Knowledge-Based Systems 217 (2021) 106799

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Double-quantitative distance measurement and classification learning based on the tri-level granular structure of neighborhood system

Xianyong Zhang [a,*], Hongyuan Gou [a], Zhiying Lv [b], Duoqian Miao [c]

[a] School of Mathematical Sciences, Sichuan Normal University, Chengdu 610066, China
[b] School of Management, Chengdu University of Information Technology, Chengdu 610225, China
[c] Department of Computer Science and Technology, Tongji University, Shanghai 201804, China

ARTICLE INFO

Article history:
Received 21 March 2020
Received in revised form 18 January 2021
Accepted 19 January 2021
Available online 23 January 2021

Keywords:
Neighborhood rough sets
Granular computing
Tri-level granular structure
Double quantification
Distance measurement
Machine learning

ABSTRACT

In terms of neighborhood rough sets, the tri-level granular structure of neighborhood system (carrying the neighborhood granule, swarm, and library) establishes a granular computing mechanism for knowledge-based learning. However, its hierarchical exploration is inadequate, while its measurement can be extended for robust applications. Regarding this tri-level granular structure, the double-quantification technology is novelly introduced to make a thorough investigation, especially on the double-quantitative distance measurement and classification learning. Firstly, the size valuation and logical operation are hierarchically supplemented at higher levels. Secondly, the relative and absolute distances of bottom neighborhood granules are linearly combined to a double-quantitative distance, and all the three types of distances are promoted to both the middle swarm level and the top library level. Finally, the double-quantitative distance powerfully characterizing the difference of neighborhood granules is utilized to generate a double-quantitative classifier KNGD, and relevant data experiments show that this new classifier outperforms or balances two existing classifiers, i.e., the relative classifier KNGR and absolute classifier KNGA. By theory, example, and experiment, this study hierarchically perfects the tri-level granular structure of neighborhood system, and the corresponding double-quantification integration and extension offer the robust knowledge measurement and effective classification learning.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Motivation from tri-level granular structure of neighborhood system

Rough sets effectively implement knowledge-based reasoning and learning. Classical rough sets consider only the partition-granulation knowledge to mainly process symbolic data with discrete values [1], and thus the relevant utilization of continuous values in information system needs the discretization pretreatment, which is usually accompanied with information loss and effect decrease. To overcome the limitation, neighborhood rough sets introduce the distance measurement and radius parameter to flexibly utilize the covering-granulation knowledge [2,3], and thus they can robustly deal with numeric and even hybrid data with continuous values. Nowadays, neighborhood rough sets have been extensively applied for classification and clustering [4–9], feature selection [10–13], attribute reduction [14–18], gene selection [19,20], and outlier detection [21] etc.

Granular computing is a methodology of cognitive computing and uncertainty analysis by virtue of granulation processing and hierarchy solving [22,23], and it involves multiple fields including fuzzy sets, rough sets, three-way decisions, etc. A granule acts as a basic information unit with characteristics, and its division processing is called information granulation; thus, granular computing means the combination and transformation at distinctive layers, and the associative structure of all granular levels constitutes a granular structure [24]. In fact, neighborhood rough sets are closely related to granular computing, and thus their underlying neighborhood system serves as an important formal context for relevant information granulation and measurement computing. The neighborhood system was proposed by Lin [25] and it motivated granular computing in terms of data mining, and it acquires subsequent studies and applications [26–29]. In particular, the neighborhood system establishes two types of tri-level granular structure by Zhou et al. [30] and Chen et al. [31]; by contrast, the former describes both the condition and decision granulation by their interaction [30], while the latter places emphasis on only the condition granulation by its aggregation [31]. Note that these tri-level granular structures closely follow the sprint of granular computing, especially the tri-level thinking discussed by Yao [32,33].

* Corresponding author.
E-mail address: xianyongzh@sina.com.cn (X. Zhang).

https://doi.org/10.1016/j.knosys.2021.106799
0950-7051/© 2021 Elsevier B.V. All rights reserved.

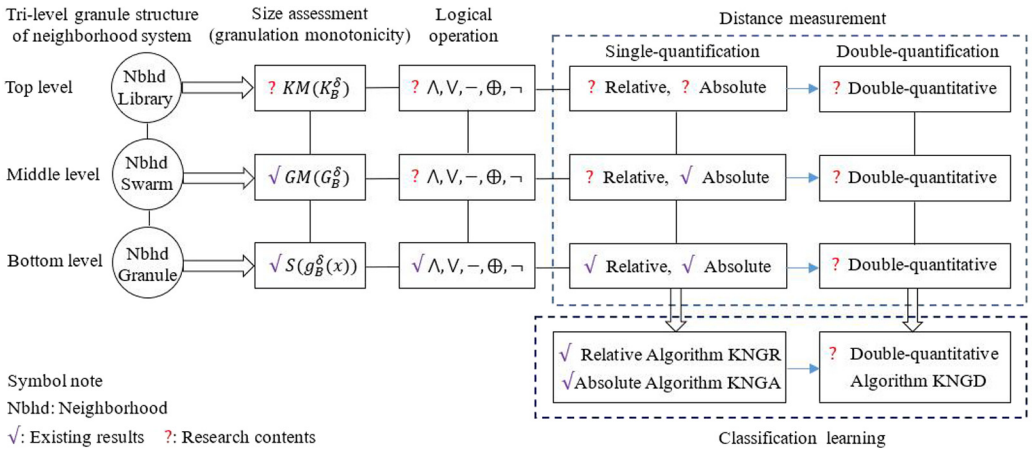


Fig. 1. Research framework of this paper.

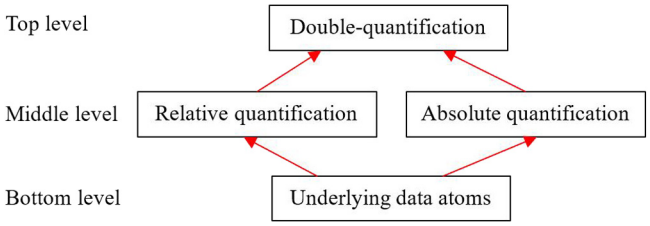


Fig. 2. Tri-level information structure of double-quantification construction.

Now focus on the tri-level granular structure of neighborhood system, proposed by Chen et al. [31] from the perspective of granular computing. By an observation point, the bottom neighborhood granule collects clustering objects in information system to become a basic cognition unit, and its set constitutes the middle neighborhood swarm; furthermore, neighborhood swarms from multiple observation points construct the top neighborhood library. As a result, this tri-level granular structure provides a granular computing mechanism for knowledge-based learning. However, its hierarchical exploration is inadequate, while its measurement can be extended for robust applications. The hierarchical blank and valuable extension generate the motivation of this paper, and they are clarified in Fig. 1 mainly by the items labeled by symbol "?".

(1) The existing studies (including those on the size assessment, logical operation, and distance measurement) stay at lower levels [31], and thus the surplus cases at higher levels are worth completing. The relevant feasibility is ensured by the hierarchy mechanism of tri-level granular structure.
(2) Distance measures, including those in neighborhood system, are important for classification learning and complexity reduction. In [31], the distance discussions concern both the relative and absolute distances and their corresponding classifiers KNGR and KNGA, and thus the double-quantification thought can be naturally introduced to establish a double-quantitative distance and its corresponding classifier for better measurement and performance.

Since the above distance construction refers to the double-quantification, let us simply introduce this methodology. Within the framework of rough sets [1], the double-quantification was first proposed by Zhang et al. [34] to quantify the approximate space. Then, the double-quantification strategy was deeply developed, and nowadays it has been extensively utilized to research the uncertainty modeling [35–43], approximation measurement [44,45], information fusion [46–48], knowledge acquisition [49,50], etc. The double-quantification mainly comes from two sides of single quantification, i.e., the relative and absolute quantification, and it embodies a basic mechanism of hierarchical construction. As shown in Fig. 2, it actually adheres to a kind of tri-level information structure.

(1) The bottom level provides basic or inseparable data atoms, and it is usually microscopic to underlie further macroscopic quantification and applications.
(2) The middle level concerns two symmetrical measurement patterns. Both the relative and absolute quantification synthesize bottom data to offer information concentration, but they have different perspectives and emphases. The relative quantification focuses on the measurement relativity regarding a particular range, and thus it usually resorts to contrastive ratios to concern the locality and restriction; in contrast, the absolute quantification considers the measurement absoluteness without a particular restriction or in a general scope, and thus it usually adopts direct and outstanding values to embody the globality and unconstraint.
(3) The relative and absolute quantification have their own representation superiorities and application environments, and they constitute two distinctive sides of measurement to observe a dialectical relationship. For example, they cannot be mutually determined or deduced to become linearly independent. Accordingly, they are worth systematically combining, and the double-quantification emerges by integrating the both single-quantification patterns. The double-quantification is located at the top level to synthesize the two types of middle measurement as well as their merits, so it comprehensively provides a more complete and robust measurement pattern.

The double-quantification carries the above tri-level structure to serve as a basic model of tri-level analysis (related to three-way decisions [32,33,51]), and its relevant tri-level thinking (such as the integrated construction) becomes innovative and significant for measurement and optimization. In this paper, the double-quantification technology will be used for expanding and deepening the tri-level granular structure of neighborhood system, mainly in terms of the distance measurement and classification learning. In particular, the double-quantitative distance refers to

and comes from the information fusion of two types of single-quantitative distances, i.e., the relative and absolute quantitative distances which are calculated by the underlying data elements, and it concerns the measure integration at the top layer within the hierarchical framework of tri-level information structure based on information granulation; as a result, the double-quantitative distance performs the powerful information measurement to underlie effective classification learning.

1.2. Research contents and innovations

Against the above background, the tri-level granular structure of neighborhood system serves as an important basis for knowledge-based learning of neighborhood covering generalization, and its hierarchy construction and measurement application are worth researching. In this paper, the tri-level granular structure of neighborhood system is investigated, and its double-quantitative distance measurement and classification learning are emphatically implemented by deftly utilizing the double-quantification, a basic methodology for measurement and reasoning.

Concretely, we will basically complete the hierarchical construction by several focuses (including the size assessment, logical operation, and distance measurement), and we will further extend the single-quantitative distances and classifiers to a double-quantitative distance and classifier, respectively. The detailed research framework about bases and developments is clarified in Fig. 1. As shown by Fig. 1, the main discussions concern four specific contents, i.e.,

- the size assessment and logical operation at higher levels,
- the double-quantification extension and hierarchical promotion of single-quantitative distances,
- the double-quantitative classifier design based on a double-quantitative distance,
- the example illustration for theoretical construction and experiment verification for double-quantitative classification performance.

In particular, the distance measurement and classification learning based on the double-quantification become the emphasis and difficulty, and thus they will fully resort to the double-quantification method. According to these research contents, our works exhibit two basic innovations regarding the tri-level granular structure of neighborhood system, and the two points or contributions exactly match the above motivation, which is also reflected by Fig. 1.

(1) In terms of theory, the size assessment, logical operation, and distance measurement of tri-level granular structure are hierarchically deepened and completed.
(2) In terms of applications, the distance measurement is double-quantitatively extended, and thus double-quantitative classifier KNGD is constructed to integrate and promote single-quantitative classifiers KNGR and KNGA.

As a result, this study hierarchically perfects the tri-level granular structure of neighborhood system, and the related double-quantification integration and extension improve the knowledge measurement and classification achievement.

The three-way decisions proposed by Yao [33,52] have many studies; for example, the three-way fuzzy concept lattice representation on the uncertainty and incompleteness was proposed by using neutrosophic sets [53]. From the perspective of three-way decisions, our research work on the tri-level thinking and analysis can be further supported by or connected with some relevant works. The double-quantitative distance measurement embraces the tri-level granular structure, but its neighborhood system mainly requires the Euclidean distance; thus, we refer to the quantitative threshold measurement in [54] and the Euclidean distance application in [55], which are both based on the three-way decisions. As a main part, the granular computing for neighborhoods is incorporated, and it is related to the processing based on interval sets in [56,57]; in particular, the usual granulation level depends on attribute subsets [58], the relevant attribute reduction process concerns the power set of attribute set to cause a major issue of exponential time complexity [59], the information granules have the multilevel and multiview [60], and all these results provide us more thoughts on the neighborhood granulation and information processing.

The remainder of this paper is organized as follows. Section 2 makes the preliminary tri-level analysis of neighborhood system, including the review of neighborhood system and its tri-level structure, the hierarchical construction of size assessment and logical operation, and the example illustration. Section 3 focuses on the distance measurement of tri-level granular structure, including the double-quantitative extension, the hierarchical promotion, and the example illustration. Section 4 turns to the classification learning based on the double-quantitative distance of neighborhood granules, including the classifier algorithm design and data experiment verification. Finally, Section 5 concludes this paper to underlie the future explorations.

2. Preliminary tri-level analysis of neighborhood system

Neighborhood rough sets come into play mainly by the neighborhood system and its inherent granule structure. As a preliminary discussion, the neighborhood system next acquires its tri-level analysis.

2.1. Granulation of neighborhood system

The neighborhood system and its granulation are first reviewed by Ref. [61].

A neighborhood system is denoted by $NS = (U, A, V, f, \delta)$. Herein, $U = \{x_1, x_2, \ldots, x_{|U|}\}$ is the universe with samples, A is a set of attributes, $V = \bigcup_{a\in A} V_a$ collects all attribute values where V_a implies attribute values of $a \in A$, $f : U \times A \to V$ is an information function, and $\delta \in \mathbb{R}^+ \cup \{0\}$ serves as a neighborhood radius. The neighborhood granulation focuses on an attribute subset and its distance function. Regarding subset $B = \{a_1, a_2, \ldots, a_n\} \subseteq A$, a usual distance function is defined by

$$D_B(x, y) = \left[\sum_{l=1}^{n}(|f(x, a_l) - f(y, a_l)|)^p\right]^{\frac{1}{p}}. \tag{1}$$

This distance satisfies the distance axioms with non-negativity (identity), symmetry, and triangular inequality; $D_B(x, y)$ is called the Manhattan distance when $p = 1$, while it becomes the Euclidean distance when $p = 2$.

In this paper, neighborhood system $NS = (U, A, V, f, \delta)$ serves as the formal context for discussions, and thus we first provide several symbol assumptions, i.e., $x, y \in U$, $B \subseteq A$, and $2^B = \{R|R \subseteq B\} = \{\emptyset, R_1, \ldots, R_m\}$ where $m = 2^{|B|} - 1$. Moreover, $|.|$ denotes the set cardinality.

Definition 1 ([61]). In $NS = (U, A, V, f, \delta)$, the δ-neighborhood (class) of object $x \in U$ regarding attribute subset $B \subseteq A$ is defined by

$$n_B^\delta(x) = \{y \in U|D_B(x, y) \le \delta\}, \tag{2}$$

the neighborhood relation regarding $B \subseteq A$ is defined by

$$NR_\delta(B) = \{(x, y) \in U \times U|D_B(x, y) \le \delta\}, \tag{3}$$

while the neighborhood covering regarding $B \subseteq A$ is defined by

$$U/NR_\delta(B) = \{n_B^\delta(x)|x \in U\}. \tag{4}$$

The granulation of neighborhood system is completed by distance function and neighborhood radius. The neighborhood acts as a basic granule unit for constructions, the neighborhood relation is a sort of similarity relation with both the reflexivity and symmetry, while the neighborhood covering gathers neighborhoods to constitute a kind of granulation structure underlying knowledge reasoning. The neighborhood granulation offers both the theoretical extension and practical adjustment, because $\delta = 0$ causes the degeneration for classical classification-granulation (with the equivalent class, relation, and partition) while $\delta > 0$ tends to yield application performances.

Definition 2 ([61]). In $NS = (U, A, V, f, \delta)$, the neighborhood lower and upper approximations of $X \subseteq U$ regarding $B \subseteq A$ are defined by

$$\begin{cases} \underline{B}(X)_\delta = \{x \in U|n_B^\delta(x) \subseteq X\}, \\ \overline{B}(X)_\delta = \{x \in U|n_B^\delta(x) \cap X \neq \emptyset\}. \end{cases} \tag{5}$$

By virtue of neighborhood granulation, the dual neighborhood approximations bi-directionally approach a basic concept to constitute neighborhood rough sets, where $\underline{B}(X)_\delta \subseteq X \subseteq \overline{B}(X)_\delta$, while the latter can induce the further dependency-based classification.

The granulation and its monotonicity play an important role in metric measurement and knowledge reasoning [62–65], and they concern two basic cases in neighborhood system.

(1) In terms of attribute granulation, $\emptyset \subset P \subseteq Q \subseteq A$ offers the refining $n_P^\delta(x) \supseteq n_Q^\delta(x)$ and monotonicity $\underline{P}(X)_\delta \subseteq \underline{Q}(X)_\delta$, $\overline{P}(X)_\delta \supseteq \overline{Q}(X)_\delta$.
(2) In terms of radius granulation, $0 \leq \gamma \leq \delta \leq 1$ leads to the coarsening $n_B^\gamma(x) \subseteq n_B^\delta(x)$ and monotonicity $\underline{B}(X)_\gamma \supseteq \underline{B}(X)_\delta$, $\overline{B}(X)_\gamma \subseteq \overline{B}(X)_\delta$.

2.2. Tri-level granular structure of neighborhood system

The neighborhood system contains many attributes; thus, a neighborhood focuses on certain attributes to establish an observation point, while further granular collections provide steady structural platforms for representations and applications. In particular, Chen et al. [31] proposed a tri-level granular structure of neighborhood system, and the relevant hierarchical study is next recalled and developed.

Definition 3 ([31]). In $NS = (U, A, V, f, \delta)$, three granular levels of neighborhood granule, swarm, and library are defined as follows.

(1) $g_B^\delta(x) = n_B^\delta(x)$ is a neighborhood granule of x regarding B;
(2) $G_B^\delta = \left(n_B^\delta(x_1), \ldots, n_B^\delta(x_{|U|})\right)$ is a neighborhood swarm of B;
(3) $K_B^\delta = \left(G_{R_1}^\delta, \ldots, G_{R_m}^\delta\right)$ is a neighborhood library of B, where $\{R_1, \ldots, R_m\} = 2^B - \{\emptyset\}$ collects all non-empty subsets of B.

To exactly present the potential repeatability regarding universe objects or attribute subsets, G_B^δ uses granule group $\left(n_B^\delta(x_1), \ldots, n_B^\delta(x_{|U|})\right)$ rather than granule set (i.e., neighborhood covering) $\{n_B^\delta(x)|\forall x \in U\}$, while K_B^δ uses group family $\left(G_{R_1}^\delta, \ldots, G_{R_m}^\delta\right)$ rather than set family $\{G_R^\delta|\forall \emptyset \neq R \subseteq B\}$.

The neighborhood granule, swarm, and library constitute a tri-level granular structure of neighborhood system. Next, we provide a relevant frame diagram – Fig. 3 – to make an in-depth clarification. Fig. 3 has two parts; subfigure (a) directly focuses on the general construction of Definition 3, while subfigure (b) deeply embodies the concrete organization of top neighborhood library. By Fig. 3's (a), $g_B^\delta(x)$, G_B^δ, K_B^δ exhibit a good granular hierarchy.

(1) At the bottom, a neighborhood granule is exactly the neighborhood class, where $g_B^\delta(x) = n_B^\delta(x)$, and thus it provides a basic observation side to gather adjacent samples.
(2) At the middle, a neighborhood swarm collects all neighborhood classes related to the universe object and neighborhood covering, and thus it forms a kind of applied knowledge with traceability.
(3) At the top, a neighborhood library includes all groups of traceable knowledge (i.e., all neighborhood swarms) on all non-empty attribute subsets, and thus it provides a sort of complete knowledge base.

The relevant construction adopts the collection function to closely follow the hierarchial integration in the bottom-middle-top direction, which is a basic strategy of granular computing, so it exhibits a development of abstraction and generalization.

(1) The bottom granule concerns a micro description. $g_B^\delta(x)$ depends on x and B to become numerous, and their number is related to sample cardinality $|U|$.
(2) The middle swarm places emphasis on a meso viewpoint. G_B^δ relies on B to offer only number 1 but carries $|U|$ neighborhoods.
(3) The top library refers to a macro change. K_B^δ is also related to B to exhibit only number 1, and it contains $m = 2^{|B|} - 1$ types of swarms or coverings.

In contrast, the reverse strategy in the top-middle-bottom can also be concerned to provide a sort of hierarchial decomposition. This decomposition direction can be used to observe the neighborhood library, which has a complex structure. By Fig. 3's (b), the top library $K_B^\delta = \left(G_{R_1}^\delta, \ldots, G_{R_m}^\delta\right)$ carries $m = 2^{|B|} - 1$ middle swarms or $m|U|$ bottom granules. In summary, the tri-level structure firmly underlies the subsequent hierarchial processing based on neighborhood knowledge, and it differs from the three-layer granular structures of a decision table [51] and of a neighborhood system [30].

2.3. Size assessment of tri-level granular structure

In neighborhood system, the above tri-level granular structure is worth measuring for applications. The size measures of both the bottom granule and middle swarm have been offered in [31], and the top level of neighborhood library will be endowed with a corresponding measure.

Definition 4 ([31]). In $NS = (U, A, V, f, \delta)$, the size of neighborhood granule $g_B^\delta(x)$ is defined by

$$S(g_B^\delta(x)) = \frac{|g_B^\delta(x)|}{|U|} = \frac{|n_B^\delta(x)|}{|U|}, \tag{6}$$

while the size of neighborhood swarm G_B^δ is defined by

$$\begin{aligned} GM(G_B^\delta) &= \frac{1}{|U|}\sum_{i=1}^{|U|} S(g_B^\delta(x_i)) = \frac{1}{|U|^2}\sum_{i=1}^{|U|} |g_B^\delta(x_i)| \\ &= \frac{1}{|U|^2}\sum_{i=1}^{|U|} |n_B^\delta(x_i)|. \end{aligned} \tag{7}$$

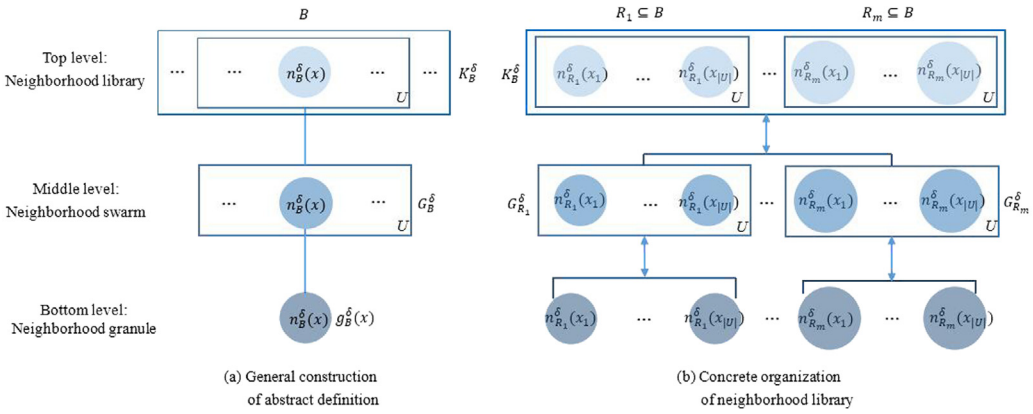


Fig. 3. Organizational chart of tri-level granular structure of neighborhood system.

Proposition 1 ([31]). *In $NS = (U, A, V, f, \delta)$, we have the following granulation monotonicity (regarding both attribute and radius).*

(1) At the bottom level, if $\emptyset \subset P \subseteq Q$ then $g_P^\delta(x) \supseteq g_Q^\delta(x)$ and $S(g_P^\delta(x)) \geq S(g_Q^\delta(x))$, while if $0 \leq \gamma \leq \delta \leq 1$ then $g_B^\gamma(x) \subseteq g_B^\delta(x)$ and $S(g_B^\gamma(x)) \leq S(g_B^\delta(x))$.

(2) At the middle level, if $\emptyset \subset P \subseteq Q$ then $GM(G_P^\delta) \geq GM(G_Q^\delta)$, while if $0 \leq \gamma \leq \delta \leq 1$ then $GM(G_B^\gamma) \leq GM(G_B^\delta)$.

Both the bottom granule and middle swarm acquire their size measures in [31], as discussed by Definition 4. Bottom measure $S(g_B^\delta(x))$ denotes the proportion of granule $g_B^\delta(x)$ in contrast to the whole universe U, while middle measure $GM(G_B^\delta)$ reflects the arithmetic mean by counting all samples and their sizes of neighborhood granules. Eqs. (6) and (7) also reflect both the bottom-middle connection and neighborhood essence, and thus the measure range and optimization implementation are given as follows:

$$\begin{aligned} &S(g_B^\delta(x)) \in [\frac{1}{|U|}, 1], \\ &\quad S(g_B^\delta(x)) = \frac{1}{|U|} \Longleftrightarrow g_B^\delta(x) = \{x\}, \\ &\quad S(g_B^\delta(x)) = 1 \Longleftrightarrow g_B^\delta(x) = U; \\ &GM(G_B^\delta) \in [\frac{1}{|U|}, 1], \\ &\quad GM(G_B^\delta) = \frac{1}{|U|} \Longleftrightarrow G_B^\delta = \big(\{x_1\}, \ldots, \{x_{|U|}\}\big), \\ &\quad GM(G_B^\delta) = 1 \Longleftrightarrow G_B^\delta = \big(U, \ldots, U\big). \end{aligned} \tag{8}$$

Furthermore, the two types of size measures have the fundamental granulation monotonicity regarding both the attribute and radius.

However, the top library never achieves its size measure [31], and here we make a corresponding realization.

Definition 5. In $NS = (U, A, V, f, \delta)$, the size of neighborhood library K_B^δ is defined by

$$KM(K_B^\delta) = \frac{1}{m} \sum_{j=1}^{m} GM(G_{R_j}^\delta), \tag{9}$$

where R_j means the non-empty subset of B to have the total number $m = 2^{|B|} - 1$.

Proposition 2. *In $NS = (U, A, V, f, \delta)$, we have*

$$\begin{aligned} KM(K_B^\delta) &= \frac{1}{m|U|} \sum_{j=1}^{m} \sum_{i=1}^{|U|} S(g_{R_j}^\delta(x_i)) = \frac{1}{m|U|^2} \sum_{j=1}^{m} \sum_{i=1}^{|U|} |g_{R_j}^\delta(x_i)| \\ &= \frac{1}{m|U|^2} \sum_{j=1}^{m} \sum_{i=1}^{|U|} |n_{R_j}^\delta(x_i)|, \end{aligned} \tag{10}$$

and thus

$$\begin{aligned} &KM(K_B^\delta) \in [\frac{1}{|U|}, 1], \\ &KM(K_B^\delta) = \frac{1}{|U|} \Longleftrightarrow \forall R_j \subseteq B \big[G_{R_j}^\delta = \big(\{x_1\}, \ldots, \{x_{|U|}\}\big)\big] \\ &\quad \Longleftrightarrow \forall R_j \subseteq B, \forall x_i \in U \big[g_B^\delta(x_i) = \{x_i\}\big], \\ &KM(K_B^\delta) = 1 \Longleftrightarrow \forall R_j \subseteq B \big[G_{R_j}^\delta = \big(U, \ldots, U\big)\big] \\ &\quad \Longleftrightarrow \forall R_j \subseteq B, \forall x_i \in U \big[g_B^\delta(x_i) = U\big]. \end{aligned} \tag{11}$$

The size construction of top library follows the hierarchical integration thought of middle swarm.

(1) Since middle swarm G_B^δ contains $|U|$ bottom granules, its size $GM(G_B^\delta)$ is defined by the arithmetic mean of size values of all neighborhood granules (Definition 4 [31]). Thus, measure $GM(G_B^\delta)$ represents the statistical characteristic of the middle swarm system.

(2) Similarly, since top library K_B^δ contains m middle swarms, its size $KM(K_B^\delta)$ is defined by the arithmetic mean of size values of all m neighborhood swarms (Definition 5), i.e., $KM(K_B^\delta)$ refers to the arithmetic mean of size values of $m|U|$ neighborhood granules (Eq. (10)). Hence, measure $KM(K_B^\delta)$ represents the statistical characteristic of the top library system.

The top size measure becomes continuous and reasonable to perfectly have the hierarchical mechanism and statistical semantics. Its top-middle and top-bottom connections are revealed in Eqs. (9) and (10), respectively. The latter formula also exhibits the statistical essence of multiple neighborhoods, which are carried by m neighborhood swarms (each swarm contains $|U|$ granules). Furthermore, $KM(K_B^\delta)$ also belongs to closed interval $[\frac{1}{|U|}, 1]$ because both $S(g_B^\delta(x))$ and $GM(G_B^\delta)$ are in this range (Eq. (8)).

Now focus on the basic topic of granulation monotonicity/non-monotonicity. Note that the top library and its size respectively concern multiple middle swarms and their size average. Thus, the change and correspondence of swarms regarding covering granulation become a key question, and they carry the complexity and difficulty. The case of radius granulation is simple, because the same observed attribute subset B has the same non-empty

attribute sets $R_1, \ldots, R_m \subseteq B$ to determine internal m swarms. However, the other case of attribute granulation becomes complex, because attribute enlargement $B \subseteq Q$ causes the subset addition $2^B \subseteq 2^Q$ and the swarm addition. For the above purpose and issue, the radius-granulation monotonicity is directly proved, while the attribute-granulation non-monotonicity will be deeply mined and proved by two lemmas. Concretely, Lemma 1 clarifies the enlargement distribution of attribute enlargement $B \subset B \cup \{a_*\} = Q \subseteq A$ when adding only a single attribute $a_* \in A$, and thus Lemma 2 provides two cases of size inequation by adopting two special cases of neighborhood knowledge (i.e., the finest and coarsest coverings). Note that proofs of the two lemmas are respectively placed in Appendices A and B.

Lemma 1. *Regarding the addition of single attribute, we have the following enlargement distribution of non-empty subsets, which refers to the structure correspondence of power sets. That is,*

$$\begin{aligned} & Q = B \cup \{a_*\} \supset B \neq \emptyset \Longrightarrow \\ 2^Q - \{\emptyset\} &= (2^B - \{\emptyset\}) \cup \{R_1 \cup \{a_*\}, \ldots, R_m \cup \{a_*\}\} \cup \{\{a_*\}\} \\ &= \{R_1, \ldots, R_m\} \cup \{R_1 \cup \{a_*\}, \ldots, R_m \cup \{a_*\}\} \cup \{\{a_*\}\}. \end{aligned} \tag{12}$$

Lemma 2. *In $NS = (U, A, V, f, \delta)$ with $\emptyset \subset B \subset A$, generally suppose that all non-empty subsets R_j ($j = 1, \ldots, m$) of B reach neither the finest nor coarsest coverings, i.e.,*

$$\begin{aligned} & \forall R_j \in 2^B - \{\emptyset\} \\ & \left(U/NR_\delta(R_j) \neq \{\{x_1\}, \{x_2\}, \ldots, \{x_{|U|}\}\} \wedge U/NR_\delta(R_j) \neq \{U\} \right). \end{aligned} \tag{13}$$

Next, an attribute $a_* \in A$ is added to B, and thus $B \subset B \cup \{a_*\} = Q \subseteq A$.

(1) If attribute a_* acquires the finest covering, i.e., $U/NR_\delta(\{a_*\}) = \{\{x_1\}, \{x_2\}, \ldots, \{x_{|U|}\}\}$, then $KM(K_B^\delta) > KM(K_Q^\delta)$.

(2) If a_* corresponds to the coarsest covering, i.e., $U/NR_\delta(\{a_*\}) = \{\{U\}\}$, then $KM(K_B^\delta) < KM(K_Q^\delta)$.

Proposition 3. *In $NS = (U, A, V, f, \delta)$, the top size exhibits the attribute-granulation non-monotonicity and radius-granulation monotonicity.*

(1) Regarding the attribute granulation, if $\emptyset \subset P \subseteq Q$ then neither $KM(K_P^\delta) \geq KM(K_Q^\delta)$ nor $KM(K_P^\delta) \leq KM(K_Q^\delta)$ necessarily holds.

(2) Regarding the radius granulation, if $0 \leq \gamma \leq \delta \leq 1$ then $KM(K_B^\gamma) \leq KM(K_B^\delta)$.

Proof. (1) According to Lemma 2, setting up $P = B \subset B \cup \{a_*\} = Q$ can theoretically acquire two possible cases: $KM(K_P^\delta) < KM(K_Q^\delta)$ and $KM(K_P^\delta) > KM(K_Q^\delta)$, so we cannot always achieve $KM(K_P^\delta) \geq KM(K_Q^\delta)$ or $KM(K_P^\delta) \leq KM(K_Q^\delta)$.

(2) Libraries K_B^γ and K_B^δ include the same number $m = 2^{|B|} - 1$ of middle swarms regarding $R_1, \ldots, R_m$. According to Proposition 1(2), we have $\forall R_j \subseteq B \left(GM(G_{R_j}^\gamma) \leq GM(G_{R_j}^\delta) \right)$; according to Definition 5,

$$KM(K_B^\gamma) = \frac{1}{m} \sum_{j=1}^{m} GM(G_{R_j}^\gamma) \leq \frac{1}{m} \sum_{j=1}^{m} GM(G_{R_j}^\delta) = KM(K_B^\delta). \quad \square$$

Lemmas 1 and 2 embody interesting and skillful constructions. They focus on a special case of attribute granulation based on the single-attribute addition, and they also adopt a special case of extreme assumptions of the finest/coarsest covering (as well as a general precondition). As a result, they fully prove the attribute-granulation non-monotonicity by mining counter examples. With the addition of radius-granulation monotonicity, Proposition 3 becomes true. The relevant conclusions of top size measure go beyond our expectation, i.e., they differ from the corresponding results of size measures at lower levels (Proposition 1). In other words, it is surprising but objective to discover that the attribute granulation has not the monotonicity but the non-monotonicity, which is mainly caused by the structure feature of attribute subsets.

We can focus on the particular feature of non-monotonicity in Proposition 3 from a statistical perspective. According to the relevant mechanism, attribute enlargement $Q \supseteq P$ may supplement non-empty subsets and their swarms, but it cannot necessarily increase/decrease the initial size average of middle swarms. Thus, the size $KM(K_Q^\delta)$ on Q cannot necessarily increase/decrease initial $KM(K_P^\delta)$ on P. In fact, if all additional swarms provide little size values (including the minimum $\frac{1}{|U|}$ related to the finest covering) then $KM(K_Q^\delta)$ more tends to be less than $KM(K_P^\delta)$, while if all additional swarms provide large size values (including the maximum 1 related to the coarsest covering) then $KM(K_Q^\delta)$ more tends to be larger than $KM(K_P^\delta)$. In summary, $KM(K_Q^\delta) < KM(K_P^\delta)$ and $KM(K_Q^\delta) > KM(K_P^\delta)$ become two possible cases, and the actual size relationships between $KM(K_Q^\delta)$ and $KM(K_P^\delta)$ depend on the average statistics to exhibit a sort of uncertainty, which generalizes and transcends the non-monotonicity.

2.4. Logical operations of tri-level granular structure

Logical operations are basic for a collection system. They are defined for the bottom granule in [31], and they are further supplemented for the remaining middle swarm and top library.

Definition 6 ([31])**.** In $NS = (U, A, V, f, \delta)$, two granules $g_B^\delta(x)$ and $g_B^\delta(y)$ regarding $B \subseteq A$ have the following logical operations:

$$\begin{aligned} &(1)\ \text{AND:}\ g_B^\delta(x) \wedge g_B^\delta(y) = n_B^\delta(x) \cap n_B^\delta(y), \\ &(2)\ \text{OR:}\ g_B^\delta(x) \vee g_B^\delta(y) = n_B^\delta(x) \cup n_B^\delta(y), \\ &(3)\ \text{DIF:}\ g_B^\delta(x) - g_B^\delta(y) = n_B^\delta(x) - n_B^\delta(y), \\ &(4)\ \text{XOR:}\ g_B^\delta(x) \oplus g_B^\delta(y) = n_B^\delta(x) \cup n_B^\delta(y) - n_B^\delta(x) \cap n_B^\delta(y), \\ &(5)\ \text{NOT:}\ \neg g_B^\delta(x) = U - n_B^\delta(x). \end{aligned} \tag{14}$$

Proposition 4. *In $NS = (U, A, V, f, \delta)$, the usual system of bottom logical operations does not have the closeness on the underlying covering, i.e.,*

$$\begin{aligned} & \left[g_B^\delta(x) \in U/n_B^\delta\right] \wedge \left[g_B^\delta(y) \in U/n_B^\delta\right] \\ \not\Rightarrow & \left[g_B^\delta(x) \wedge g_B^\delta(y) \in U/n_B^\delta\right] \wedge \left[g_B^\delta(x) \vee g_B^\delta(y) \in U/n_B^\delta\right] \\ & \wedge \left[g_B^\delta(x) - g_B^\delta(y) \in U/n_B^\delta\right] \\ & \wedge \left[g_B^\delta(x) \oplus g_B^\delta(y) \in U/n_B^\delta\right] \wedge \left[\neg g_B^\delta(x) \in U/n_B^\delta\right]. \end{aligned} \tag{15}$$

At the bottom level, the neighborhood granule has the set essence, so the five usual types of logical operations are naturally defined by the set operations regarding intersection, union, and complement. However, there is no the operation closeness regarding a fixed attribute subset, i.e., the result of logical operation may not be necessarily a neighborhood.

Definition 7. In $NS = (U, A, V, f, \delta)$, swarm G_B^δ (regarding $B \subseteq A$) and swarm $G_{B'}^\delta$ (regarding $B' \subseteq A$) have the following logical

operations:

$$\begin{aligned}&(1)\ \text{AND:}\ \ G_B^\delta \wedge G_{B'}^\delta = G_{B\cap B'}^\delta,\\ &(2)\ \text{OR:}\ \ \ G_B^\delta \vee G_{B'}^\delta = G_{B\cup B'}^\delta,\\ &(3)\ \text{DIF:}\ \ G_B^\delta - G_{B'}^\delta = G_{B-B'}^\delta,\\ &(4)\ \text{XOR:}\ \ G_B^\delta \oplus G_{B'}^\delta = G_{B\oplus B'}^\delta,\\ &(5)\ \text{NOT:}\ \ \neg G_B^\delta = G_{A-B}^\delta,\end{aligned} \tag{16}$$

while library K_B^δ (regarding $B \subseteq A$) and library $K_{B'}^\delta$ (regarding $B' \subseteq A$) have the following logical operations:

$$\begin{aligned}&(1)\ \text{AND:}\ \ K_B^\delta \wedge K_{B'}^\delta = K_{B\cap B'}^\delta,\\ &(2)\ \text{OR:}\ \ \ K_B^\delta \vee K_{B'}^\delta = K_{B\cup B'}^\delta,\\ &(3)\ \text{DIF:}\ \ K_B^\delta - K_{B'}^\delta = K_{B-B'}^\delta,\\ &(4)\ \text{XOR:}\ \ K_B^\delta \oplus K_{B'}^\delta = K_{B\oplus B'}^\delta,\\ &(5)\ \text{NOT:}\ \ \neg K_B^\delta = K_{A-B}^\delta.\end{aligned} \tag{17}$$

Proposition 5. *In $NS = (U, A, V, f, \delta)$, the usual system of middle/top logical operations has the closeness regarding all attribute subsets. Let $\mathcal{G}^\delta = \{G_B^\delta | B \subseteq A\}$ and $\mathcal{K}^\delta = \{K_B^\delta | B \subseteq A\}$, where $G_\emptyset^\delta = null$ and $K_\emptyset^\delta = null$ represent the emptiness, and thus*

$$\begin{aligned}&\left[G_B^\delta \in \mathcal{G}^\delta\right] \wedge \left[G_{B'}^\delta \in \mathcal{G}^\delta\right]\\ \Rightarrow&\left[G_B^\delta \wedge G_{B'}^\delta \in \mathcal{G}^\delta\right] \wedge \left[G_B^\delta \vee G_{B'}^\delta \in \mathcal{G}^\delta\right]\\ \wedge&\left[G_B^\delta - G_{B'}^\delta \in \mathcal{G}^\delta\right] \wedge \left[G_B^\delta \oplus G_{B'}^\delta \in \mathcal{G}^\delta\right] \wedge \left[\neg G_B^\delta \in \mathcal{G}^\delta\right],\\ &\left[K_B^\delta \in \mathcal{K}^\delta\right] \wedge \left[K_{B'}^\delta \in \mathcal{K}^\delta\right]\\ \Rightarrow&\left[K_B^\delta \wedge K_{B'}^\delta \in \mathcal{K}^\delta\right] \wedge \left[K_B^\delta \vee K_{B'}^\delta \in \mathcal{K}^\delta\right] \wedge \left[K_B^\delta - K_{B'}^\delta \in \mathcal{K}^\delta\right]\\ \wedge&\left[K_B^\delta \oplus K_{B'}^\delta \in \mathcal{K}^\delta\right] \wedge \left[\neg K_B^\delta \in \mathcal{K}^\delta\right].\end{aligned} \tag{18}$$

Since middle swarms are completely determined by attribute subsets, their logical operations are basically defined by corresponding logical operations regarding attribute subsets. This definition strategy is not only simple and direct but also effective, and thus Proposition 5 offers the closeness merit on the collection set $\mathcal{G}^\delta$. The top library has the similar case and analysis for logical operations. In terms of knowledge base, the top level can also consider interaction between internal swarms on 2^B and $2^{B'}$, so to explore other logical operations between K_B^δ and $K_{B'}^\delta$. Moreover, these logical operations come from the set operations rather than the hierarchical integration; in contrast, the hierarchical integration is more related to the tri-level structure, so it is worth introducing and discussing for further works on logical operations.

2.5. Illustrative example I

Based on a neighborhood system, an example is provided to illustrate the above relevant notions and results, including the tri-level granular structure and its size assessment and logical operation.

Example 1. The neighborhood decision system $(U, A\cup\{d\}, V, f, \delta)$ is given in Table 1, and it internally includes a neighborhood system (U, A, V, f, δ). Herein, $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, $A = \{a, b, c\}$, and d implies the sole decision attribute. For neighborhood granulation, the Euclidean distance (i.e., $p = 2$ in Eq. (1)) and distance radius $\delta = 0.3$ are utilized; next, each non-empty subset $B \subseteq A$ is focused on, and there are in total 7 subsets.

As a basis, Table 2 exhibits neighborhoods of all 6 universe objects regarding 7 attribute subsets. Thus, an arbitrary attribute subset $B \subseteq A$ can generate a tri-level granular structure; furthermore, the corresponding neighborhood granule, swarm, and library can be measured, and Table 3 provides all size values. As an example, we choose a representative subset $B = \{a, b\}$ to make detailed explanations.

Table 1
A neighborhood decision system.

U	a	b	c	d
x_1	1	0.8	1	N
x_2	1	0	0.3	Y
x_3	0.8	1	0.6	N
x_4	0.3	0.8	0	Y
x_5	0.2	0.6	0	Y
x_6	0	0.4	0	N

(1) The bottom granules concern 6 neighborhoods, i.e., $g_B^\delta(x_1) = \{x_1, x_3\}, \ldots, g_B^\delta(x_6) = \{x_5, x_6\}$. They exhibit 6 values, i.e., $S(g_B^\delta(x_1)) = 0.3333, \ldots, S(g_B^\delta(x_6)) = 0.3333$.

(2) The middle swarm has only number 1 to carry 6 neighborhoods, i.e., $G_B^\delta = \left(\{x_1, x_3\}, \ldots, \{x_5, x_6\}\right)$, and it corresponds to covering $U/NR_\delta(B) = \{\{x_1, x_3\}, \{x_2\}, \{x_4, x_5\}, \{x_4, x_5, x_6\}, \{x_5, x_6\}\}$ with only 5 different neighborhoods. The size $GM(G_B^\delta)$ becomes the arithmetic mean of all 6 values of $S(g_B^\delta(x_i))$ $(i = 1, 2, \ldots, 6)$ to gain value 0.3333.

(3) The top library has only number 1 to carry 3 swarms regarding non-empty subsets $\{a\}, \{b\}, \{a, b\}$. It offers

$$\begin{aligned}K_B^\delta &= \left(G_{\{a\}}^\delta, G_{\{b\}}^\delta, G_{\{a,b\}}^\delta\right)\\ &= \big(\ (\{x_1, x_2, x_3\}, \ldots, \{x_4, x_5, x_6\}),\\ &\quad (\{x_1, x_3, x_4, x_5\}, \ldots, \{x_5, x_6\}), (\{x_1, x_3\}, \ldots, \{x_5, x_6\})\ \big),\end{aligned}$$

and it is related to knowledge base (with 3 coverings)

$$\begin{aligned}&\left(U/NR_\delta(\{a\}), U/NR_\delta(\{b\}), U/NR_\delta(\{a, b\})\right)\\ &= \big(\ \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6\}\}, \{\{x_1, x_3, x_4, x_5\}, \ldots, \{x_5, x_6\}\},\\ &\quad \{\{x_1, x_3\}, \ldots, \{x_5, x_6\}\}\ \big).\end{aligned}$$

Furthermore, the size $KM(K_B^\delta)$ becomes the arithmetic mean of all 3 values of $G_{\{a\}}^\delta$, $G_{\{b\}}^\delta$, $G_{\{a,b\}}^\delta$ to gain value $\frac{1}{3}(0.5 + 0.5 + 0.3333) = 0.4444$.

The tri-level results regarding $B = \{a, b\}$ come from the relevant definitions, and the last top library and its size need the complex computation. Moreover, the relevant hierarchical relationship can be clearly observed.

Herein, Table 3 is utilized to analyze the size features, especially the granulation monotonicity/non-monotonicity. First, the range $[\frac{1}{|U|}, 1] = [0.1667, 1]$ indeed includes all values. All subsets of A constitute multiple types of covering-knowledge granulation, and thus we easily verify the attribute-granulation monotonicity in Proposition 1. Regarding Proposition 3, the radius-granulation monotonicity can be verified by adding other radii. In contrast, the attribute-granulation non-monotonicity can present only one manifestation by observing all subsets, i.e.,

$$\forall\ \emptyset \subset P \subseteq Q \subseteq A\left[KM(K_P^\delta) > KM(K_Q^\delta)\right].$$

On the other side, we want to provide a further case by showing and verifying Lemmas 1 and 2.

Table 3 can be updated to a new table by replacing c with a_*. Thus, we let $B = \{a, b\} \subset \{a, b\}\cup\{a_*\} = \{a, b, a_*\} = Q$. $2^B - \{\emptyset\} = \{\{a\}, \{b\}, \{a, b\}\}$ and $2^Q - \{\emptyset\} = \{\{a\}, \{b\}, \{a, b\}, \{a, a_*\}, \{b, a_*\}, \{a, b, a_*\}, \{a_*\}\}$, and the latter $2^3 - 1 = 7$ subsets enlarge the former $2^2 - 1 = 3$ subsets to exhibit 3 groups; this result verifies Lemma 1. Regarding Lemma 2, there are two cases.

(1) a_* can enough discretize each sample to become a neighborhood (such as $f(x_i, a_*) = 0.33(i-1)$ $(i = 1, \ldots 6)$), i.e., it

Table 2
All neighborhoods of Table 1.

U	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{b,c\}$	$\{a,c\}$	$\{a,b,c\}$
x_1	$\{x_1,x_2,x_3\}$	$\{x_1,x_3,x_4,x_5\}$	$\{x_1\}$	$\{x_1,x_3\}$	$\{x_1\}$	$\{x_1\}$	$\{x_1\}$
x_2	$\{x_1,x_2,x_3\}$	$\{x_2\}$	$\{x_2,x_3,x_4,x_5,x_6\}$	$\{x_2\}$	$\{x_2\}$	$\{x_2\}$	$\{x_2\}$
x_3	$\{x_1,x_2,x_3\}$	$\{x_1,x_3,x_4\}$	$\{x_2,x_3\}$	$\{x_1,x_3\}$	$\{x_3\}$	$\{x_3\}$	$\{x_3\}$
x_4	$\{x_4,x_5,x_6\}$	$\{x_1,x_3,x_4,x_5\}$	$\{x_2,x_4,x_5,x_6\}$	$\{x_4,x_5\}$	$\{x_4,x_5\}$	$\{x_4,x_5,x_6\}$	$\{x_4,x_5\}$
x_5	$\{x_4,x_5,x_6\}$	$\{x_1,x_4,x_5,x_6\}$	$\{x_2,x_4,x_5,x_6\}$	$\{x_4,x_5,x_6\}$	$\{x_4,x_5,x_6\}$	$\{x_4,x_5,x_6\}$	$\{x_4,x_5,x_6\}$
x_6	$\{x_4,x_5,x_6\}$	$\{x_5,x_6\}$	$\{x_2,x_4,x_5,x_6\}$	$\{x_5,x_6\}$	$\{x_5,x_6\}$	$\{x_4,x_5,x_6\}$	$\{x_5,x_6\}$

Table 3
All size values of tri-level granular structure of Table 1.

Hierarchical subject	Size measure	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{b,c\}$	$\{a,c\}$	$\{a,b,c\}$
Bottom granule	$S(g_B^\delta(x_1))$	0.5	0.6667	0.1667	0.3333	0.1667	0.1667	0.1667
	$S(g_B^\delta(x_2))$	0.5	0.1667	0.8333	0.1667	0.1667	0.1667	0.1667
	$S(g_B^\delta(x_3))$	0.5	0.5	0.3333	0.3333	0.1667	0.1667	0.1667
	$S(g_B^\delta(x_4))$	0.5	0.6667	0.6667	0.3333	0.3333	0.5	0.3333
	$S(g_B^\delta(x_5))$	0.5	0.6667	0.6667	0.5	0.5	0.5	0.5
	$S(g_B^\delta(x_6))$	0.5	0.3333	0.6667	0.3333	0.3333	0.5	0.3333
Middle swarm	$GM(G_B^\delta)$	0.5	0.5	0.5556	0.3333	0.2778	0.3333	0.2778
Top library	$KM(K_B^\delta)$	0.5	0.5	0.5556	0.4444	0.4444	0.4630	0.3968

leads to the finest covering. Thus,

$$KM(K_Q^\delta) = \frac{1}{7}\big(0.4444\times 3 + 0.1667\times 3 + 0.1667\times 1\big) = 0.2857 < 0.4444 = KM(K_B^\delta).$$

(2) When a_* endows all objects with the same value (such as $f(x_i,a_*) = 0.5$ $(i = 1,\ldots,6)$), it causes the coarsest covering. Thus,

$$KM(K_Q^\delta) = \frac{1}{7}\big(0.4444\times 3 + 0.4444\times 3 + 1\times 1\big) = 0.5238 > 0.4444 = KM(K_B^\delta).$$

Therefore, Lemma 2 is verified to be true, and the above two inequations also support the attribute-granulation non-monotonicity in Proposition 3.

Finally, we analyze the logical operations at the three levels, and $B = \{a,b\}$ and $B' = \{b,c\}$ are utilized to focus on some fragments.

(1) At the middle level of granule, $g_B^\delta(x_4) = \{x_4,x_5\}$ and $g_B^\delta(x_6) = \{x_5,x_6\}$, and thus

$$g_B^\delta(x_4)\wedge g_B^\delta(x_6) = \{x_5\}\notin U/n_B^\delta,$$
$$g_B^\delta(x_4)\vee g_B^\delta(x_6) = \{x_4,x_5,x_6\}\in U/n_B^\delta,$$
$$g_B^\delta(x_4) - g_B^\delta(x_6) = \{x_4\}\notin U/n_B^\delta,$$
$$g_B^\delta(x_4)\oplus g_B^\delta(x_6) = \{x_4,x_6\}\notin U/n_B^\delta,$$
$$\neg g_B^\delta(x_4) = \{x_1,x_2,x_3,x_6\}\notin U/n_B^\delta.$$

These results verify the bottom conclusion, as given in Definition 6 and Proposition 4.

(2) At the middle level of swarm, $\mathcal{G}^\delta = \{null, G_{\{a\}}^\delta, G_{\{b\}}^\delta, G_{\{c\}}^\delta, G_{\{a,b\}}^\delta, G_{\{b,c\}}^\delta, G_{\{a,c\}}^\delta, G_{\{a,b,c\}}^\delta\}$, and thus

$$G_B^\delta\wedge G_{B'}^\delta = G_{\{b\}}^\delta\in\mathcal{G}^\delta,\ G_B^\delta\vee G_{B'}^\delta = G_{\{a,b,c\}}^\delta\in\mathcal{G}^\delta,$$
$$G_B^\delta - G_{B'}^\delta = G_{\{a\}}^\delta\in\mathcal{G}^\delta,$$
$$G_B^\delta\oplus G_{B'}^\delta = G_{\{a,c\}}^\delta\in\mathcal{G}^\delta,\ \neg G_B^\delta = G_{\{c\}}^\delta\in\mathcal{G}^\delta.$$

These results verify the middle conclusion, as given in Definition 7 and Proposition 5.

(3) The top case of library is similar to the above item (2).

3. Distance measurement of tri-level granular structure

Distance measurement is fundamental for the uncertainty quantification and system classification, and thus it is emphatically researched for the above tri-level granular structure, including the bottom neighborhood granule, middle neighborhood swarm, and top neighborhood library. Only the bottom and middle levels have acquired their distances in [31]. Next, distances at the two levels will be powerfully enlarged by the double-quantification fusion, while distances at the surplus top level will be further supplemented by hierarchical construction.

3.1. Distance measurement of bottom neighborhood granules

Aiming at bottom-level granules, the relative and absolute distances are recalled to be integrated into a double-quantitative distance. According to the granular essence, the neighborhood granules mainly use their set operations, which equivalently replace the previous style of logical operations [31].

Definition 8 ([31]). In $NS = (U,A,V,f,\delta)$, the relative and absolute distances between two bottom neighborhood granules $s = g_B^\delta(x)$ and $t = g_B^\delta(y)$ are respectively defined by

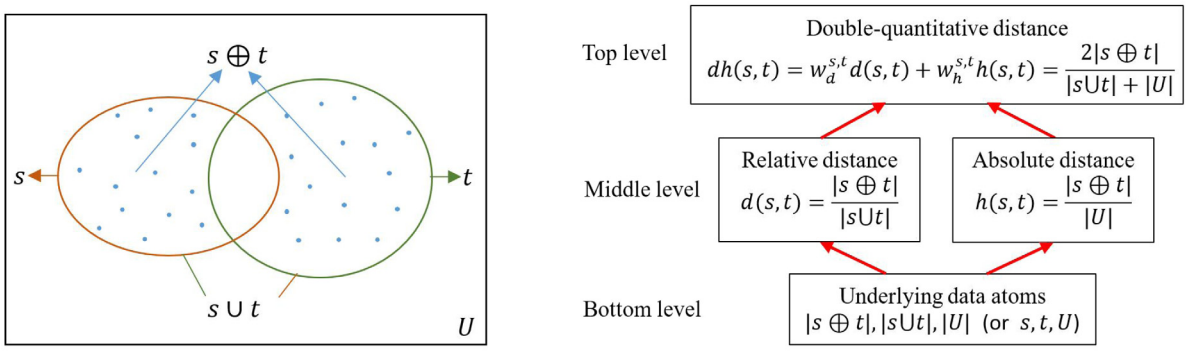
$$d(s,t) = \frac{|s\oplus t|}{|s\cup t|} = \frac{|g_B^\delta(x)\cup g_B^\delta(y) - g_B^\delta(x)\cap g_B^\delta(y)|}{|g_B^\delta(x)\cup g_B^\delta(y)|},$$
$$h(s,t) = \frac{|s\oplus t|}{|U|} = \frac{|g_B^\delta(x)\cup g_B^\delta(y) - g_B^\delta(x)\cap g_B^\delta(y)|}{|U|}. \tag{19}$$

Proposition 6 ([31]). *In $NS = (U,A,V,f,\delta)$, measures $d(s,t)\in[0,1]$ and $h(s,t)\in[0,1]$ conform to the distance axioms regarding non-negativity (identity), symmetry, and triangular inequality. In terms of triviality based on $\emptyset$, U, we have*

$$d(s,\emptyset) = 1,\ \ h(s,\emptyset) = \frac{|s|}{|U|},$$
$$d(s,U) = 1 - \frac{|s|}{|U|},\ \ h(s,U) = 1 - \frac{|s|}{|U|},$$

and thus they represent the neighborhood granule size as follows:

$$S(s) = h(s,\emptyset), S(s) = 1 - d(s,U) = 1 - h(s,U). \tag{20}$$



(a) Bottom information structure (b) Tri-level construction

Fig. 4. Construction mechanism of double-quantitative distance.

The structural framework of information atom regarding system (U, s, t) is described in Fig. 4(a). Here, $s \oplus t = (s - t) \cup (t - s)$ implies the difference and separability between two neighborhood granules s and t, and thus its cardinality $|s \oplus t| = |s - t| + |t - s|$ closely adheres to the granule distance. To fall into the normal metric interval $[0, 1]$, different ranges are chosen as denominators to construct two types of basic distances, as given in Definition 8 and Proposition 6.

(1) $s \oplus t$ has a relative range $s \cup t$, i.e., $s \oplus t$ is in union $s \cup t$ while the latter $s \cup t$ relies on s, t to have the locality and variability. Thus, $d(s, t)$ adopts the proportion between $|s \oplus t|$ and $|s \cup t|$ to correspond to the relative distance and quantification.
(2) $s \oplus t$ has an absolute range $|U|$, i.e., $s \oplus t$ is in universe U while the latter U never depends on s, t to exhibit the globality and stationarity. Thus, $h(s, t)$ adopts the proportion between $|s \oplus t|$ and $|U|$ to embody the absolute distance and quantification.

According to the double-quantification thought (given in Section 1.1 and Fig. 2), a double-quantitative distance is worth mining by integrating $d(s, t)$ and $h(s, t)$, thus providing more powerful measurement and effective applications.

Lemma 3. *The relative and absolute distances have two systematic relationships, i.e.,*

$$\begin{aligned} &\forall s = g_B^{\delta}(x) \in U/NR_{\delta}(B), \forall t = g_B^{\delta}(y) \in U/NR_{\delta}(B), \\ &(1)\ \alpha_d^{s,t} \times d(s,t) + \alpha_h^{s,t} \times h(s,t) = 0 \Longleftrightarrow \alpha_d^{s,t} \equiv 0 \equiv \alpha_h^{s,t}, \\ &(2)\ d(s,t) \geq h(s,t), \end{aligned} \tag{21}$$

where combination coefficients satisfy $\alpha_d^{s,t}, \alpha_h^{s,t} \in \mathbb{R}^+ \cup \{0\}$.

According to Lemma 3, the two distances are linearly independent, and they follow a fixed size relationship. We next construct a good double-quantitative distance by the weighted combination, and the relevant mechanism is clarified as follows, where Fig. 4(b) can provide some auxiliary instructions.

(1) A linear combination serves as a basic operation and an efficient technology for metric fusion, and thus it is chosen to fuse $d(s, t)$ and $h(s, t)$. In terms of measurement, $d(s, t)$ and $h(s, t)$ mean two types of distance measures in metric range $[0, 1]$, but they place emphasis on two dialectical sides of relativity and absoluteness. They have the homogeneity and heteromorphism to offer the linear independence, so their linear combination is feasible.
(2) Regarding two granules, their systematic action area (or their whole function scope) usually goes beyond their union but also is less than the universe, i.e., we can offer a general description: $s \cup t \subseteq Domain(s, t) \subseteq U$. Based on the core information $|s \oplus t|$ related to distance difference semantics, the former $d(s, t)$ adopts a stricter range $s \cup t$ or $|s \cup t|$ to tend to a greater value, while the latter $h(s, t)$ adopts a looser range U or $|U|$ to tend to a less value; this fact can be partly supported by the identical relationship $d(s, t) \geq h(s, t)$. Thus, a compromise based on linear combination is required and valuable. At first, the linear combination can resort to equal and fixed coefficient 0.5 to provide a common and simple statistic, i.e., arithmetic mean $\frac{d(s,t)+h(s,t)}{2} = 0.5 \times d(s, t) + 0.5 \times h(s, t)$. By corresponding adjustments, we next choose distinctive weights to generate a reasonable compromise distance. Concretely, $d(s, t)$ comes from local region $s \cup t$, while $h(s, t)$ is related to global universe U; thus, they respectively use weights $\frac{|s \cup t|}{|s \cup t| + |U|}$ and $\frac{|U|}{|s \cup t| + |U|}$, which respectively reflect the corresponding relativeness and absoluteness.

The above double-quantification development of distance measurement is systematically reflected by Fig. 4, and the relevant tri-level structure of double-quantification concentrate mainly on Fig. 4(b), which accords with the double-quantification framework of Fig. 2.

Definition 9. In $NS = (U, A, V, f, \delta)$, the double-quantitative distance between two bottom neighborhood granules $s = g_B^{\delta}(x)$ and $t = g_B^{\delta}(y)$ is defined by

$$dh(s,t) = w_d^{s,t} \times d(s,t) + w_h^{s,t} \times h(s,t) = \frac{2|s \oplus t|}{|s \cup t| + |U|}, \tag{22}$$

where two weight coefficients submit to

$$w_d^{s,t} = \frac{|s \cup t|}{|s \cup t| + |U|} \in (0,1), \quad w_h^{s,t} = \frac{|U|}{|s \cup t| + |U|} \in (0,1), \quad w_d^{s,t} + w_h^{s,t} = 1. \tag{23}$$

Corollary 1. $dh(s, t) = 2\frac{|s \cup t| - |s \cap t|}{|s \cup t| + |U|}$.

Lemma 4. *If* $0 \leq p \leq q$ *and* $r \geq 0$, *then* $\frac{p}{q} \leq \frac{p+r}{q+r}$.

Proposition 7. *In* $NS = (U, A, V, f, \delta)$, *measure* $dh(s, t)$ *has the metric range* $[0, 1]$, *and it conforms to the distance axioms regarding non-negativity (identity), symmetry, and triangular inequality, i.e.,*

(1) $dh(s, t) \geq 0$, *and* $dh(s, t) = 0$ *if and only if* $s = t$;
(2) $dh(s, t) = dh(t, s)$;
(3) $dh(s, k) + dh(k, t) \geq dh(s, t)$.

In terms of triviality based on $\emptyset, U$, *we have*

$$dh(s, \emptyset) = \frac{2|s|}{|s| + |U|}, dh(s, U) = 1 - \frac{|s|}{|U|},$$

and thus they represent the neighborhood granule size as follows:

$$S(s) = \frac{dh(s, \emptyset)|U|}{2 - dh(s, \emptyset)}, S(s) = 1 - dh(s, U). \tag{24}$$

Proof. According to weight descriptions in Eq. (23), the features of integrated distance $dh(s, t)$ can be derived from corresponding properties of initial distances $d(s, t)$ and $h(s, t)$.

As an example, we prove only the triangular inequality:

$$dh(s, k) + dh(k, t) \geq dh(s, t) \tag{25}$$

in a direct way based on essential analytical formulas. According to Corollary 1 and Lemma 4, we have the equation given in Box I Therefore, the required Eq. (25) holds. □

Proposition 8. *In* $NS = (U, A, V, f, \delta)$*, the double-quantitative distance is located between the relative and absolute distances. That is,*

$$\begin{aligned} &\forall s = g_B^\delta(x) \in U/NR_\delta(B), \forall t = g_B^\delta(y) \in U/NR_\delta(B), \\ &dh(s, t) \in [h(s, t), d(s, t)], \textit{i.e.,} d(s, t) \geq dh(s, t) \geq h(s, t). \end{aligned} \tag{26}$$

Regarding the double-quantitative distance, Definition 9 and Corollary 1 respectively offer the defined and equivalent formulas, while Propositions 7 and 8 respectively provide the metric axiom and size property. According to these discussions, $dh(s, t)$ utilizes distributional weights $w_d^{s,t}, w_h^{s,t}$ to linearly combine two distances $d(s, t), h(s, t)$, and it exactly becomes a distance called as the double-quantitative distance. As a result, $dh(s, t) \in [h(s, t), d(s, t)]$ makes a systematic and skillful compromise of $d(s, t), h(s, t)$, and it becomes a robust measure to better characterize the granule distance by integrating two merits of relative and absolute measurement. Herein, the advance of double-quantitative measurement is basically illustrated by two special cases. Let (s_1, t_1) and (s_2, t_2) be two groups of neighborhood granules.

(1) If $\frac{|s_1 \oplus t_1|}{|s_1 \cup t_1|} = \frac{|s_2 \oplus t_2|}{|s_2 \cup t_2|}$ but $|s_1 \oplus t_1| \neq |s_2 \oplus t_2|$, i.e., $d(s_1, t_1) = d(s_2, t_2)$ but $h(s_1, t_1) \neq h(s_2, t_2)$, then the two granule groups (s_1, t_1) and (s_2, t_2) have the necessary difference but they cannot be effectively discriminated by only the relative measure. Accordingly, $dh(s_1, t_1)$ and $dh(s_2, t_2)$ respectively supplement the absolute measurement to resolve this issue.
(2) The opposite case concerns $|s_1 \oplus t_1| = |s_2 \oplus t_2|$ but $\frac{|s_1 \oplus t_1|}{|s_1 \cup t_1|} \neq \frac{|s_2 \oplus t_2|}{|s_2 \cup t_2|}$, i.e., $h(s_1, t_1) = h(s_2, t_2)$ but $d(s_1, t_1) \neq d(s_2, t_2)$. Thus, (s_1, t_1) and (s_2, t_2) can be effectively identified by not the absolute distance but the double-quantitative distance, and the latter introduces the relative measurement information to improve the former.

In the later studies on classification learning, the double-quantitative distance will be verified to can have optimal performances, where three types of distances are adopted and compared. In summary, the double-quantitative distance $dh(s, t)$ resorts to a weighted combination to act as a fusion measure of relative-quantitative distance $d(s, t)$ and absolute-quantitative distance $h(s, t)$, while the latter two distances are computed by bottom neighborhood granules s, t and their cardinalities; thus, the distance development benefits from the tri-level construction based on information granulation shown in Fig. 4, and $dh(s, t)$ becomes powerful and advanced based on and in contrast to $d(s, t), h(s, t)$, especially in later classification discussions.

3.2. Distance measurement of both middle neighborhood swarms and top neighborhood libraries

The above bottom distances, including the relative, absolute, and double-quantitative types, characterize the difference between neighborhood granules. As shown by Fig. 3, neighborhood granules constitute neighborhood swarms, while the latter further construct neighborhood libraries. Therefore, these granule distances at the bottom level can be naturally promoted to establish distances at middle and top levels, so to evaluate corresponding differences between both swarms and libraries. Next, the middle swarms and top libraries are respectively endowed with three corresponding types of distances, where only the absolute distance at the swarm level has been obtained in [31].

Definition 10. In $NS = (U, A, V, f, \delta)$ with $S, T \subseteq A$, the relative, absolute, and double-quantitative distances between two middle neighborhood swarms $G_S^\delta = \left(g_S^\delta(x_1), g_S^\delta(x_2), \ldots, g_S^\delta(x_{|U|})\right)$ and $G_T^\delta = \left(g_T^\delta(x_1), g_T^\delta(x_2), \ldots, g_T^\delta(x_{|U|})\right)$ are respectively defined by

$$\begin{aligned} D(G_S^\delta, G_T^\delta) &= \frac{1}{|U|}\sum_{i=1}^{|U|} d(g_S^\delta(x_i), g_T^\delta(x_i)), \\ H(G_S^\delta, G_T^\delta) &= \frac{1}{|U|}\sum_{i=1}^{|U|} h(g_S^\delta(x_i), g_T^\delta(x_i)), \\ DH(G_S^\delta, G_T^\delta) &= \frac{1}{|U|}\sum_{i=1}^{|U|} dh(g_S^\delta(x_i), g_T^\delta(x_i)). \end{aligned} \tag{27}$$

Lemma 5. *In* $NS = (U, A, V, f, \delta)$ *with* $S, T \subseteq A$*,*

$$\begin{aligned} &\forall x_i \in U, d(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}], \\ &h(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}], \\ &dh(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}]. \end{aligned} \tag{28}$$

Proof. $\forall x_i \in U$, we have $x_i \in g_S^\delta(x_i)$ and $x_i \in g_T^\delta(x_i)$, so $x_i \in g_S^\delta(x_i) \cap g_T^\delta(x_i) \subseteq g_S^\delta(x_i) \cup g_T^\delta(x_i)$. Thus, $g_S^\delta(x_i) \oplus g_T^\delta(x_i) = g_S^\delta(x_i) \cup g_T^\delta(x_i) - g_S^\delta(x_i) \cap g_T^\delta(x_i)$ is located between $\emptyset$ and $U - \{x_i\}$, and we have $|g_S^\delta(x_i) \oplus g_T^\delta(x_i)| \in [0, |U| - 1]$.

According to relevant definitions, we can obtain $d(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}]$ and $h(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}]$. By definition and property of weighted combination, we further obtain $dh(g_S^\delta(x_i), g_T^\delta(x_i)) \in [0, 1 - \frac{1}{|U|}]$. □

Proposition 9. *The three types of middle distances have the following properties, where* $S, K, T \subseteq A$ *in* $NS = (U, A, V, f, \delta)$*.*

(1) They have the same range $[0, 1 - \frac{1}{|U|}]$*, such as* $DH(G_S^\delta, G_T^\delta) \in [0, 1 - \frac{1}{|U|}]$*.*

(2) They satisfy the distance axioms (i.e., the non-negativity (identity), symmetry, and triangular inequality), such as the double-quantitative case:

$$\begin{aligned} &DH(G_S^\delta, G_T^\delta) \geq 0,\ DH(G_S^\delta, G_T^\delta) = 0 \Longleftrightarrow G_S^\delta = G_T^\delta, \\ &DH(G_S^\delta, G_T^\delta) = DH(G_T^\delta, G_S^\delta), \\ &DH(G_S^\delta, G_K^\delta) + DH(G_K^\delta, G_T^\delta) \geq DH(G_S^\delta, G_T^\delta). \end{aligned} \tag{29}$$

(3) In terms of triviality, we have

$$\begin{aligned} &D(G_S^\delta, G_{null}) = 1, H(G_S^\delta, G_{null}) = \frac{1}{|U|}\sum_{i=1}^{|U|} \frac{|g_S^\delta(x_i)|}{|U|}, \\ &DH(G_S^\delta, G_{null}) = \frac{1}{|U|}\sum_{i=1}^{|U|} \frac{2|g_S^\delta(x_i)|}{|g_S^\delta(x_i)| + |U|}, \\ &D(G_S^\delta, G_{full}) = H(G_S^\delta, G_{full}) = DH(G_S^\delta, G_{full}) = 1 - \frac{1}{|U|}\sum_{i=1}^{|U|} \frac{|g_S^\delta(x_i)|}{|U|}, \end{aligned}$$

$$
\begin{aligned}
&dh(s,k)+dh(k,t)-dh(s,t)\\
=&2\frac{|s\cup k|-|s\cap k|}{|s\cup k|+|U|}+2\frac{|k\cup t|-|k\cap t|}{|k\cup t|+|U|}-2\frac{|s\cup t|-|s\cap t|}{|s\cup t|+|U|}\\
=&2\left[\left(1-\frac{|s\cap k|+|U|}{|s\cup k|+|U|}\right)+\left(1-\frac{|k\cap t|+|U|}{|k\cup t|+|U|}\right)-\left(1-\frac{|s\cap t|+|U|}{|s\cup t|+|U|}\right)\right]\\
=&2\left(1-\frac{|s\cap k|+|U|}{|s\cup k|+|U|}-\frac{|k\cap t|+|U|}{|k\cup t|+|U|}+\frac{|s\cap t|+|U|}{|s\cup t|+|U|}\right)\\
\geq&2\left(1-\frac{|s\cap k|+|U|+(|t|-|t\cap(s\cup k)|)}{|s\cup k|+|U|+(|t|-|t\cap(s\cup k)|)}-\frac{|k\cap t|+|U|+(|s|-|s\cap(k\cup t)|)}{|k\cup t|+|U|+(|s|-|s\cap(k\cup t)|)}+\frac{|s\cap t|+|U|}{|s\cup k\cup t|+|U|}\right)\\
=&2\left(1-\frac{|s\cap k|+|t|-(|s\cap t|+|k\cap t|-|s\cap k\cap t|)+|U|}{|s\cup k\cup t|+|U|}-\frac{|k\cap t|+|s|-(|s\cap k|+|s\cap t|-|s\cap k\cap t|)+|U|}{|s\cup k\cup t|+|U|}+\frac{|s\cap t|+|U|}{|s\cup k\cup t|+|U|}\right)\\
=&2\frac{|s\cup k\cup t|-|s\cap k|-|t|+|s\cap t|+|k\cap t|-|s\cap k\cap t|-|k\cap t|-|s|+|s\cap k|+|s\cap t|-|s\cap k\cap t|+|s\cap t|}{|s\cup k\cup t|+|U|}\\
=&2\frac{|s\cup k\cup t|-(|s|+|t|)+3|s\cap t|-2|s\cap k\cap t|}{|s\cup k\cup t|+|U|}\\
=&2\frac{|s\cup k\cup t|-(|s\cup t|+|s\cap t|)+3|t\cap s|-2|s\cap t\cap k|}{|s\cup k\cup t|+|U|}\\
=&2\left(\frac{|s\cup k\cup t|-|s\cup t|}{|s\cup k\cup t|+|U|}+2\frac{|s\cap t|-|s\cap k\cap t|}{|s\cup k\cup t|+|U|}\right)\\
\geq&0.
\end{aligned}
$$

Box I.

and thus they represent the neighborhood swarm size as follows:

$$
\begin{aligned}
&GM(G_S^\delta)=H(G_S^\delta,G_{null}),\\
&GM(G_S^\delta)=1-D(G_S^\delta,G_{full})=1-H(G_S^\delta,G_{full})=1-DH(G_S^\delta,G_{full}),
\end{aligned}
\tag{30}
$$

where $G_{null}=(\emptyset,\emptyset,\ldots,\emptyset)$ *and* $G_{full}=(U,U,\ldots,U)$.

Proof. (1) This result directly comes from the basic range $[0,1-\frac{1}{|U|}]$ in Lemma 5 and average operation $\frac{1}{|U|}\sum_{i=1}^{|U|}$ in Definition 10.

(2) The distance axioms of middle swarms can be derived from the distance axioms and average integration of bottom granules. As an example, By Proposition 7 and Definition 10, we can deduce triangular inequality regarding the double-quantitative distance, i.e.,

$$
\begin{aligned}
&DH(G_S^\delta,G_K^\delta)+DH(G_K^\delta,G_T^\delta)\\
=&\frac{1}{|U|}\sum_{i=1}^{|U|}dh(g_S^\delta(x_i),g_K^\delta(x_i))+\frac{1}{|U|}\sum_{i=1}^{|U|}dh(g_K^\delta(x_i),g_T^\delta(x_i))\\
=&\frac{1}{|U|}\sum_{i=1}^{|U|}\left[dh(g_S^\delta(x_i),g_K^\delta(x_i))+dh(g_K^\delta(x_i),g_T^\delta(x_i))\right]\\
\geq&\frac{1}{|U|}\sum_{i=1}^{|U|}dh(g_S^\delta(x_i),g_T^\delta(x_i))\\
=&DH(G_S^\delta,G_T^\delta).
\end{aligned}
$$

(3) This triviality assertion can be derived from the corresponding pairs in Propositions 6 and 7, the average function $\frac{1}{|U|}\sum_{i=1}^{|U|}$ in Definition 10, as well as the size formula in Eq. (7). □

Corollary 2. *At the middle swarm level, the relative, absolute, and double-quantitative distances follow an inequation, i.e.,*

$$
D(G_S^\delta,G_T^\delta)\geq DH(G_S^\delta,G_T^\delta)\geq H(G_S^\delta,G_T^\delta).
\tag{31}
$$

According to the bottom-middle construction from granules to swarms, the middle distances are established by hierarchically integrating bottom distances of all samples, as given in Definition 10; their arithmetic mean form $\frac{1}{|U|}\sum_{i=1}^{|U|}$ provides the good fusion mechanism and clear measurement semantics, so they effectively characterize the difference between swarms. Of course, they have three types of different viewpoints and forms. By quantification delivery or feature heredity, the middle single-quantitative and double-quantitative distances respectively adhere to the relativity/absoluteness and completeness, and they also have similar but developmental properties, as shown in Proposition 9 and Corollary 2. In particular, the absolute distance and its properties have been given in [31], so the relative and double-quantitative distances become new contents for perfection. In terms of system, the relative distance is parallel and symmetrical to the absolute distance, while the comprehensive double-quantitative distance conducts a double-quantification compromise.

Thus far, the relative, absolute, and double-quantitative distances are offered at the bottom granule level, and then the three types are hierarchically promoted to the middle swarm level by the average integration, where two arbitrary neighborhood swarms adhere to a one-to-one correspondence due to the same granular dimensionality regarding $|U|$. Now, we sequentially consider the systematic distances at the library level by the middle-top integration. However, two neighborhood libraries may have different numbers of internal swarms, so the correspondence complexity causes a difficulty of definition. Next, a special case of number correspondence is preliminarily considered.

Lemma 6. *Let* $S,T\subseteq A$. *If* $|S|=|T|$, *then* $2^S-\{\emptyset\}$ *and* $2^T-\{\emptyset\}$ *follow a one-to-one correspondence regarding non-empty subsets, and this correspondence can be concretized by the complete lattice isomorphism* $(2^S,\cap,\cup)\cong(2^T,\cap,\cup)$, *which is derived from a one-to-one correspondence by sorting* S *and* T.

Definition 11. In $NS = (U, A, V, f, \delta)$ with $S, T \subseteq A$, suppose $|S| = |T|$ and denote $m' = 2^{|S|}-1 = 2^{|T|}-1$. The relative, absolute, and double-quantitative distances between top neighborhood libraries $K_S^\delta = \left(G_{S_1}^\delta, g_{S_2}^\delta, \ldots, G_{S_{m'}}^\delta\right)$ and $K_T^\delta = \left(G_{T_1}^\delta, G_{T_2}^\delta, \ldots, G_{T_{m'}}^\delta\right)$ are respectively defined by

$$\begin{aligned} D(K_S^\delta, K_T^\delta) &= \frac{1}{m'}\sum_{j=1}^{m'} D(G_{S_j}^\delta, G_{T_j}^\delta), \\ H(K_S^\delta, K_T^\delta) &= \frac{1}{m'}\sum_{j=1}^{m'} H(G_{S_j}^\delta, G_{T_j}^\delta), \\ DH(K_S^\delta, K_T^\delta) &= \frac{1}{m'}\sum_{j=1}^{m'} DH(G_{S_j}^\delta, G_{T_j}^\delta), \end{aligned} \tag{32}$$

where $\{S_1, S_2, \ldots, S_{m'}\} = 2^S - \{\emptyset\}$ and $\{T_1, T_2, \ldots, T_{m'}\} = 2^T - \{\emptyset\}$ adopt the one-to-one correspondence in Lemma 6.

Corollary 3. *The three types of top distances have the following properties, where $S, K, T \subseteq A$, $|S| = |K| = |T|$, $m' = 2^{|S|} - 1 = 2^{|K|} - 1 = 2^{|T|} - 1$ in $NS = (U, A, V, f, \delta)$.*

(1) They have the same range $[0, 1 - \frac{1}{|U|}]$.

(2) They satisfy the distance axioms (i.e., the non-negativity (identity), symmetry, and triangular inequality), such as the double-quantitative case:

$$\begin{aligned} &DH(K_S^\delta, K_T^\delta) \geq 0,\ DH(K_S^\delta, K_T^\delta) = 0 \Longleftrightarrow K_S^\delta = K_T^\delta, \\ &DH(K_S^\delta, K_T^\delta) = DH(K_T^\delta, K_S^\delta), \\ &DH(K_S^\delta, K_K^\delta) + DH(K_K^\delta, K_T^\delta) \geq DH(K_S^\delta, K_T^\delta). \end{aligned} \tag{33}$$

(3) They represent the neighborhood library size as follows:

$$\begin{aligned} &KM(K_S^\delta) = H(K_S^\delta, K_{null}), \\ &KM(K_S^\delta) = 1 - D(K_S^\delta, K_{full}) = 1 - H(K_S^\delta, K_{full}) = 1 - DH(K_S^\delta, K_{full}), \end{aligned} \tag{34}$$

where $K_{null} = \left((\emptyset, \emptyset, \ldots, \emptyset), \ldots, (\emptyset, \emptyset, \ldots, \emptyset)\right)$ and $K_{full} = \left((U, U, \ldots, U), \ldots, (U, U, \ldots, U)\right)$.

Corollary 4. *At the top library level, the relative, absolute, and double-quantitative distances follow an inequation, i.e.,*

$$D(K_S^\delta, K_T^\delta) \geq DH(K_S^\delta, K_T^\delta) \geq H(K_S^\delta, K_T^\delta), \tag{35}$$

where $|S| = |T|$.

In Definition 11, we characterize the library difference by virtue of the same swarm scale (i.e., by virtue of the same cardinality of attribute subset), and the one-to-one correspondence mainly adopts the strategy of Lemma 6 where the set sorting needs to be provided in advance. Thus, the single-quantitative and double-quantitative distances naturally utilize the average integration regarding all internal swarms, and they transmit the middle distance features, as given by Corollaries 3 and 4. As an example, the top size valuation utilizes the same average structure (Eq. (9)), so its relationship with the top distance also directly promotes the previous connection at the middle level (Eq. (30)). Moreover, the general distance studies on different swarm numbers (i.e., $|S| \neq |T|$) remain to become a new question.

3.3. Illustrative example II

Herein, an example is given to illustrate the distance measurement, including the tri-level and double-quantification measurement.

Example 2. The previous Example 1 (with radius $\delta = 0.3$) is unceasingly considered for distance measurement.

(1) At the granule level, representative subset $B = \{a, b\} \subset A$ is still chosen for calculation and verification, and we focus on three granules:

$$s = g_B^\delta(x_1) = \{x_1, x_3\},\ k = g_B^\delta(x_2) = \{x_2\},\ t = g_B^\delta(x_3) = \{x_1, x_3\}.$$

Thus, the relative, absolute, and double-quantitative distances offer

$$\begin{aligned} &d(s,k) = \frac{|\{x_1,x_3\} - \{x_2\}| + |\{x_2\} - \{x_1,x_3\}|}{|\{x_1,x_3\} \cup \{x_2\}|} = \frac{3}{3} = 1, \\ &d(k,t) = 1,\ d(s,t) = 0, \\ &h(s,k) = \frac{|\{x_1,x_3\} - \{x_2\}| + |\{x_2\} - \{x_1,x_3\}|}{|U|} = \frac{3}{6} = 0.5, \\ &h(k,t) = 0.5,\ h(s,t) = 0, \\ &dh(s,k) = 2\frac{|\{x_1,x_3\} - \{x_2\}| + |\{x_2\} - \{x_1,x_3\}|}{|\{x_1,x_3\} \cup \{x_2\}| + |U|} = \frac{6}{9} = 0.6667, \\ &dh(k,t) = 0.6667,\ dh(s,t) = 0, \end{aligned}$$

where weighted coefficients become

$$\begin{aligned} &(w_d^{s,k}, w_h^{s,k}) = (\frac{3}{9}, \frac{6}{9}) = (0.3333, 0.6667), \\ &(w_d^{k,t}, w_h^{k,t}) = (0.3333, 0.6667), \\ &(w_d^{s,t}, w_h^{s,t}) = (0.25, 0.75). \end{aligned}$$

These results accord with the range [0, 1], triangular inequalities:

$$\begin{aligned} d(s,k) + d(k,t) \geq d(s,t),\ h(s,k) + h(k,t) \geq h(s,t), \\ dh(s,k) + dh(k,t) \geq dh(s,t), \end{aligned}$$

and the relationship inequation:

$$\begin{aligned} d(s,k) \geq dh(s,k) \geq h(s,k),\ d(k,t) \geq dh(k,t) \geq h(k,t), \\ d(s,t) \geq dh(s,t) \geq h(s,t). \end{aligned}$$

(2) At the swarm level, consider three attribute subsets $B' = \{a\}$, $B = \{a, b\}$, $A = \{a, b, c\}$. They generate three middle neighborhood swarms related to the hierarchical structure of bottom granules, i.e.,

$$\begin{aligned} G_{B'}^\delta &= (\{x_1,x_2,x_3\}, \{x_1,x_2,x_3\}, \{x_1,x_2,x_3\}, \{x_4,x_5,x_6\}, \\ &\quad \{x_4,x_5,x_6\}, \{x_4,x_5,x_6\}), \\ G_B^\delta &= (\{x_1,x_3\}, \{x_2\}, \{x_1,x_3\}, \{x_4,x_5\}, \{x_4,x_5,x_6\}, \{x_5,x_6\}), \\ G_A^\delta &= (\{x_1\}, \{x_2\}, \{x_3\}, \{x_4,x_5\}, \{x_4,x_5,x_6\}, \{x_5,x_6\}), \end{aligned}$$

and each one resorts to its neighborhood covering to exhibit a one-to-one correspondence to universe

$$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}.$$

Thus, the relative, absolute, and double-quantitative distances become

$$\begin{aligned} &D(G_{B'}^\delta, G_B^\delta) = \frac{1}{6}\sum_{i=1}^{6} d(g_{B'}^\delta(x_i), g_B^\delta(x_i)) = 0.3333, \\ &D(G_B^\delta, G_A^\delta) = 0.1667,\ D(G_{B'}^\delta, G_A^\delta) = 0.4444, \\ &H(G_{B'}^\delta, G_B^\delta) = \frac{1}{6}\sum_{i=1}^{6} h(g_{B'}^\delta(x_i), g_B^\delta(x_i)) = 0.1667, \\ &H(G_B^\delta, G_A^\delta) = 0.0278,\ H(G_{B'}^\delta, G_A^\delta) = 0.2222, \\ &DH(G_{B'}^\delta, G_B^\delta) = \frac{1}{6}\sum_{i=1}^{6} dh(g_{B'}^\delta(x_i), g_B^\delta(x_i)) = 0.2222, \\ &DH(G_B^\delta, G_A^\delta) = 0.0833,\ DH(G_{B'}^\delta, G_A^\delta) = 0.2963, \end{aligned}$$

and the hierarchical calculation regarding arithmetic mean is clarified in Table 4.

These results can verify all properties of swarm distances. For example, they embody the theoretical range $[0, 1-\frac{1}{|U|}] = [0, 0.8333]$, the triangular inequalities based on

$$D(G^\delta_{B'}, G^\delta_B) + D(G^\delta_B, G^\delta_A) \geq D(G^\delta_{B'}, G^\delta_A),$$
$$H(G^\delta_{B'}, G^\delta_B) + H(G^\delta_B, G^\delta_A) \geq H(G^\delta_{B'}, G^\delta_A),$$
$$DH(G^\delta_{B'}, G^\delta_B) + DH(G^\delta_B, G^\delta_A) \geq DH(G^\delta_{B'}, G^\delta_A),$$

and the relationship inequation based on

$$D(G^\delta_{B'}, G^\delta_B) \geq DH(G^\delta_{B'}, G^\delta_B) \geq H(G^\delta_{B'}, G^\delta_B),$$
$$D(G^\delta_B, G^\delta_A) \geq DH(G^\delta_B, G^\delta_A) \geq H(G^\delta_B, G^\delta_A),$$
$$D(G^\delta_{B'}, G^\delta_A) \geq DH(G^\delta_{B'}, G^\delta_A) \geq H(G^\delta_{B'}, G^\delta_A).$$

(3) At the library level, we offer $S = \{a, b\}$, $K = \{b, c\}$, $T = \{a, c\}$ with the same set cardinality 2 and non-empty subset number $m' = 3$. The above natural order and correspondence are chosen for S, K, T, and thus Lemma 6 generates a one-to-one correspondence:

$$2^S = \{\{a\}, \{b\}, \{a, b\}\} \longleftrightarrow 2^K = \{\{b\}, \{c\}, \{b, c\}\}$$
$$\longleftrightarrow 2^T = \{\{a\}, \{c\}, \{a, c\}\}.$$

Thus, S, K, T produce 3 neighborhood libraries with corresponding 9 swarms:

$$K^\delta_S = (G^\delta_{\{a\}}, G^\delta_{\{b\}}, G^\delta_{\{a,b\}}), K^\delta_K = (G^\delta_{\{b\}}, G^\delta_{\{c\}}, G^\delta_{\{b,c\}}),$$
$$K^\delta_T = (G^\delta_{\{a\}}, G^\delta_{\{c\}}, G^\delta_{\{a,c\}}).$$

The relative, absolute, and double-quantitative distances become

$$D(K^\delta_S, K^\delta_K) = 0.4343, D(K^\delta_K, K^\delta_T) = 0.2009, D(K^\delta_S, K^\delta_T) = 0.3074,$$
$$H(K^\delta_S, K^\delta_K) = 0.2963, H(K^\delta_K, K^\delta_T) = 0.1296, H(K^\delta_S, K^\delta_T) = 0.2037,$$
$$DH(K^\delta_S, K^\delta_K) = 0.3453, DH(K^\delta_K, K^\delta_T) = 0.1557,$$
$$DH(K^\delta_S, K^\delta_T) = 0.2390.$$

As an explanation, we provide the relevant formulas for the double-quantitative part, i.e.,

$$\begin{aligned} DH(K^\delta_S, K^\delta_K) &= \frac{1}{m'}\sum_{j=1}^{n} DH(G^\delta_{S_j}, G^\delta_{K_j}) \\ &= \frac{1}{3}\left[DH(G^\delta_{\{a\}}, G^\delta_{\{b\}}) + DH(G^\delta_{\{b\}}, G^\delta_{\{c\}}) + DH(G^\delta_{\{a,b\}}, G^\delta_{\{b,c\}})\right] \\ &= 0.3453, \end{aligned}$$

$$\begin{aligned} DH(K^\delta_K, K^\delta_T) &= \frac{1}{m'}\sum_{j=1}^{n} DH(G^\delta_{K_j}, G^\delta_{T_j}) \\ &= \frac{1}{3}\left[DH(G^\delta_{\{a\}}, G^\delta_{\{b\}}) + DH(G^\delta_{\{c\}}, G^\delta_{\{c\}}) + DH(G^\delta_{\{b,c\}}, G^\delta_{\{a,c\}})\right] \\ &= 0.1557, \end{aligned}$$

$$\begin{aligned} DH(K^\delta_S, K^\delta_T) &= \frac{1}{m'}\sum_{j=1}^{n} DH(G^\delta_{S_j}, G^\delta_{T_j}) \\ &= \frac{1}{3}\left[DH(G^\delta_{\{a\}}, G^\delta_{\{a\}}) + DH(G^\delta_{\{b\}}, G^\delta_{\{c\}}) + DH(G^\delta_{\{a,b\}}, G^\delta_{\{a,c\}})\right] \\ &= 0.2390. \end{aligned}$$

More details of hierarchical calculation regarding arithmetic mean are given in Table 5. These results effectively verify those properties of top library distances, such as the range, the triangular inequality, and the relationship inequation.

4. Classification learning based on the double-quantitative distance of neighborhood granules

In [31], the above relative and absolute distances of bottom neighborhood granules are utilized to implement classification learning, and they respectively induce classifiers KNGR and KNGA to outperform the traditional classifier KNN, which denotes the K-nearest neighbor classifier based on the Euclidean distance [66]. In this section, the above double-quantitative distance of granules is naturally adopted to further focus on classification learning, and thus it produces a new classifier KNGD, i.e., a K-nearest neighbor classifier based on the double-quantitative distance. The better or compromised classification achievement of classifier KNGD (in contrast to classifiers KNGR and KNGA) will be finally verified by data experiments.

4.1. Classification algorithm design

In machine learning, classifier KNN contains a simple but mature strategy for classification tasks. Concretely, for a sample, if its K nearest neighbors in the feature space belong to a certain category, then it also belongs to this category. On this basis, a general algorithm framework can be generated by introducing the granulation and distance, and it can use different types of distances to construct multiple classifiers. For comparison and generation, we provide the following computing procedure by referring to that in [31]. The relevant algorithm framework is given in Fig. 5, and there are 7 steps including (1) data preprocessing, (2) training–testing division, (3) training set granulation, (4) testing set granulation, (5) K-nearest neighbor discriminant, (6) testing sample determination, (7) discriminant-determination circulation.

(1) Remove data related to missing values, and make the data normalization to reach the standard range [0, 1].
(2) Divide the normal data into a training set and a testing set. The percentage of both parts can be artificially set up, and we adopt a special value 4 : 1.
(3) Granulate the training set, and generate training granules by setting up the neighborhood radius.
(4) Granulate the testing set, and each testing sample resorts to its Euclidean distance regarding the training objects and radius threshold to generate a testing granule.
(5) Aiming at a testing object, compute its granular distance with each training granule (by a specific distance function). By the total order of distance values, acquire the K-nearest neighbors from the training set.
(6) Endow the testing object with the optimal category label by virtue of the maximal distribution of the K-nearest neighbors.
(7) Completely classify all testing samples. Concretely, turn back to Step 5 to circularly classify the surplus testing object.

The above calculation process mainly includes four parts of preprocessing, granulation, matching, and classification, and they correspond to Step (1), Steps (2)–(4), Step (5), Step (6), respectively. This classification method refers to the classical thought of KNN, and it can use different distance functions to generate different classifiers. As a result, the relative (absolute) distance is utilized to give classifier KNGR (KNGA) in [31], while the double-quantitative distance naturally motivates a classifier noted as KNGD. Regarding algorithm parameters, there are only two, i.e., neighborhood radius δ and nearest-neighbor number K.

As a detailed example, Algorithm 1 provides the pseudo-code description of classifier KNGD by concretizing relevant parts of Fig. 5. Steps 3–10 complete the training set granulation by Euclidean distance D_C, where the training granule $n^\delta_C(x_{Train})|_{TrainSet}$

Table 4
Hierarchical construction of middle swarm distances regarding three groups of swarms.

	(B', B)			(B, A)			(B', A)		
U	d	h	dh	d	h	dh	d	h	dh
x_1	0.3333	0.1667	0.2222	0.5	0.1667	0.25	0.6667	0.3333	0.4444
x_2	0.6667	0.3333	0.4444	0	0	0	0.6667	0.3333	0.4444
x_3	0.3333	0.1667	0.2222	0.5	0.1667	0.25	0.6667	0.3333	0.4444
x_4	0.3333	0.1667	0.2222	0	0	0	0.3333	0.1667	0.2222
x_5	0	0	0	0	0	0	0	0	0
x_6	0.3333	0.1667	0.2222	0	0	0	0.3333	0.1667	0.2222
Average	$D = \overline{d}$: 0.3333	$H = \overline{h}$: 0.1667	$DH = \overline{dh}$: 0.2222	$D = \overline{d}$: 0.1667	$H = \overline{h}$: 0.0278	$DH = \overline{dh}$: 0.0833	$D = \overline{d}$: 0.4444	$H = \overline{h}$: 0.2222	$DH = \overline{dh}$: 0.2963

Table 5
Hierarchical construction of top library distances regarding three groups of libraries.

	(S, K)				(K, T)				(S, T)			
Distance	$(\{a\}, \{b\})$	$(\{b\}, \{c\})$	$(\{a, b\}, \{b, c\})$	Average	$(\{a\}, \{b\})$	$(\{c\}, \{c\})$	$(\{b, c\}, \{a, c\})$	Average	$(\{a\}, \{a\})$	$(\{b\}, \{c\})$	$(\{a, b\}, \{a, c\})$	Average
D	0.4917	0.6	0.1667	$\overline{D}$: 0.4343	0.4917	0	0.1111	$\overline{D}$: 0.2009	0	0.6	0.2778	$\overline{D}$: 0.3074
H	0.3333	0.5	0.0556	$\overline{H}$: 0.2963	0.3333	0	0.0556	$\overline{H}$: 0.1296	0	0.5	0.1111	$\overline{H}$: 0.2037
DH	0.3929	0.5596	0.0833	$\overline{DH}$: 0.3453	0.3929	0	0.0741	$\overline{DH}$: 0.1557	0	0.5596	0.1574	$\overline{DH}$: 0.2390

Algorithm 1 Classification algorithm of KNGD

Input: Neighborhood system $NS = (U, A, V, f, \delta)$ with radius δ and nearest-neighbor parameter K.
Output: Classification of testing set.

1: Data preprocessing.
2: Make an object division to establish a training set *TrainSet* and a testing set *TestSet*.
3: **for** $x_{Train} \in TrainSet$ **do**
4: $n_C^{\delta}(x_{Train})|_{TrainSet} = \emptyset$.
5: **for** $y_{Train} \in TrainSet$ **do**
6: **if** $D_C(x_{Train}, y_{Train}) \leq \delta$ **then**
7: $n_C^{\delta}(x_{Train})|_{TrainSet} \leftarrow n_C^{\delta}(x_{Train})|_{TrainSet} \cup \{y_{Train}\}$.
8: **end if**
9: **end for**
10: **end for**
11: **for** $x_{Test} \in TestSet$ **do**
12: $n_C^{\delta}(x_{Test})|_{TrainSet} = \emptyset$.
13: **for** $x_{Train} \in TrainSet$ **do**
14: **if** $D_C(x_{Test}, x_{Train}) \leq \delta$ **then**
15: $n_C^{\delta}(x_{Test})|_{TrainSet} \leftarrow n_C^{\delta}(x_{Test})|_{TrainSet} \cup \{x_{Train}\}$.
16: **end if**
17: **end for**
18: **end for**
19: **for** $x_{Test} \in TestSet$ **do**
20: Set up an initial distance-queue $Q(x_{Test}) = null$.
21: **for** $x_{Train} \in TrainSet$ **do**
22: By Eq. (22), calculate the granular double-quantitative distance:

$$dh(n_C^{\delta}(x_{Test})|_{TrainSet}, n_C^{\delta}(x_{Train})|_{TrainSet}) = \frac{2\left|n_C^{\delta}(x_{Test})|_{TrainSet} \oplus n_C^{\delta}(x_{Train})|_{TrainSet}\right|}{\left|n_C^{\delta}(x_{Test})|_{TrainSet} \cup n_C^{\delta}(x_{Train})|_{TrainSet}\right| + \left|U\right|}. \tag{36}$$

23: The above distance $dh(n_C^{\delta}(x_{Test})|_{TrainSet}, n_C^{\delta}(x_{Train})|_{TrainSet})$ is added to update list $Q(x_{Test})$ by relevant sorting or positioning.
24: **end for**
25: According to the final distance-list $Q(x_{Test})$, extract the smallest K distances and their corresponding neighbors (i.e., training granules), and the decision label with maximum distribution regarding K nearest neighbors is given for sample x_{Test}.
26: **end for**
27: **return** Corresponding category labels for all testing objects.

becomes $\{y_{Train} \in TrainSet | D_C(x_{Train}, y_{Train}) \leq \delta\} = n_C^{\delta}(x_{Train}) \cap TrainSet$. Steps 11–18 realize the testing set granulation, where $n_C^{\delta}(x_{Test})|_{TrainSet} = \{x_{Train} \in TrainSet | D_C(x_{Test}, x_{Train}) \leq \delta\} = n_C^{\delta}(x_{Test}) \cap TrainSet$. In Steps 19–26, a testing sample x_{Test} is embraced, and its granular distances with all training granules induce its K nearest neighbors, where Eq. (36) is the double-quantitative distance from Eq. (22); thus, x_{Test} is endowed with the category label by considering the maximum distribution of

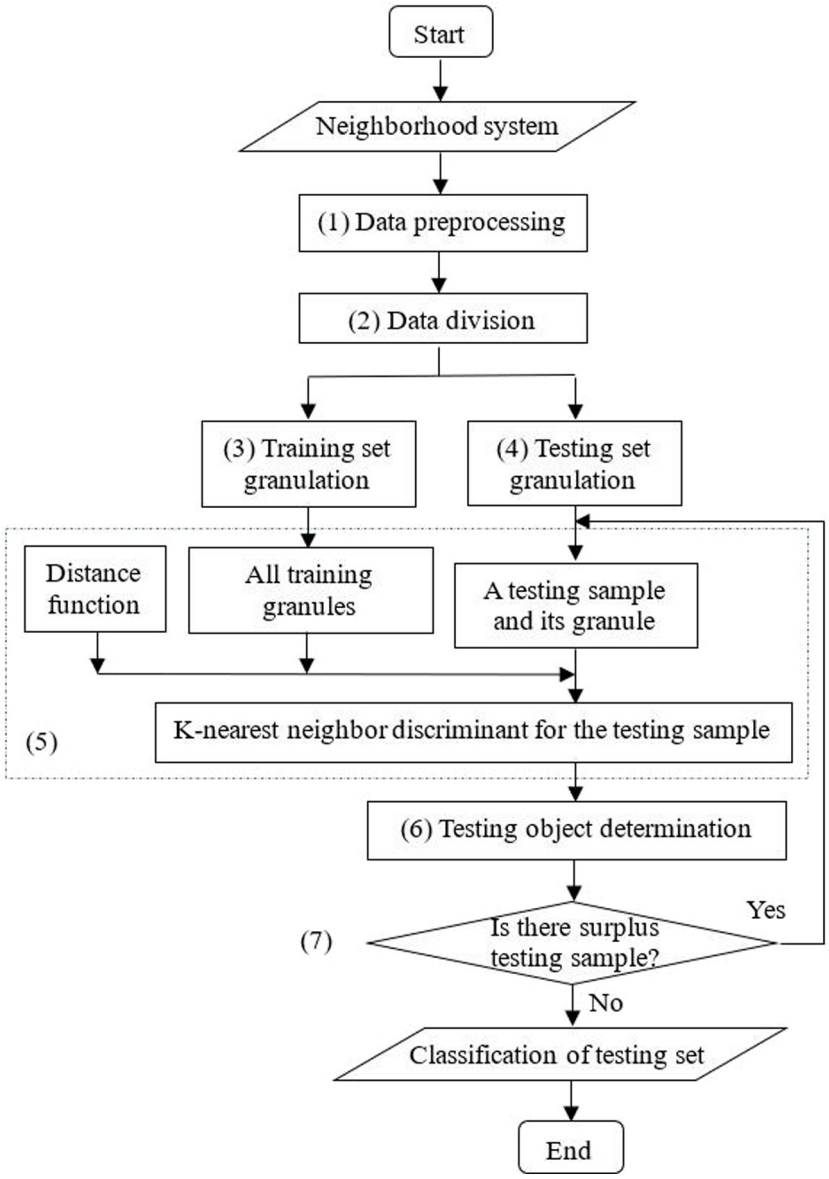


Fig. 5. Algorithm framework of relevant classifiers related to K-nearest neighbor.

the above K nearest neighbors. In particular, Algorithm 1 can motivate the other two classifiers of KNGR and KNGA by alternatively using the single-quantitative distances (Eq. (19)) in Step 22, and it can simultaneously provide the three classifiers of KNGR, KNGA, KNGD if the three related types of quantitative distance are parallelly adopted in Step 22.

4.2. Data experiments

To verify the effectiveness and advantage of classifier KNGD, six UCI datasets are chosen to make classification experiments, i.e., Fertility, Wdbc, Wine, Lymphography, SolarFlare, Cmc [67], and their basic information is presented in Table 6. The classification ability serves as a main estimation index in theory and applications, so it is still utilized for comparison and revelation. According to the two classifier parameters, we next design sufficient experiments by three cases changing only δ, only K, and the two parameters, respectively.

First consider experiments impacted by only radius δ. We use a radius-increase sequence from 0.05 to 1 with step length 0.05 and step number 20, i.e.,

$$\delta : 0.05 \to 0.1 \to 0.15 \to \cdots \to 0.95 \to 1. \tag{37}$$

On the other hand, fixed nearest-neighbor number K is determined by experiments, and its concrete value respectively becomes 5, 5, 1, 10, 5, 5 for the six datasets.

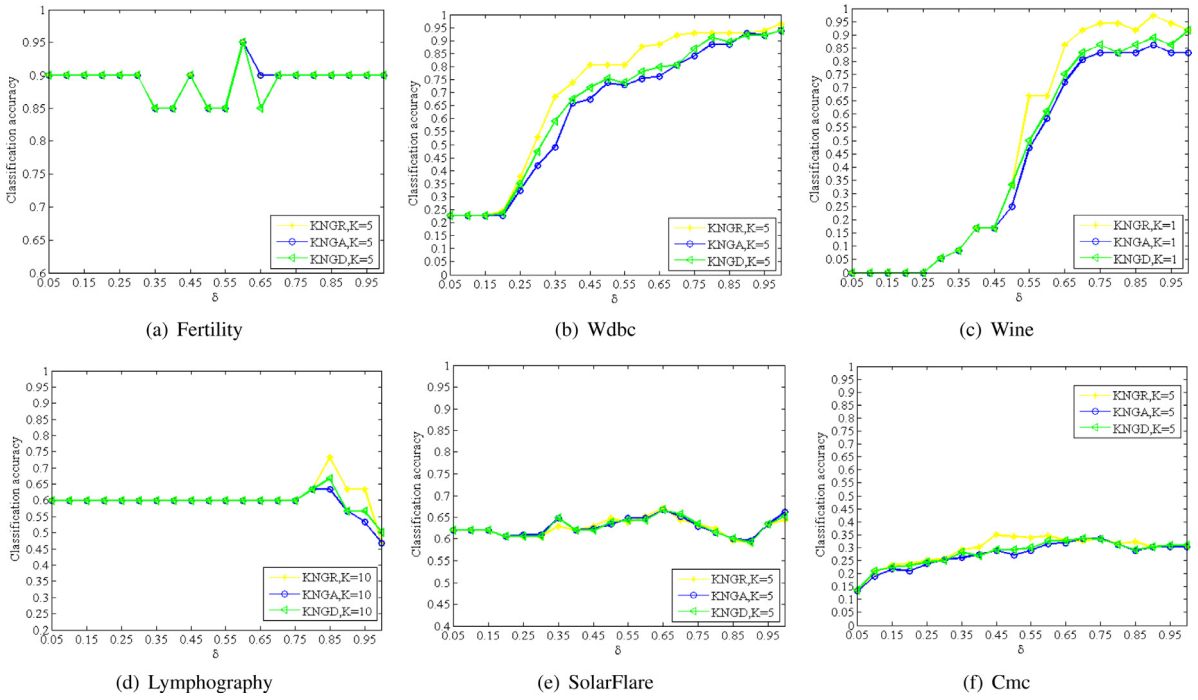
Regarding the six datasets, relevant classification accuracy values based on classifiers KNGR, KNGA, KNGD are given in Table 7, and corresponding two-dimensional figures are shown in Fig. 6. Regarding (1) Fertility, KNGR, KNGA, KNGD acquire the almost coincident effect, except when $\delta = 0.65$. Regarding (2) Wdbc, (3) Wine, (4) Lymphography, the three algorithms exhibit some common phenomena; KNGR, KNGA, KNGD exhibit the same accuracy when δ is small, and for the consistency, the three datasets respectively correspond to radius intervals $\delta \in [0.05, 0.2]$, $\delta \in [0.05, 0.45]$, $\delta \in [0.05, 0.8]$; when δ is great, KNGR is usually superior to KNGA, while KNGD reaches a midvalue to approach KNGR and outperform KNGA, and the three algorithms present the significant difference for accuracy. Regarding surplus (5) SolarFlare and (6) Cmc, the three algorithms

Table 6
Basic descriptions of six UCI datasets.

No.	Name or abbreviation	Sample number	Condition attribute number	Decision class number
(1)	Fertility	100	9	2
(2)	Wdbc	569	31	2
(3)	Wine	178	13	3
(4)	Lymphography	148	18	4
(5)	Solar Flare (SolarFlare)	1066	9	6
(6)	Contraceptive Method Choice (Cmc)	1473	9	3

Table 7
Classification accuracy with changed δ and fixed K.

Dataset (Parameter K)	Algorithm	$\delta = 0.05$	0.1	0.15	0.2	$\cdots$	0.7	0.75	0.8	0.85	0.9	0.95	$\delta = 1$
(1) Fertility ($K = 5$)	KNGR	0.9000	0.9000	0.9000	0.9000	$\cdots$	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
	KNGA	0.9000	0.9000	0.9000	0.9000	$\cdots$	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
	KNGD	0.9000	0.9000	0.9000	0.9000	$\cdots$	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
(2) Wdbc ($K = 5$)	KNGR	0.2281	0.2281	0.2281	0.2456	$\cdots$	0.9211	0.9298	0.9298	0.9298	0.9298	0.9386	0.9649
	KNGA	0.2281	0.2281	0.2281	0.2281	$\cdots$	0.8070	0.8421	0.8860	0.8860	0.9298	0.9211	0.9386
	KNGD	0.2281	0.2281	0.2281	0.2368	$\cdots$	0.8070	0.8684	0.9123	0.8947	0.9211	0.9211	0.9386
(3) Wine ($K = 1$)	KNGR	0.0000	0.0000	0.0000	0.0000	$\cdots$	0.9167	0.9444	0.9444	0.9167	0.9722	0.9444	0.9167
	KNGA	0.0000	0.0000	0.0000	0.0000	$\cdots$	0.8056	0.8333	0.8333	0.8333	0.8611	0.8333	0.8333
	KNGD	0.0000	0.0000	0.0000	0.0000	$\cdots$	0.8333	0.8611	0.8333	0.8611	0.8889	0.8611	0.9167
(4) Lymphography ($K = 10$)	KNGR	0.6000	0.6000	0.6000	0.6000	$\cdots$	0.6000	0.6000	0.6333	0.7333	0.6333	0.6333	0.4667
	KNGA	0.6000	0.6000	0.6000	0.6000	$\cdots$	0.6000	0.6000	0.6333	0.6333	0.5667	0.5333	0.4667
	KNGD	0.6000	0.6000	0.6000	0.6000	$\cdots$	0.6000	0.6000	0.6333	0.6667	0.5667	0.5667	0.5000
(5) SolarFlare ($K = 5$)	KNGR	0.6197	0.6197	0.6197	0.6056	$\cdots$	0.6432	0.6338	0.6244	0.5962	0.5915	0.6338	0.6432
	KNGA	0.6197	0.6197	0.6197	0.6056	$\cdots$	0.6526	0.6291	0.6150	0.6009	0.5962	0.6338	0.6620
	KNGD	0.6197	0.6197	0.6197	0.6056	$\cdots$	0.6573	0.6338	0.6150	0.6009	0.5915	0.6338	0.6526
(6) Cmc ($K = 5$)	KNGR	0.1356	0.2034	0.2339	0.2373	$\cdots$	0.3254	0.3356	0.3153	0.3220	0.3017	0.3085	0.3051
	KNGA	0.1322	0.1898	0.2169	0.2102	$\cdots$	0.3356	0.3356	0.3119	0.2881	0.3017	0.3017	0.3017
	KNGD	0.1356	0.2102	0.2237	0.2305	$\cdots$	0.3356	0.3322	0.3119	0.2915	0.3017	0.3085	0.3085



(a) Fertility (b) Wdbc (c) Wine

(d) Lymphography (e) SolarFlare (f) Cmc

Fig. 6. Contrast diagram of classification results with changed δ and fixed K.

generally have no significant difference; KNGA is sometimes better than KNGR while KNGD sometimes becomes the best, and each algorithm actually has the region and possibility for the optimal performance.

Then consider experiments by changed parameter K. We use a natural series from 1 to 10 with step length 1 and step number 10, i.e.,

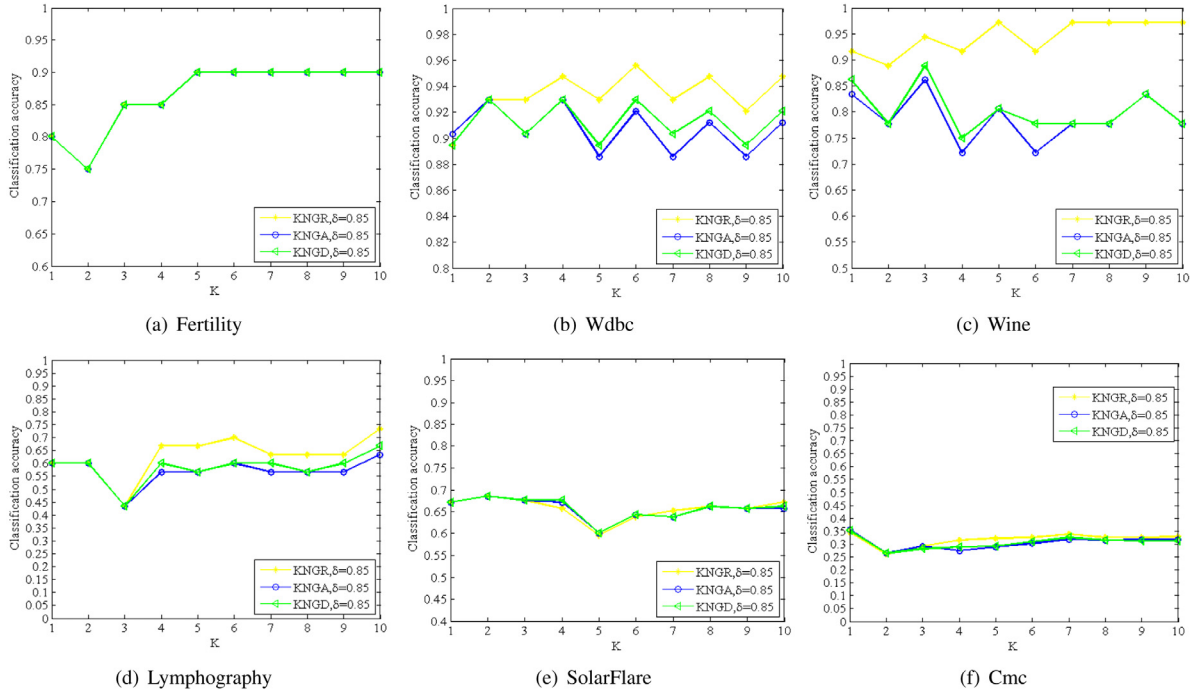
$$K : 1 \to 2 \to 3 \to \cdots \to 9 \to 10. \tag{38}$$

On the other hand, the neighborhood radius is fixed at $\delta = 0.85$ for each dataset.

Regarding the six datasets, relevant classification accuracy values based on classifiers KNGR, KNGA, KNGD are given in Table 8

Table 8
Classification accuracy with fixed δ and changed K.

Dataset (Parameter δ)	Algorithm	$K=1$	2	3	4	5	6	7	8	9	$K=10$
(1) Fertility	KNGR	0.8000	0.7500	0.8500	0.8500	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
($\delta=0.85$)	KNGA	0.8000	0.7500	0.8500	0.8500	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
	KNGD	0.8000	0.7500	0.8500	0.8500	0.9000	0.9000	0.9000	0.9000	0.9000	0.9000
(2) Wdbc	KNGR	0.8947	0.9298	0.9298	0.9474	0.9298	0.9561	0.9298	0.9474	0.9211	0.9474
($\delta=0.85$)	KNGA	0.9035	0.9298	0.9035	0.9298	0.8860	0.9211	0.8860	0.9123	0.8860	0.9123
	KNGD	0.8947	0.9298	0.9035	0.9298	0.8947	0.9298	0.9035	0.9211	0.8947	0.9211
(3) Wine	KNGR	0.9167	0.8889	0.9444	0.9167	0.9722	0.9167	0.9722	0.9722	0.9722	0.9722
($\delta=0.85$)	KNGA	0.8333	0.7778	0.8611	0.7222	0.8056	0.7222	0.7778	0.7778	0.8333	0.7778
	KNGD	0.8611	0.7778	0.8889	0.7500	0.8056	0.7778	0.7778	0.7778	0.8333	0.7778
(4) Lymphography	KNGR	0.6000	0.6000	0.4333	0.6667	0.6667	0.7000	0.6333	0.6333	0.6333	0.7333
($\delta=0.85$)	KNGA	0.6000	0.6000	0.4333	0.5667	0.5667	0.6000	0.5667	0.5667	0.5667	0.6333
	KNGD	0.6000	0.6000	0.4333	0.6000	0.5667	0.6000	0.6000	0.5667	0.6000	0.6667
(5) SolarFlare	KNGR	0.6714	0.6854	0.6761	0.6573	0.5962	0.6385	0.6526	0.6620	0.6573	0.6714
($\delta=0.85$)	KNGA	0.6714	0.6854	0.6761	0.6714	0.6009	0.6432	0.6385	0.6620	0.6573	0.6573
	KNGD	0.6714	0.6854	0.6761	0.6761	0.6009	0.6432	0.6385	0.6620	0.6573	0.6620
(6) Cmc	KNGR	0.3458	0.2610	0.2915	0.3153	0.3220	0.3254	0.3390	0.3254	0.3254	0.3288
($\delta=0.85$)	KNGA	0.3559	0.2644	0.2915	0.2746	0.2881	0.3017	0.3186	0.3153	0.3186	0.3186
	KNGD	0.3525	0.2644	0.2814	0.2881	0.2915	0.3085	0.3254	0.3153	0.3119	0.3119



(a) Fertility (b) Wdbc (c) Wine

(d) Lymphography (e) SolarFlare (f) Cmc

Fig. 7. Contrast diagram of classification results with fixed δ and changed K.

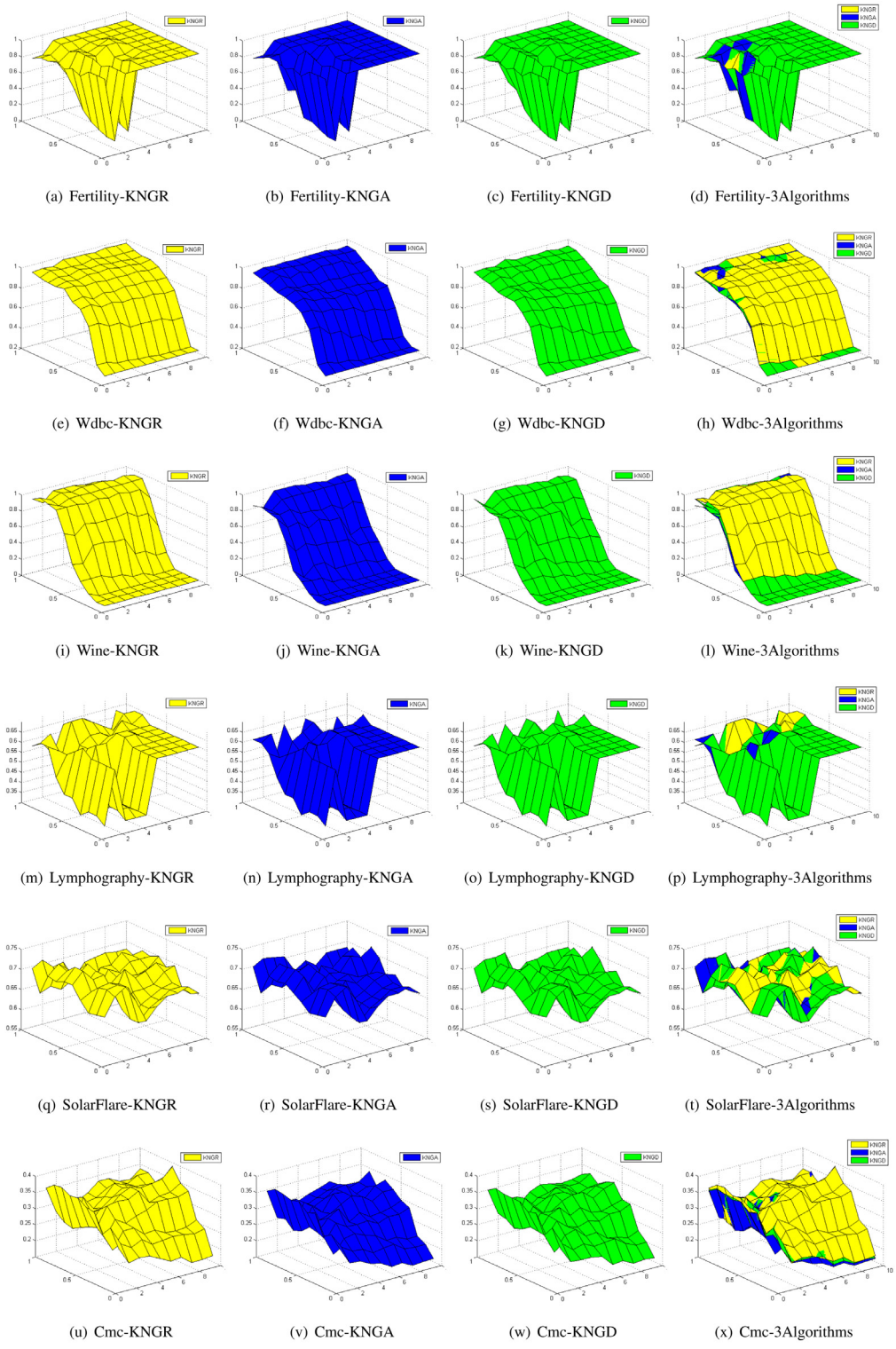
and are described in Fig. 7. Regarding (1) Fertility, KNGR, KNGA, KNGD acquire the same result. Regarding (2) Wdbc, (3) Wine, (4) Lymphography, KNGD becomes suboptimal to fall in between KNGR and KNGA, where KNGR is usually better than KNGA, and the relevant difference of accuracy is obvious. Regarding surplus (5) SolarFlare and (6) Cmc, the three algorithms generally have no significant difference, and KNGR, KNGA, KNGD all have the case and possibility to achieve the optimal result.

Finally, test the influence from both parameters δ and K. Herein, a new radius chain

$$\delta : 0.1 \rightarrow 0.2 \rightarrow 0.3 \rightarrow \cdots \rightarrow 0.9 \rightarrow 1 \tag{39}$$

is considered by coarsening the previous radius chain in Eq. (37), and it is further combined with the previous K sequence in Eq. (38). In other words, we will resort to Eqs. (38) and (39) to generate a two-dimensional parameter net with scale 10×10, and relevant classification accuracy values in the third dimension can more thoroughly reveal the achievement relationships of the three algorithms.

All accuracy values of three methods KNGR, KNGA, KNGD are vividly depicted in Fig. 8 in the three-dimensional space. For the six datasets, Fig. 8 can yield the following observation and analysis on the statistical area of parameters $\delta \in [0.1, 1]$ and $K \in [1, 10]$, especially from the viewpoint of algorithm KNGD. Regarding (1) Fertility, KNGD can acquire the maximal accuracy in most cases, such as when δ is small or K is great, while KNGR or KNGA can also reach the relevant optimal accuracy for the surplus cases. Regarding (2) Wdbc, only KNGR acquires the best effect in main regions, while KNGA or KNGD can generate the optimal result in a very few part. Regarding (3) Wine, only KNGR achieves the optimal result in majority regions, while KNGD (rather than KNGA) gains the classification performance in the surplus part. Regarding (4) Lymphography, KNGD can optimally cover the most



(a) Fertility-KNGR (b) Fertility-KNGA (c) Fertility-KNGD (d) Fertility-3Algorithms

(e) Wdbc-KNGR (f) Wdbc-KNGA (g) Wdbc-KNGD (h) Wdbc-3Algorithms

(i) Wine-KNGR (j) Wine-KNGA (k) Wine-KNGD (l) Wine-3Algorithms

(m) Lymphography-KNGR (n) Lymphography-KNGA (o) Lymphography-KNGD (p) Lymphography-3Algorithms

(q) SolarFlare-KNGR (r) SolarFlare-KNGA (s) SolarFlare-KNGD (t) SolarFlare-3Algorithms

(u) Cmc-KNGR (v) Cmc-KNGA (w) Cmc-KNGD (x) Cmc-3Algorithms

Fig. 8. Contrast diagram of classification results with changed δ and K.

cases, while KNGR or KNGA determines the surplus region. Regarding (5) SolarFlare, KNGD can play the optimal role in almost half regions, while KNGR, KNGA can come into play in the surplus part. Regarding (6) Cmc, KNGR, KNGA, KNGD respectively play a main, minor, and auxiliary roles. By these discussions, we can draw a general conclusion, which also summarizes the previous two cases with only a single-parameter change. That is, KNGD can reach the optimum accuracy in many cases to outperform KNGR or KNGA, and it sometimes achieves a compromised classification performance between KNGR and KNGA; thus, the double-quantitative classifier KNGD has the practical application space, where the single-quantification measurement is weakened.

In summary, algorithm KNGD generally acquires the improvement or comprise for existing two algorithms KNGR and KNGA, so it is effective for classification learning. Moreover, the relevant efficiency can be analyzed. For the same dataset (even exceeding

1000 objects), three algorithms KNGR, KNGA, and KNGD almost have the similar level of run time (where algorithm KNGD is slightly slower), and this is because they have the same algorithm framework but the different quantification distances. For the same classification algorithm (of KNGR, KNGA, or KNGD), the change of object scale necessarily causes the obvious time difference, while the change of attribute scale does not easily lead to the significant time change. In other words, the sample size of dataset becomes the main factor of time complexity for these classification algorithms, and this is because relevant classifiers concern multiple times of object loops to get the relevant covering granulation and distance utilization (where attributes are fixed at the whole set).

5. Conclusions

The tri-level granular structure of neighborhood system becomes significant by its bottom neighborhood granule, middle neighborhood swarm, and top neighborhood library [31], but it has the theoretical incomplete construction and applied potential promotion. As shown by Fig. 1, this paper mainly makes both the theory perfection and application extension.

(1) The size assessment and relevant granulation non-monotonicity are supplemented at the top library level, while the logical operation and single-quantitative distance are supplemented at both the middle swarm level and top library level.
(2) The tri-level double-quantitative distances utilize the double-quantification technology to combine and extend corresponding single-quantitative distances, so they offer the more powerful measurement and more robust applications.
(3) The double-quantitative distance of neighborhood granules is utilized to conduct the classification learning, and its corresponding classifier KNGD outperforms or balances the existing two classifiers KNGR and KNGA, as verified by data experiments. The good performance of KNGD mainly benefits from the optimization of reasonable double-quantification fusion and compromise.

By virtue of relevant theory discussion, example illustration, and experiment verification, this study hierarchically perfects the tri-level granular structure of neighborhood system, and the corresponding double-quantification integration and extension offer the robust knowledge measurement and effective classification learning. In view of the fusion characteristic and measurement superiority of double-quantification, the double-quantitative measurement and classifier are mainly applicable for a comprehensive and systematic scenario, where neither the relative quantification nor the absolute quantification plays a leading role.

According to relevant contents, there are three issues are remained for future explorations.

(1) Regarding the logical operations, they lack the closeness at the bottom level, and they are directly defined by the initial attribute subsets. Thus, they are worth further researching by essential neighborhood coverings and hierarchical integration connections, based on the tri-level granular structure of neighborhood system.
(2) Regarding the distance measurement at the top library level, a basic precondition of equal cardinality is used to establish a one-to-one correspondence between internal swarms. Thus, the general case needs in-depth considerations. In other words, how to define and develop distances between top neighborhood libraries K_S^δ and K_T^δ (where $|S| \neq |T|$) becomes an open question.
(3) Regarding tri-level applications, only the bottom distance of neighborhood granules is concerned, i.e., the single-quantitative and double-quantitative distances and their induced classifiers are effectively used for classification learning. Furthermore, potential applications should be mined at the middle and top levels. The middle swarm is related to the knowledge covering, and thus relevant attribute significance and uncertainty information can be concerned; the top library implies the knowledge base, and thus optimal processing and reasoning can be discussed. For example, the feature selection, which is related to the middle knowledge and the top knowledge base, can be utilized in advance before the classification learning at the bottom level, and thus, the corresponding attribute reduction is worth hierarchically exploring at the tri-level granular structure of neighborhood system; in particular, a recent study on hierarchical, systematic, and informational attribute reducts [68] can provide some reference.

CRediT authorship contribution statement

Xianyong Zhang: Conceptualization, Methodology, Formal analysis. **Hongyuan Gou:** Software, Visualization. **Zhiying Lv:** Validation, Investigation. **Duoqian Miao:** Conceptualization, Supervision.

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 authors thank both the editors and reviewers for their valuable suggestions, which substantially improve this paper.

This work was supported by National Natural Science Foundation of China (61673285, 61673301, 61976158), Sichuan Science and Technology Program of China (2021YJ0085, 2019YJ0529), and A Joint Research Project of Laurent Mathematics Center of Sichuan Normal University and National-Local Joint Engineering Laboratory of System Credibility Automatic Verification.

Appendix A. Proof of Lemma 1

Proof. Regarding non-empty subsets and their numbers, B has $m = 2^{|B|} - 1$ ones: $R_1, \ldots, R_m$, while $Q = B \cup \{a_*\}$ has $2^{|B|+1} - 1$ ones. Note that $2^{|B|+1}-1 = (2^{|B|}-1)+(2^{|B|}-1)+1$. By combination observation, non-empty subsets of Q can be divided into three groups related to $R_1, \ldots, R_m$.

(1) The first group concerns initial subsets of Q, i.e., $R_1, \ldots, R_m$, and it offers number $2^{|B|} - 1$.
(2) The second group involves the union of initial $R_1, \ldots, R_m$ and additional a_*, i.e., $R_1 \cup \{a_*\}, \ldots, R_m \cup \{a_*\}$, and it also offers number $2^{|B|} - 1$.
(3) The third group contains not $R_1, \ldots, R_m$ but the additional attribute a_*, i.e., $\{a_*\}$, and it submits to number 1.

By repeatability examination, the above three groups of non-empty subsets exactly constitute $2^Q - \{\emptyset\}$, i.e., Eq. (12) holds. □

Appendix B. Proof of Lemma 2

Proof. At first, a general range $KM(K_B^\delta) \in (\frac{1}{|U|}, 1)$ is determined by the basic range in Eq. (11) and the general assumption in Eq. (13).

Regarding initial B, library K_B^δ has m swarms based on $R_1, \ldots, R_m$. The other Q has three groups of non-empty subsets according to Lemma 1, so library K_Q^δ further adds both m swarms based on $R_1, \ldots, R_m$ and 1 swarm based on $\{a_*\}$. Thus, $KM(K_B^\delta)$ and $KM(K_Q^\delta)$ can be calculated and compared, and we mainly utilize the neighborhood realization about single attributes, i.e.,

$$n_{B'}^\delta(x) = \bigcap_{a \in B' \subseteq A} n_{\{a\}}^\delta(x). \tag{B.1}$$

(1) When the added attribute a_* corresponds to $U/NR_\delta(\{a_*\}) = \{\{x_1\}, \{x_2\}, \ldots, \{x_{|U|}\}\}$, each increased swarm becomes $\left(\{x_1\}, \{x_2\}, \ldots, \{x_{|U|}\}\right)$ (according to Eq. (B.1)) to offer size value $\frac{1}{|U|}$. Since $\frac{1}{|U|} < KM(K_B^\delta)$, we can achieve

$$\begin{aligned} KM(K_Q^\delta) &= \frac{1}{2m+1}\left(KM(K_B^\delta) \times m + \frac{1}{|U|} \times m + \frac{1}{|U|} \times 1\right) \\ &< \frac{1}{2m+1}\left(KM(K_B^\delta) \times m + KM(K_B^\delta) \times m + KM(K_B^\delta) \times 1\right) \\ &= KM(K_B^\delta), \end{aligned} \tag{B.2}$$

i.e., $KM(K_B^\delta) > KM(K_Q^\delta)$.

(2) When a_* is accompanied with $U/NR_\delta(\{a_*\}) = \{U\}$, Eq. (B.1) leads to two cases for increased swarms. The increased swarm on the latter subset $R_j \cup \{a_*\}$ is equivalent to the former swarm on subset R_j to provide the repetitive size value, where $j = 1, \ldots, m$, while the surplus increased swarm on $\{a_*\}$ becomes $\left(U, U, \ldots, U\right)$ to offer size value 1. Since $1 > KM(K_B^\delta)$, we can achieve

$$\begin{aligned} KM(K_Q^\delta) &= \frac{1}{2m+1}\left(KM(K_B^\delta) \times m + KM(K_B^\delta) \times m + 1 \times 1\right) \\ &> \frac{1}{2m+1}\left(KM(K_B^\delta) \times m + KM(K_B^\delta) \times m + KM(K_B^\delta) \times 1\right) \\ &= KM(K_B^\delta), \end{aligned} \tag{B.3}$$

i.e., $KM(K_B^\delta) < KM(K_Q^\delta)$. □

References

[1] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning About Data, vol. 9, Kluwer Academic Publishers, Dordrecht, 1991.
[2] Y.Y. Yao, Relational interpretations of neighborhood operators and rough set approximation operators, Inform. Sci. 111 (1–4) (1998) 239–259.
[3] Q.H. Hu, D.R. Yu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Inform. Sci. 178 (2008) 3577–3594.
[4] F.Y. Cao, J.Y. Liang, G. Jiang, An initialization method for the K-means algorithm using neighborhood model, Comput. Math. Appl. 58 (2009) 474–483.
[5] P. Yao, Y.H. Lu, Neighborhood rough set and SVM based hybrid credit scoring classifier, Expert Syst. Appl. 38 (9) (2011) 11300–11304.
[6] P. Zhou, X.G. Hu, P.P. Li, X.D. Wu, Online feature selection for high-dimensional class-imbalanced data, Knowl.-Based Syst. 9 (2017) 1–13.
[7] Y.M. Chen, K.S. Wu, X.H. Chen, C.H. Tang, Q.X. Zhu, An entropy-based uncertainty measurement approach in neighborhood systems, Inform. Sci. 279 (2014) 239–250.
[8] S.U. Kumar, H.H. Inbarani, PSO-Based feature selection and neighborhood rough set-based classification for BCI multiclass motor imagery task, Neural Comput. Appl. 28 (11) (2017) 3239–3258.
[9] X.D. Yue, Y.F. Chen, D.Q. Miao, H. Fujita, Fuzzy neighborhood covering for three-way classification, Inform. Sci. 507 (2020) 795–808.
[10] D.B. Chakraborty, S.K. Pal, Neighborhood rough filter and intuitionistic entropy in unsupervised tracking, IEEE Trans. Fuzzy Syst. 26 (4) (2018) 2188–2200.
[11] Q.H. Hu, W. Pedrycz, D.R. Yu, J. Lang, Selecting discrete and continuous features based on neighborhood decision error minimization, IEEE Trans. Syst. Man Cybern. B 40 (1) (2010) 137–150.
[12] J.H. Liu, Y.J. Lin, Y.W. Li, W. Weng, S.X. Wu, Online multi-label streaming feature selection based on neighborhood rough set, Pattern Recognit. 84 (2018) 273–287.
[13] C.Z. Wang, Y. Huang, M.W. Shao, Q.H. Hu, D.G. Chen, Feature selection based on neighborhood self-information, IEEE Trans. Cybern. 50 (9) (2020) 4031–4042.
[14] H.M. Chen, T.R. Li, Y. Cai, C. Luo, H. Fujita, Parallel attribute reduction in dominance-based neighborhood rough set, Inform. Sci. 373 (2016) 351–368.
[15] X.D. Fan, W.D. Zhao, C.Z. Wang, Y. Huang, Attribute reduction based on max-decision neighborhood rough set model, Knowl.-Based Syst. 151 (2018) 16–23.
[16] Q. Wang, Y.H. Qian, X.Y. Liang, Q. Guo, J.Y. Liang, Local neighborhood rough set, Knowl.-Based Syst. 153 (2018) 53–64.
[17] Y. Liu, W.L. Huang, Y.L. Jiang, Z.Y. Zeng, Quick attribute reduct algorithm for neighborhood rough set model, Inform. Sci. 271 (7) (2014) 65–81.
[18] Y.M. Chen, Z.Q. Zeng, J.W. Lu, Neighborhood rough set reduction with fish swarm algorithm, Soft Comput. 21 (23) (2017) 6907–6918.
[19] J. Meng, J. Zhang, Y.S. Luan, Gene selection integrated with biological knowledge for plant stress response using neighborhood system and rough set theory, IEEE/ACM Trans. Comput. Biol. Bioinform. 12 (2) (2015) 433–444.
[20] Y.M. Chen, Z.J. Zhang, J.Z. Zheng, Y. Ma, Y. Xue, Gene selection for tumor classification using neighborhood rough sets and entropy measures, J. Biomed. Inform. 67 (2017) 59–68.
[21] Z. Yuan, X.Y. Zhang, S. Feng, Hybrid data-driven outlier detection based on neighborhood information entropy and its developmental measures, Expert Syst. Appl. 112 (2018) 243–257.
[22] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (2) (1997) 111–127.
[23] T.Y. Lin, Granular computing on binary relations II: Rough set representations and belief functions, in: A. Skowron, L. Polkowski (Eds.), Rough Sets in Knowledge Discovery, Physica-Verlag, Heidelberg, 1998, pp. 121–140.
[24] Y.Y. Yao, Information granulation and rough set approximation, Int. J. Intell. Syst. 16 (1) (2001) 87–104.
[25] T.Y. Lin, Neighborhood systems and approximation inrelational databases and knowledge bases, in: Proceedings of the Fourth International Symposium on Methodologies of Intelligent Systems, 1989, pp. 75–86.
[26] X.B. Yang, M. Zhang, H.L. Dou, Y.J. Yang, Neighborhood systems-based rough sets in incomplete information system, Knowl.-Based Syst. 24 (6) (2011) 858–867.
[27] L.J. Wang, X.B. Yang, J.Y. Yang, W. Chen, Relationships among generalized rough sets in six coverings and pure reflexive neighborhood system, Inform. Sci. 207 (2012) 66–78.
[28] T.T. Zheng, L.Y. Zhu, Uncertainty measures of neighborhood system-based rough sets, Knowl.-Based Syst. 86 (2015) 57–65.
[29] Y.L. Zhang, C.Q. Li, M.L. Lin, Y.J. Lin, Relationships between generalized rough sets based on covering and reflexive neighborhood system, Inform. Sci. 319 (2015) 56–67.
[30] Y.H. Zhou, X.Y. Zhang, Z.M. Mo, Granular monotonic conditional neighborhood entropy and its related attribute reduction, Comput. Res. Develop. 55 (11) (2018) 2395–2405, (in Chinese).
[31] Y.M. Chen, N. Qin, W. Li, F.F. Xu, Granule structures distances and measures in neighborhood systems, Knowl.-Based Syst. 165 (2019) 268–281.
[32] Y.Y. Yao, Three-way granular computing, rough sets, and formal concept analysis, Internat. J. Approx. Reason. 116 (2020) 106–125.
[33] Y.Y. Yao, Tri-level thinking: Models of three-way decision, Int. J. Mach. Learn. Cybern. 11 (2020) 947–959.
[34] X.Y. Zhang, Z.W. Mo, F. Xiong, W. Cheng, Comparative study of variable precision rough set model and graded rough set model, Internat. J. Approx. Reason. 53 (1) (2012) 104–116.
[35] B. Huang, H.X. Li, G.F. Feng, C.X. Guo, D.F. Chen, Double-quantitative rough sets, optimal scale selection and reduction in multi-scale dominance IF decision tables, Internat. J. Approx. Reason. 130 (2021) 170–191.
[36] X.Y. Hu, B.Z. Sun, T. Wang, Double-quantitative decision rough set over two universes and application to african swine fever decision-making, J. Exp. Theoret. Artif. Intell. (2021) http://dx.doi.org/10.1080/0952813X.2020.1744195.
[37] Y.T. Guo, E.C.C. Tsang, W.H. Xu, D.G. Chen, Local logical disjunction double-quantitative rough sets, Inform. Sci. 500 (2019) 87–112.
[38] B.B. Sang, L. Yang, H.M. Chen, W.H. Xu, Y.T. Guo, Z. Yuan, Generalized multi-granulation double-quantitative decision-theoretic rough set of multi-source information system, Internat. J. Approx. Reason. 115 (2019) 157–179.

[39] M.M. Li, M.H. Chen, W.H. Xu, Double-quantitative multigranulation decision-theoretic rough fuzzy set model, Int. J. Mach. Learn. Cybern. 10 (11) (2019) 3225–3244.
[40] W.T. Li, X.P. Xue, W.H. Xu, T. Zhan, B.J. Fan, Double-quantitative variable consistency dominance-based rough set approach, Internat. J. Approx. Reason. 124 (2020) 1–26.
[41] W.H. Xu, Y.T. Guo, Generalized multigranulation double-quantitative decision-theoretic rough set, Knowl.-Based Syst. 105 (2016) 190–205.
[42] B.W. Fang, B.Q. Hu, Probabilistic graded rough set and double relative quantitative decision-theoretic rough set, Internat. J. Approx. Reason. 74 (2016) 1–12.
[43] X.Y. Zhang, D.Q. Miao, Two basic double-quantitative rough set models of precision and grade and their investigation using granular computing, Internat. J. Approx. Reason. 54 (8) (2013) 1130–1148.
[44] Y.T. Guo, E.C.C. Tsang, M. Hu, X.X. Lin, D.G. Chen, W.H. Xu, B.B. Sang, Incremental updating approximations for double-quantitative decision-theoretic rough sets with the variation of objects, Knowl.-Based Syst. 189 (2020) 105082.
[45] J.H. Yu, B. Zhang, M.H. Chen, W.H. Xu, Double-quantitative decision-theoretic approach to multigranulation approximate space, Internat. J. Approx. Reason. 98 (2018) 236–258.
[46] W.T. Li, W. Pedrycz, X.P. Xue, W.H. Xu, B.J. Fan, Fuzziness and incremental information of disjoint regions in double-quantitative decision-theoretic rough set model, Int. J. Mach. Learn. Cybern. 10 (10) (2019) 2669–2690.
[47] X.Y. Zhang, D.Q. Miao, Quantitative information architecture, granular computing and rough set models in the double-quantitative approximation space of precision and grade, Inform. Sci. 268 (2014) 147–168.
[48] X.Y. Zhang, D.Q. Miao, Double-quantitative fusion of accuracy and importance: Systematic measure mining, benign integration construction, hierarchical attribute reduction, Knowl.-Based Syst. 91 (2016) 219–240.
[49] X.Y. Hu, B.Z. Sun, X.T. Chen, Double quantitative fuzzy rough set-based improved AHP method and application to supplier selection decision making, Int. J. Mach. Learn. Cybern. 11 (1) (2020) 153–167.
[50] B.J. Fan, E.C.C. Tsang, W.H. Xu, J.H. Yu, Double-quantitative rough fuzzy set based decisions: A logical operations method, Inform. Sci. 378 (2017) 264–281.
[51] X.Y. Zhang, D.Q. Miao, Three-layer granular structures and three-way informational measures of a decision table, Inform. Sci. 412–413 (2017) 67–86.
[52] Y.Y. Yao, Three-way decisions and cognitive computing, Cogn. Comput. 8 (4) (2016) 543–554.
[53] P.K. Singh, Three-way fuzzy concept lattice representation using neutrosophic set, Int. J. Mach. Learn. Cybern. 8 (2017) 69–79.
[54] X.F. Deng, Y.Y. Yao, Decision-theoretic three-way approximations of fuzzy sets, Inform. Sci. 279 (2014) 702–715.
[55] P.K. Singh, Medical diagnoses using three-way fuzzy concept lattice and their Euclidean distance, Comput. Appl. Math. 37 (2018) 3283–3306.
[56] Y.Y. Yao, Interval sets and three-way concept analysis in incomplete contexts, Int. J. Mach. Learn. Cybern. 8 (2017) 3–20.
[57] P.K. Singh, Interval-valued neutrosophic graph representation of concept lattice and its (α, β, γ)-decomposition, Arab. J. Sci. Eng. 43 (2018) 723–740.
[58] P.K. Singh, C.A. Kumar, Concept lattice reduction using different subset of attributes as information granules, Gran. Comput. 2 (2017) 159–173.
[59] X.Y. Zhang, D.Q. Miao, Three-way attribute reducts, Internat. J. Approx. Reason. 88 (2017) 401–434.
[60] D. Liu, X. Yang, T.R. Li, Three-way decisions: beyond rough sets and granular computing, Int. J. Mach. Learn. Cybern. 11 (2020) 989–1002.
[61] Q.H. Hu, D.R. Yu, Z.X. Xie, Neighborhood classifiers, Expert Syst. Appl. 34 (2) (2008) 866–876.
[62] J. Blaszczynski, S. Greco, R. Slowinski, M. Szelg, Monotonic variable consistency rough set approaches, Internat. J. Approx. Reason. 50 (7) (2009) 979–999.
[63] C. Gao, Z.H. Lai, J. Zhou, J.J. Wen, W.K. Wong, Granular maximum decision entropy-based monotonic uncertainty measure for attribute reduction, Internat. J. Approx. Reason. 104 (2019) 9–24.
[64] D.Q. Miao, Y. Zhao, Y.Y. Yao, H.X. Li, F.F. Xu, Relative reducts in consistent and inconsistent decision tables of the pawlak rough set model, Inform. Sci. 179 (24) (2009) 4140–4150.
[65] G.Y. Wang, X.A. Ma, H. Yu, Monotonic uncertainty measures for attribute reduction in probabilistic rough set model, Internat. J. Approx. Reason. 59 (2015) 41–67.
[66] P.E. Hart, The condensed nearest neighbor rule, IEEE Trans. Inform. Theory 14 (3) (1968) 515–516.
[67] D. Dua, C. Graff, UCI Machine Learning Repository, University of California, School of Information and Computer Science, Irvine, CA, 2019, http://archive.ics.uci.edu/ml.
[68] X.Y. Zhang, H. Yao, Z.Y. Lv, D.Q. Miao, Class-specific information measures and attribute reducts for hierarchy and systematicness, Inform. Sci. (2021) http://dx.doi.org/10.1016/j.ins.2021.01.080.