, Sequential three-way decision and granulation for cost-sensitive face recognition, Knowl-Based Syst. 91 (2016) 241–251.
[7] X.Y. Zhang, Y.Y. Yao, Tri-level attribute reduction in rough set theory, Expert Syst. Appl. 190 (2022) 116187.
[8] X.Y. Zhang, J.F. Jiang, Measurement, modeling, reduction of decision-theoretic multigranulation fuzzy rough sets based on three-way decisions, Inf. Sci. 607 (2022) 1550–1582.
[9] J.M. Zhan, H.B. Jiang, Y.Y. Yao, Three-Way Multiattribute Decision-Making Based on Outranking Relations, IEEE Trans. Fuzzy Syst. 29 (2021) 2844–2858.
[10] D.C. Liang, M.W. Wang, Z.S. Xu, D. Liu, Risk appetite dual hesitant fuzzy three-way decisions with TODIM, Inf. Sci. 507 (2020) 585–605.
[11] X.D. Yue, Y.F. Chen, B. Yuan, Y. Lv, Three-way image classification with evidential deep convolutional neural networks, Cognit. Comput. 1 (2021) 1–13.
[12] X. Yang, Y.J. Li, Q.K. Li, D. Liu, T.R. Li, Temporal-spatial three-way granular computing for dynamic text sentiment classification, Inf. Sci. 596 (2022) 551–566.
[13] H.R. Zhang, F. Min, B. Shi, Regression-based three-way recommendation, Inf. Sci. 378 (2017) 444–461.
[14] X.Q. Ye, D. Liu, A cost-sensitive temporal-spatial three-way recommendation with multi-granularity decision, Inf. Sci. 589 (2022) 670–689.
[15] Y.Y. Yao, Stratified rough sets and granular computing, in: 18th International Conference of the North American Fuzzy Information Processing Society - NAFIPS (Cat. No.99TH8397), 1999, pp. 800–804.
[16] Y.Y. Yao, Three-way decision and granular computing, Int. J. Approx. Reason. 103 (2018) 107–123.
[17] D.Q. Miao, G.Y. Wang, Q. Liu, e. al., Granular Computing Past Nowday and Future, Science Publisher, Beijing, 2007.
[18] A. Bargiela, W. Pedrycz, Toward a theory of granular computing for human-centered information processing, IEEE Trans. Fuzzy Syst. 16 (2008) 320–330.
[19] L.A. Zadeh, Fuzzy sets and information granularity, in: M. Gupta, R. Ragade, R. Yager (Eds.), Advances in Fuzzy Set Theory and Applications, North-Holland Publishing, Amsterdam, 1979, pp. 3–18.
[20] H.F. de Arruda, F.N. Silva, L.D. Costa, D.R. Amancio, Knowledge acquisition: A Complex networks approach, Inf. Sci. 421 (2017) 154–166.
[21] J.H. Li, C.L. Mei, W.H. Xu, Y.H. Qian, Concept learning via granular computing: A cognitive viewpoint, Inf. Sci. 298 (2015) 447–467.
[22] S.F. Ding, H.J. Huang, J.Z. Yu, H. Zhao, Research on the hybrid models of granular computing and support vector machine, Artif. Intell. Rev. 43 (2015) 565–577.
[23] H.Q. Truong, L.T. Ngo, W. Pedrycz, Granular Fuzzy Possibilistic C-Means Clustering approach to DNA microarray problem, Knowl-Based Syst. 133 (2017) 53–65.
[24] H. Fujita, A. Gaeta, V. Loia, F. Orciuoli, Resilience Analysis of Critical Infrastructures: A Cognitive Approach Based on Granular Computing, IEEE T. Cybern. 49 (2019) 1835–1848.
[25] X.B. Zhu, W. Pedrycz, Z.W. Li, A Development of Granular Input Space in System Modeling, IEEE T. Cybern. 51 (2021) 1639–1650.
[26] S.D. Hu, D.Q. Miao, W. Pedrycz, Multi granularity based label propagation with active learning for semi-supervised classification, Expert Syst. Appl. 192 (2022) 116276.
[27] Y.Y. Yao, Granular computing using neighborhood systems, in: Advances in Soft Computing, Springer-Verlag, London, 1999, pp. 539–553.
[28] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: A multi-granulation rough set, Inf. Sci. 180 (2010) 949–970.
[29] Q.R. Feng, D.Q. Miao, Y. Cheng, Hierarchical decision rules mining, Expert Syst. Appl. 37 (2010) 2081–2091.

[30] J. Qian, P. Lv, X.D. Yue, C.H. Liu, Z.J. Jing, Hierarchical attribute reduction algorithms for big data using MapReduce, Knowl-Based Syst. 73 (2015) 18–31.
[31] W.Z. Wu, Y. Leung, Theory and applications of granular labelled partitions in multi-scale decision tables, Inf. Sci. 181 (2011) 3878–3897.
[32] W.Z. Wu, Y. Leung, Optimal scale selection for multi-scale decision tables, Int. J. Approx. Reason. 54 (2013) 1107–1129.
[33] Z.H. Huang, J.J. Li, W.Z. Dai, R.D. Lin, Generalized multi-scale decision tables with multi-scale decision attributes, Int. J. Approx. Reason. 115 (2019) 194–208.
[34] C. Hao, J.H. Li, M. Fan, W.Q. Liu, E.C.C. Tsang, Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions, Inf. Sci. 415 (2017) 213–232.
[35] J. Qian, C.H. Liu, D.Q. Miao, X.D. Yue, Sequential three-way decisions via multi-granularity, Inf. Sci. 507 (2020) 606–629.
[36] J. Qian, D.W. Tang, Y. Yu, X.B. Yang, S. Gao, Hierarchical sequential three-way decision model, Int. J. Approx. Reason. 140 (2022) 156–172.
[37] W. Shen, Z.H. Wei, Q.W. Li, H.Y. Zhang, D.Q. Miao, Three-way decisions based blocking reduction models in hierarchical classification, Inf. Sci. 523 (2020) 63–76.
[38] S. Guo, H. Zhao, Hierarchical classification with multi-path selection based on granular computing, Artif. Intell. Rev. 54 (2021) 2067–2089.
[39] Y. Wang, Q.H. Hu, H. Chen, Y.H. Qian, Uncertainty instructed multi-granularity decision for large-scale hierarchical classification, Inf. Sci. 586 (2022) 644–661.
[40] S. Li, J. Yang, G.Y. Wang, Q.H. Zhang, J. Hu, Granularity Selection for Hierarchical Classification Based on Uncertainty Measure, IEEE Transactions on Fuzzy Systems, doi: 10.1109/TFUZZ.2022.3161747.
[41] Y.Y. Yao, X.F. Deng, Sequential three-way decisions with probabilistic rough sets, in: Proceedings of the 10th IEEE International Conference on Cognitive Informatics and Cognitive Computing, 2011, pp. 120–125.
[42] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (1982) 341–356.
[43] X.B. Yang, Y.H. Qian, J.Y. Yang, Hierarchical Structures on Multigranulation Spaces, J. Comput. Sci. Technol. 27 (2012) 1169–1183.
[44] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (2014) 225–237.
[45] Y.Y. Yao, Y.H. She, Rough set models in multigranulation spaces, Inf. Sci. 327 (2016) 40–56.
[46] Y.Y. Yao, Rough sets and three-way decisions, in: Proceedings of the 10th International Conference on Rough Sets and Knowledge Technology (RSKT2015), Tianjin, China, 2015, pp. 62–73.