International Journal of Machine Learning and Cybernetics (2020) 11:1141–1157
https://doi.org/10.1007/s13042-020-01089-4

ORIGINAL ARTICLE

Knowledge granularity based incremental attribute reduction for incomplete decision systems

Chucai Zhang[1] · Jianhua Dai[1] · Jiaolong Chen[1]

Received: 23 June 2019 / Accepted: 12 February 2020 / Published online: 27 February 2020

Abstract
Attribute reduction is an important application of rough set theory. With the dynamic changes of data becoming more and more common, traditional attribute reduction, also called static attribute reduction, is no longer efficient. How to update attribute reducts efficiently gets more and more attention. In the light of the variation about the number of objects, we focus on incremental attribute reduction approaches based on knowledge granularity which can be used to measure the uncertainty in incomplete decision systems. We first introduce incremental mechanisms to calculate knowledge granularity for incomplete decision systems when multiple objects vary dynamically. Then, incremental attribute reduction algorithms for incomplete decision systems when adding multiple objects and when deleting multiple objects are proposed respectively. Finally, comparative experiments on different real-life data sets are conducted to demonstrate the effectiveness and efficiency of the proposed incremental algorithms for updating attribute reducts with the variation of multiple objects in incomplete decision systems.

Keywords Incremental attribute reduction · Knowledge granularity · Incomplete decision system · Rough sets

1 Introduction

Rough set theory [1], as one of the effective tools to handle uncertainty requiring no preliminary or additional information about data, has been applied successfully in various applications, such as machine learning [2–4], rule mining [5], forecast [6], decision supporting [7–9], knowledge engineering [10, 11], disease diagnosis [12–14], genetic engineering [15, 16], uncertainty modelling [17–21], and other applications [22–25].

Attribute reduction, also called attribute selection or feature selection, as one of the most important core applications of rough sets, can select a few attributes with high importance from all attributes of information systems for many practical applications. Selected attributes are retained and unselected attributes are removed, the process of which is treated as a pre-processing stage of data analysis without decreasing the accuracy of classification. Many researchers have deeply studied non-incremental attribute reduction, also called classic attribute reduction or static attribute reduction, and have obtained many achievements. In general, non-incremental attribute reduction approaches can be divided into four categories: methods based on positive domain [26, 27], methods based on discernibility matrix [28–32], methods based on information entropy [15, 33, 34] and other attribute reduction methods [35–41]. For example, Fan et al. [26] presented systematic acceleration policies that can reduce the computational domain and optimize the computation of the positive region. Ni et al. [27] presented a novel accelerator based on the positive region for attribute reduction. Konecny [30] studied the problem of attribute reduction by using discernibility matrix in various extensions of formal concept analysis. Wang et al. [31] proposed a new feature evaluation function for feature selection by using discernibility matrix. Dai and Tian [33] defined the concepts of knowledge information entropy, knowledge rough entropy, knowledge granulation and knowledge granularity measure in set-valued information systems and studied relationship between these concepts. Wang et al. [34] developed an information entropy based attribute reduction algorithm for data sets with dynamically varying data values. Dai et al. [36] put forward two new attribute reduction algorithms based on quick neighbor inconsistent

✉ Jianhua Dai
jhdai@hunnu.edu.cn

1 Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, Hunan Normal University, Changsha 410081, China

pairs, which can reduce the time for finding a reduct. Wang et al. [40] introduced distance measures into fuzzy rough sets by constructing a fuzzy rough set model based on distance measure with a fixed parameter and proposed a greedy convergent algorithm for attribute reduction. Dai et al. [37] proposed two attribute selection algorithms from the viewpoint of the object pair which can ignore the object pairs that were already discerned by the selected attribute subsets and needed only to deal with part of object pairs instead of the whole object pairs from the discourse.

Unlike non-incremental reduction which needs to be calculated from scratch, incremental reduction approaches can use the previous reduction results which have been obtained from the original decision system, to get new reduct of the changed decision system. This advantage of incremental reduction approaches greatly satisfies current situation that data sets usually vary with time, which has aroused the interest of researchers. There are several useful findings that can handle three conditions: variation of object sets [42–47], variation of attribute sets [48] and variation of attribute values [49, 50]. For example, Yang et al. [46] presented an efficient incremental algorithm which includes active sample selection process and incremental attribute reduction process for dynamic data sets with increasing samples. Ma et al. [44] introduced a compressed binary discernibility matrix and developed an incremental attribute reduction algorithm for group dynamic data considering both situations of single dynamic object and group dynamic objects. Jing et al. [48]considered the variation of attribute sets in complete decision systems and presented the corresponding incremental algorithms for attribute reduction. Wei et al. [50] proposed a discernibility matrix based incremental attribute reduction algorithm which can obtain all the reducts of dynamic data. There are a few algorithms that can deal with more complex situations, for example, Jing et al. [51] developed incremental methods to update reducts when attributes and objects of the decision system increased simultaneously. Yang et al. [52] presented a unified incremental reduction algorithm to deal with three kinds of variations of decision systems, such as adding an object, deleting an object and modifying the object's value.

Although researches on incremental attribute reduction can obtain reducts effectively and efficiently, they are based on the equivalence relation that could not be applied directly to incomplete decision systems with missing attribute values. To our best knowledge, a few achievements of the research on attribute reduction for incomplete decision systems have been obtained but not sufficient. Shu et al. [53, 54] focused on positive region-based incremental mechanisms for attribute reduction algorithms to get new attribute reducts when attribute values varied, attribute set varied and object set varied. Xie and Qin [55] introduced inconsistency degree to update reduct incrementally for dynamic incomplete decision and proved that the attribute reduction based on the inconsistency degree was equivalent to that based on the positive region. Luo et al. [56] proposed a model of dynamic probabilistic rough sets with incomplete data and presented incremental updating strategies and algorithms when adding and removing objects, respectively. Attribute reduction methods listed above obtain reducts for incomplete decision systems from the perspective of positive domain, information entropy or discrimination matrix. Zhang and Dai [57] proposed knowledge granularity based incremental attribute reduction methods from the perspective of granular computing for incomplete decision systems with objects varying and obtained satisfactory effects. However, they only considered the situation that objects varied one by one. Therefore, we investigate incremental mechanisms of knowledge granularity to get the reducts efficiently when multiple objects are added to or removed from incomplete decision systems simultaneously in this paper.

The remainder of the paper is organized as follows. Section 2 reviews basic concepts in rough set theory and knowledge granulation related concepts for incomplete decision systems. Section 3 investigates incremental mechanisms to calculate knowledge granularity for the cases of adding and deleting multiple objects respectively at a time and corresponding reduction algorithms for incomplete decision systems. Experiments and comparisons on practical data sets have been conducted in Sect. 4. Section 5 presents conclusions and points out our further research direction.

2 Preliminary knowledge

In this section, we review some basic concepts in rough set theory [1, 58], definitions of knowledge granulation, relative knowledge granularity, inner significance and outer significance respectively for incomplete decision systems, which can be found in [33, 48, 59].

2.1 Basic concepts in rough set theory

An information system is a quadruple $IIS = \langle U, A, V, f \rangle$, where U denotes a non-empty finite set of objects, which is called the universe; A denotes a non-empty finite set of condition attributes; V is the union of attribute domains, $V = \bigcup_{a\in A} V_a$, where V_a is the value set of attribute a, called the domain of a; $f : U \times A \rightarrow V$ is an information function which assigns particular values from domains of attribute to objects such as $a \in A, x \in U, f(a,x) \in V_a$, where $f(a, x)$ denotes the value of attribute a for object x. Each attribute subset $B \subseteq A$ determines a binary indiscernible relation as follows:

$$IND(B) = \left\{ \left(u_i, u_j\right) \in U^2 | \forall a \in B, a\left(u_i\right) = a\left(u_j\right) \right\} \tag{1}$$

By the relation $IND(B)$, we obtain the partition of U denoted by $U/IND(B)$ or U/B. For $B \subseteq A$ and $X \subseteq U$, the lower approximation and the upper approximation of X can be defined as follows:

$$\underline{B}X = \{u_i \in U | [u_i]_B \subseteq X\}$$
$$\overline{B}X = \{u_i \in U | [u_i]_B \cap X \neq \emptyset\},$$

where $\underline{B}X$ is a set of objects that belong to X with certainty, while $\overline{B}X$ is a set of objects that possibly belong to X.

A decision system is a quadruple $DS = \langle U, C \cup D, V, f \rangle$, where D is the decision attribute set, C is the condition attribute set, and $C \cap D = \emptyset$; V is the union of attribute domain, i.e., $V = V_C \cup V_D$. In general, we assume that $D = \{d\}$. If there exists an $a \in A$, $x \in U$ such that $f(a,x)$ is equal to a missing value, then the decision system is called an incomplete decision system (IDS). Thus, an incomplete decision system can be denoted as: $IDS = \langle U, C \cup D, V, f \rangle$, where $* \in V_C$, $* \notin V_D$.

Definition 1 [58] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $\forall B \subseteq C$, the binary tolerance relations between objects that are possibly indiscernible in terms of values of attributes in B is defined as

$$T(B) = \{(x,y) | \forall a \in B, f(a,x) = f(a,y) \vee f(a,x) = * \vee f(a,y) = *\} \quad (2)$$

$T(B)$ is reflexive and symmetric, but not necessarily transitive.

Definition 2 [58] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $x \in U$ and $B \subseteq C$, the tolerance class of the object x with respect to attribute set B is defined by:

$$T_B(x) = \{y | (x,y) \in T(B)\} \quad (3)$$

2.2 Knowledge granulation in incomplete decision systems

Definition 3 [33] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $B \subseteq C$. $T_B(u_i)$ is the tolerance class of object u_i with respect to B. The knowledge granularity of B on U is defined as follows:

$$GK_U(B) = \frac{1}{|U|^2} \sum_{i=1}^{|U|} |T_B(u_i)| \quad (4)$$

where $|U|$ stands for the number of objects in U.

Example 1 Example for computing of the knowledge granularity.

Table 1 shows an incomplete decision system $IDS = \langle U, C \cup D, V, f \rangle$, where $U = \{u_1, u_2, u_3, u_4, u_5\}$, $C = \{a_1, a_2, a_3, a_4, a_5\}$ and $D = \{d\}$. According to Definition 2, we have $T_C(u_1) = \{u_1\}$, $T_C(u_2) = \{u_2, u_4\}$, $T_C(u_3) = \{u_3, u_4\}$, $T_C(u_4) = \{u_2, u_3, u_4\}$, $T_C(u_5) = \{u_5\}$. According to Definition 3, we have $GK_U(C) = \frac{1}{5^2}(1+2+2+3+1) = \frac{9}{25}$. Similarly, we can get $GK_U(C \cup D) = \frac{5}{25}$.

Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete system, $P, Q \subseteq C$. If $P \subseteq Q$, then $GK_U(P) \geq GK_U(Q)$ [57]. Thus, the measure of knowledge granularity has the monotonicity with respect to attributes and is reasonable for the uncertainty measure in rough set theory.

Definition 4 [48] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $B \subseteq C$. $T(B)$ and $T(B \cup D)$ are the tolerance relations for attribute set B and $B \cup D$, respectively. The knowledge granularity of B relative to D on U is defined as follows:

$$GK_U(D|B) = GK_U(B) - GK_U(B \cup D) \quad (5)$$

Definition 5 [48] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $B \subseteq C$ and $a \in B$. T_B, $T_{B-\{a\}}$, $T_{B \cup D}$ and $T_{(B-\{a\}) \cup D}$ are the tolerance relations for B, $B - \{a\}$, $B \cup D$ and $(B - \{a\}) \cup D$, respectively. The significance measure (inner significance) of a in B on U is defined as follows:

$$Sig_U^{inner}(a, B, D) = GK_U(D|(B - \{a\})) - GK_U(D|B) \quad (6)$$

Definition 6 [59] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, the core of IDS is defined as follows:

$$Core_C = \{a \in C | Sig_U^{inner}(a, C, D) > 0\} \quad (7)$$

If $\forall a \in C$, $Sig_U^{inner}(a, C, D) = 0$, then $Core_C = \emptyset$.

Where a denotes any one attribute in C.

Example 2 (Continued from Example 1) According to Definition 4, we have $GK_U(D|C) = \frac{9}{25} - \frac{5}{25} = \frac{4}{25}$. Similarly, we can get $GK_U(D|C - \{a_1\}) = \frac{4}{25}$, $GK_U(D|C - \{a_2\}) = \frac{6}{25}$, $GK_U(D|C - \{a_3\}) = \frac{4}{25}$, $GK_U(D|C - \{a_4\}) = \frac{4}{25}$ and $GK_U(D|C - \{a_5\}) = \frac{4}{25}$. Then according to Definition 5, we can get $Sig_U^{inner}(a_1, C, D) = \frac{4}{25} - \frac{4}{25} = 0$, $Sig_U^{inner}(a_2, C, D) =$

Table 1 Incomplete decision system 1

U	a_1	a_2	a_3	a_4	a_5	d
u_1	1	*	0	0	1	0
u_2	0	0	1	*	1	0
u_3	0	0	*	1	0	0
u_4	*	0	1	1	*	1
u_5	0	1	1	1	0	1

$\frac{6}{25} - \frac{4}{25} = \frac{2}{25}$, $Sig_U^{inner}(a_3, C, D) = \frac{4}{25} - \frac{4}{25} = 0$, $Sig_U^{inner}(a_4, C, D) = \frac{4}{25} - \frac{4}{25} = 0$, $Sig_U^{inner}(a_5, C, D) = \frac{4}{25} - \frac{4}{25} = 0$. Then, we have $Core_C = \{a_2\}$.

Definition 7 [48] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $B \subseteq C$. $T_B, T_{B\cup D}, T_{B\cup\{a\}}$ and $T_{(B\cup\{a\})\cup D}$ are the tolerance relations for B, $B \cup D$, $B \cup \{a\}$ and $(B \cup \{a\}) \cup D$, respectively. Then $\forall a \in (C - B)$, the significance measure (outer significance) of a in B on U is defined as follows:

$$Sig_U^{outer}(a, B, D) = GK_U(D|B) - GK_U(D|(B \cup \{a\})) \tag{8}$$

Definition 8 [48] Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system, $B \subseteq C$. B is a relative reduct based on the knowledge granularity of *IDS* if

(1) $GK_U(D|B) = GK_U(D|C)$
(2) $\forall a \in B, GK_U(D|(B - \{a\})) \neq GK_U(D|B)$.

3 Incremental attribute reduction for incomplete decision systems when multiple objects arrive simultaneously

Incremental attribute reduction algorithms for incomplete decision systems which objects are added to and deleted from one by one were discussed in [57]. After investigating the incremental mechanisms to compute knowledge granularity for incomplete decision systems when multiple objects are added at a time and multiple objects are deleted at a time, this section introduces two incremental attribute reduction algorithms based on knowledge granularity for the addition of multiple objects and the deletion of multiple objects respectively.

3.1 An incremental mechanism to calculate knowledge granularity for IDS when adding multiple objects

This section investigates the changes of knowledge granularity, relative knowledge granularity, inner significance and outer significance, and then introduces the incremental mechanisms for calculating knowledge granularity when multiple objects are added to an incomplete decision system at a time.

Proposition 1 *Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system. $U = \{u_1, u_2, \dots, u_n\}$ denotes a nonempty finite set containing n objects. $U' = \{u'_1, u'_2, \dots, u'_s\}$ denotes the incremental object set containing s objects that will be added to IDS, and $T_C^{U^+}$ is the tolerance relation on $U^+ = U \cup U'$. The knowledge granularity of C on U^+ is*

$$GK_{U^+}(C) = \frac{1}{(n+s)^2}\left(n^2 GK_U(C) + 2\sum_{j=1}^{s} |T_C^U(u'_j)| + \sum_{j=1}^{s} |T_C^{U'}(u'_j)|\right) \tag{9}$$

where $T_C^U(u'_j) = \left\{u_i \middle| \left(u'_j, u_i\right) \in T_C^{U^+}, 1 \le i \le n, 1 \le j \le s\right\}$ and $T_C^{U'}(u'_j) = \left\{u'_i \middle| \left(u'_j, u'_i\right) \in T_C^{U^+}, 1 \le i \le s, 1 \le j \le s\right\}$.

Proof When $u'_j (1 \le j \le s)$ is adding to U, the tolerance class of $u_i (1 \le i \le n)$ on U^+ is

$$T_C^{U^+}(u_i) = \begin{cases} T_C^U(u_i) \cup \{u'_j\}, & (u_i, u'_j) \in T_C^{U^+} \\ T_C^U(u_i), & (u_i, u'_j) \notin T_C^{U^+} \end{cases}$$

We can get

$$T_C^{U^+}(u_i) = T_C^U(u_i) \cup T_C^{U'}(u_i), 1 \le i \le n$$

and

$$T_C^{U^+}(u'_j) = T_C^U(u'_j) \cup T_C^{U'}(u'_j), 1 \le j \le s.$$

Then we have

$$|T_C^{U^+}(u_i)| = |T_C^U(u_i)| + |T_C^{U'}(u_i)|, 1 \le i \le n$$

and

$$|T_C^{U^+}(u'_j)| = |T_C^U(u'_j)| + |T_C^{U'}(u'_j)|, 1 \le j \le s.$$

If $u'_j \in T_C^{U'}(u_i)$ then $u_i \in T_C^U(u'_j)$, $1 \le i \le n$, $1 \le j \le s$ and vice versa, because tolerance relation is symmetric. Obviously, we can get $\sum_{i=1}^{|U|} |T_C^{U'}(u_i)| = \sum_{j=1}^{|U'|} |T_C^U(u'_j)|$. According to Definition 3, the knowledge granularity of C on U^+ is described as following:

$$\begin{aligned} GK_{U^+}(C) &= \frac{1}{|U^+|^2} \sum_{i=1}^{|U^+|} |T_C^{U^+}(u_i)| \\ &= \frac{1}{|U^+|^2}\left(\sum_{i=1}^{|U|} |T_C^{U^+}(u_i)| + \sum_{j=1}^{|U'|} |T_C^{U^+}(u'_j)|\right) \\ &= \frac{1}{|U^+|^2}\left(\sum_{i=1}^{|U|} |T_C^U(u_i)| + \sum_{i=1}^{|U|} |T_C^{U'}(u_i)| + \sum_{j=1}^{|U'|} |T_C^U(u'_j)| \right. \\ &\quad \left. + \sum_{j=1}^{|U'|} |T_C^{U'}(u'_j)|\right) \\ &= \frac{1}{(n+s)^2}\left(n^2 GK_U(C) + 2\sum_{j=1}^{s} |T_C^U(u'_j)| + \sum_{j=1}^{s} |T_C^{U'}(u'_j)|\right) \end{aligned}$$

□

If $U' = \{u'_1\}$, that denotes adding objects one by one to the original object set U, then Eq. (9) can be written as $\frac{1}{(n+1)^2}(n^2 GK_U(C) + 2|T_C^U(u'_1)| + |T_C^{U'}(u'_1)|)$. From $T_C^U(u'_1) = T_C^{U^+}(u'_1) - \{u'_1\}$ and $u'_1 \in T_C^{U^+}(u'_1)$, we get $|T_C^U(u'_1)| = |T_C^{U^+}(u'_1)| - 1$. Since $T_C^{U'}(u'_1) = \{u'_1\}$, we have $|T_C^{U'}(u'_1)| = 1$. Then we get $2|T_C^U(u'_1)| + |T_C^{U'}(u'_1)| = 2(|T_C^{U^+}(u'_1)| - 1) + 1 = 2|T_C^{U^+}(u'_1)| - 1$. Therefore, Eq. (9) degenerates into Eq. (9) in [57]. In other words, the case of incrementally getting the knowledge granularity under the circumstance of adding objects one by one is a special case of adding multiple objects at a time.

Example 3 (Continued from Example 1) Table 2 shows an incremental data set $U' = \{u'_1, u'_2, u'_3, u'_4\}$. According to Definition 2, we have $T_C^U(u'_1) = \{u_1\}$, $T_C^U(u'_2) = \{u_1\}$, $T_C^U(u'_3) = \{u_2, u_3, u_4, u_5\}$, $T_C^U(u'_4) = \emptyset$, $T_C^{U'}(u'_1) = \{u'_1, u'_2\}$, $T_C^{U'}(u'_2) = \{u'_1, u'_2\}, T_C^{U'}(u'_3) = \{u'_3\}, T_C^{U'}(u'_4) = \{u'_4\}$. According to Example 1, we have $GK_U(C) = \frac{9}{25}$. Then, we can get $GK_{U^+}(C) = \frac{1}{(5+4)^2}(5^2 \times \frac{9}{25} + 2(1+1+4+0) + (2+2+1+1)) = \frac{27}{81}$.

Proposition 2 *Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system. $U' = \{u'_1, u'_2, \ldots, u'_s\}$ is the incremental object set. $T_C^U(u'_j)$ and $T_{C\cup D}^U(u'_j)$ are the tolerance classes of u'_j with respect to attribute sets C and $C \cup D$ on U, respectively. $T_C^{U'}(u'_j)$ and $T_{C\cup D}^{U'}(u'_j)$ are the tolerance classes of u'_j with respect to attribute sets C and $C \cup D$ on U', respectively. The knowledge granularity of C relative to D on U^+ is*

$$GK_{U^+}(D|C) = \frac{1}{(n+s)^2}(n^2 GK_U(D|C) + 2\sum_{j=1}^{s}(|T_C^U(u'_j)| - |T_{C\cup D}^U(u'_j)|) + \sum_{j=1}^{s}(|T_C^{U'}(u'_j)| - |T_{C\cup D}^{U'}(u'_j)|)) \tag{10}$$

where $|.|$ denotes the cardinality of a set.

Proof According to Definition 4 and Proposition 1, we have

$$\begin{aligned}
&GK_{U^+}(D|C) \\
&= GK_{U^+}(C) - GK_{U^+}(C \cup D) \\
&= \frac{1}{(n+s)^2}(n^2 GK_U(C) + 2\sum_{j=1}^{s}|T_C^U(u'_j)| \\
&\quad + \sum_{j=1}^{s}|T_C^{U'}(u'_j)|) - \frac{1}{(n+s)^2}(n^2 GK_U(C \cup D) \\
&\quad + 2\sum_{j=1}^{s}|T_{C\cup D}^U(u'_j)| + \sum_{j=1}^{s}|T_{C\cup D}^{U'}(u'_j)|) \\
&= \frac{1}{(n+s)^2}(n^2 GK_U(D|C) + 2\sum_{j=1}^{s}(|T_C^U(u'_j)| \\
&\quad - |T_{C\cup D}^U(u'_j)|) + \sum_{j=1}^{s}(|T_C^{U'}(u'_j)| - |T_{C\cup D}^{U'}(u'_j)|))
\end{aligned}$$

□

If $U' = \{u'_1\}$, that denotes adding objects one by one to the original object set U, then Eq. (10) can be written as

$$\frac{1}{(n+1)^2}(n^2 GK_U(D|C) + 2(|T_C^U(u'_1)| - |T_{C\cup D}^U(u'_1)|) + (|T_C^{U'}(u'_1)| - |T_{C\cup D}^{U'}(u'_1)|)).$$

From $T_C^U(u'_1) = T_C^{U^+}(u'_1) - \{u'_1\}$ and $u'_1 \in T_C^{U^+}(u'_1)$, we get $|T_C^U(u'_1)| = |T_C^{U^+}(u'_1)| - 1$. Similarly, $|T_{C\cup D}^U(u'_1)| = |T_{C\cup D}^{U^+}(u'_1)| - 1$. Since $T_C^{U'}(u'_1) = \{u'_1\}$, we have $|T_C^{U'}(u'_1)| = 1$. Similarly, $|T_{C\cup D}^{U'}(u'_1)|)| = 1$. Finally, Eq. (10) can be written as $\frac{1}{(n+1)^2}(n^2 GK_U(D|C) + 2(|T_C^{U^+}(u'_1)| - |T_{C\cup D}^{U^+}(u'_1)|))$. Therefore, Eq. (10) can degenerate into Eq. (10) in [57]. In other words, the case of incrementally getting the relative knowledge granularity under the circumstance of adding objects one by one is a special case of adding multiple objects at a time.

Example 4 (Continued from Example 3) According to Definition 2, we have $T_{C\cup D}^U(u'_1) = \{u_1\}$, $T_{C\cup D}^U(u'_2) = \emptyset$, $T_{C\cup D}^U(u'_3) = \{u_2, u_3\}$, $T_{C\cup D}^U(u'_4) = \emptyset$, $T_{C\cup D}^{U'}(u'_1) = \{u'_1\}$, $T_{C\cup D}^{U'}(u'_2) = \{u'_2\}, T_{C\cup D}^{U'}(u'_3) = \{u'_3\}, T_{C\cup D}^{U'}(u'_4) = \{u'_4\}$. According to Example 2, we have $GK_U(D|C) = \frac{4}{25}$. Then, we get $GK_{U^+}(D|C) = \frac{1}{(5+4)^2}(5^2 \times \frac{4}{25} + 2(1-1+1-0+4-2+0-0) + (2-1+2-1+1-1+1-1)) = \frac{12}{81}$.

Proposition 3 *Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system. $U' = \{u'_1, u'_2, \ldots, u'_s\}$ is the incremental object set. $T_C^U(u'_j), T_{C\cup D}^U(u'_j), T_{C-\{a\}}^U(u'_j)$ and $T_{(C-\{a\})\cup D}^U(u'_j)$ are the tolerance classes of u'_j with respect to attribute sets C, $C \cup D$, $C - \{a\}$ and $(C - \{a\}) \cup D$ on U, respectively. $T_C^{U'}(u'_j), T_{C\cup D}^{U'}(u'_j), T_{C-\{a\}}^{U'}(u'_j)$ and $T_{(C-\{a\})\cup D}^{U'}(u'_j)$ are the tolerance classes of u'_j with respect to attribute sets C, $C \cup D$,*

Table 2 Incremental data set to the incomplete decision system 1 in Table 1

U	a_1	a_2	a_3	a_4	a_5	d
u'_1	1	0	0	0	*	0
u'_2	1	0	0	*	0	1
u'_3	0	*	1	0	1	0
u'_4	*	1	1	0	1	1

$C-\{a\}$ and $(C-\{a\})\cup D$ on U', respectively. Then $\forall a \in C$, the inner significance of a in C on U^{+} is

$$Sig_{U^+}^{inner}(a,C,D) = \frac{1}{(n+s)^2}(n^2 Sig_U^{inner}(a,C,D) + 2\sum_{j=1}^{s}(|T_{C-\{a\}}^{U}(u_j')| - |T_{(C-\{a\})\cup D}^{U}(u_j')| - |T_C^U(u_j')| + |T_{C\cup D}^{U}(u_j')|) + \sum_{j=1}^{s}(|T_{C-\{a\}}^{U'}(u_j')| - |T_{(C-\{a\})\cup D}^{U'}(u_j')| - |T_C^{U'}(u_j')| + |T_{C\cup D}^{U'}(u_j')|)) \tag{11}$$

Proof According to Definition 5 and Proposition 2, we can get

$$\begin{aligned}
&Sig_{U^+}^{inner}(a,C,D)\\
&= GK_{U^+}(D|C-\{a\}) - GK_{U^+}(D|C)\\
&= \frac{1}{(n+s)^2}(n^2 GK_U(D|C-\{a\}) + 2\sum_{j=1}^{s}(|T_{C-\{a\}}^{U}(u_j')| - |T_{(C-\{a\})\cup D}^{U}(u_j')|) + \sum_{j=1}^{s}(|T_{C-\{a\}}^{U'}(u_j')| - |T_{(C-\{a\})\cup D}^{U'}(u_j')|))\\
&\quad - \frac{1}{(n+s)^2}(n^2 GK_U(D|C) + 2\sum_{j=1}^{s}(|T_C^U(u_j')| - |T_{C\cup D}^{U}(u_j')|) + \sum_{j=1}^{s}(|T_C^{U'}(u_j')| - |T_{C\cup D}^{U'}(u_j')|))\\
&= \frac{1}{(n+s)^2}(n^2 Sig_U^{inner}(a,C,D) + 2\sum_{j=1}^{s}(|T_{C-\{a\}}^{U}(u_j')| - |T_{(C-\{a\})\cup D}^{U}(u_j')| - |T_C^U(u_j')| + |T_{C\cup D}^{U}(u_j')|)\\
&\quad + \sum_{j=1}^{s}(|T_{C-\{a\}}^{U'}(u_j')| - |T_{(C-\{a\})\cup D}^{U'}(u_j')| - |T_C^{U'}(u_j')| + |T_{C\cup D}^{U'}(u_j')|))
\end{aligned}$$

□

Similarly, if $U' = \{u_1'\}$, then Eq. (11) can be written as

$$\frac{1}{(n+s)^2}(n^2 Sig_U^{inner}(a,C,D) + 2(|T_{C-\{a\}}^{U}(u_1')| - |T_{(C-\{a\})\cup D}^{U}(u_1')| - |T_C^U(u_1')| + |T_{C\cup D}^{U}(u_1')|)).$$

From $T_C^U(u_1') = T_C^{U^+}(u_1') - \{u_1'\}$ and $u_1' \in T_C^{U^+}(u_1')$, we get $|T_C^U(u_1')| = |T_C^{U^+}(u_1')| - 1$. Similarly, $|T_{C\cup D}^{U}(u_1')| = |T_{C\cup D}^{U^+}(u_1')| - 1$, $|T_{C-\{a\}}^{U}(u_1')| = |T_{C-\{a\}}^{U^+}(u_1')| - 1$ and $|T_{(C-\{a\})\cup D}^{U}(u_1')| = |T_{(C-\{a\})\cup D}^{U^+}(u_1')| - 1$. Since $T_C^{U'}(u_1') = \{u_1'\}$, we have $|T_C^{U'}(u_1')| = 1$. Similarly, $|T_{C\cup D}^{U'}(u_1')|)| = 1$, $|T_{C-\{a\}}^{U'}(u_j')| = 1$ and $|T_{(C-\{a\})\cup D}^{U'}(u_j')| = 1$. Finally, Eq. (11) can be written as

$$\frac{1}{(n+s)^2}(n^2 Sig_U^{inner}(a,C,D) + 2(|T_{C-\{a\}}^{U^+}(u_1')| - |T_C^{U^+}(u_1')| - |T_{(C-\{a\})\cup D}^{U^+}(u_1')| + |T_{C\cup D}^{U^+}(u_1')|)).$$

Therefore, Eq. (11) degenerates into Eq. (11) in [57]. In other words, the case of incrementally getting the inner significance under the circumstance of adding objects one by one is a special case of adding multiple objects at a time.

Proposition 4 *Let $IDS = \langle U, C\cup D, V, f\rangle$ be an incomplete decision system. $U' = \{u_1', u_2', \dots, u_s'\}$ is the incremental object set. $T_B^U(u_j')$, $T_{B\cup D}^{U}(u_j')$, $T_{B\cup\{a\}}^{U}(u_j')$ and $T_{B\cup\{a\}\cup D}^{U}(u_j')$ are the tolerance classes of u_j' with respect to attribute sets B, $B\cup D$, $B\cup\{a\}$ and $B\cup\{a\}\cup D$ on U, respectively. $T_B^{U'}(u_j')$, $T_{B\cup D}^{U'}(u_j')$, $T_{B\cup\{a\}}^{U'}(u_j')$ and $T_{B\cup\{a\}\cup D}^{U'}(u_j')$ are the tolerance classes of u_j' with respect to attribute sets B, $B\cup D$, $B\cup\{a\}$ and $B\cup\{a\}\cup D$ on U', respectively. Then $\forall a \in (C-B)$, the outer significance of a in B on U^+ is*

$$Sig_{U^+}^{outer}(a,B,D) = \frac{1}{(n+s)^2}(n^2 Sig_U^{outer}(a,B,D) + 2\sum_{j=1}^{s}(|T_B^U(u_j')| - |T_{B\cup D}^{U}(u_j')| - |T_{B\cup\{a\}}^{U}(u_j')| + |T_{B\cup\{a\}\cup D}^{U}(u_j')|) + \sum_{j=1}^{s}(|T_B^{U'}(u_j')| - |T_{B\cup D}^{U'}(u_j')| - |T_{B\cup\{a\}}^{U'}(u_j')| + |T_{B\cup\{a\}\cup D}^{U'}(u_j')|)) \tag{12}$$

Proof According to Definition 7 and Proposition 2, we can get

$$
\begin{aligned}
&Sig_{U^+}^{outer}(a,B,D)\\
&= GK_{U^+}(D|B) - GK_{U^+}(D|(B\cup\{a\}))\\
&= \frac{1}{(n+s)^2}(n^2GK_U(D|B) + 2\sum_{j=1}^{s}(|T_B^U(u_j')|\\
&\quad - |T_{B\cup D}^U(u_j')|) + \sum_{j=1}^{s}(|T_B^{U'}(u_j')| - |T_{B\cup D}^{U'}(u_j')|))\\
&\quad - \frac{1}{(n+s)^2}(n^2GK_U(D|B\cup\{a\})\\
&\quad + 2\sum_{j=1}^{s}(|T_{B\cup\{a\}}^U(u_j')| - |T_{B\cup\{a\}\cup D}^U(u_j')|)\\
&\quad + \sum_{j=1}^{s}(|T_{B\cup\{a\}}^{U'}(u_j')| - |T_{B\cup\{a\}\cup D}^{U'}(u_j')|))\\
&= \frac{1}{(n+s)^2}(n^2Sig_U^{outer}(a,B,D) + 2\sum_{j=1}^{s}(|T_B^U(u_j')|\\
&\quad - |T_{B\cup D}^U(u_j')| - |T_{B\cup\{a\}}^U(u_j')| + |T_{B\cup\{a\}\cup D}^U(u_j')|)\\
&\quad + \sum_{j=1}^{s}(|T_B^{U'}(u_j')| - |T_{B\cup D}^{U'}(u_j')| - |T_{B\cup\{a\}}^{U'}(u_j')|\\
&\quad + |T_{B\cup\{a\}\cup D}^{U'}(u_j')|))
\end{aligned}
$$

□

Similarly, if $U' = \{u_1'\}$, then Eq. (12) can be written as

$$
\frac{1}{(n+1)^2}(n^2Sig_U^{outer}(a,B,D) + 2(|T_B^{U^+}(u_j')| - |T_{B\cup D}^{U^+}(u_j')|
- |T_{B\cup\{a\}}^{U^+}(u_j')| + |T_{B\cup\{a\}\cup D}^{U^+}(u_j')|).
$$

Therefore, the case of incrementally getting the outer significance under the circumstance of adding objects one by one is a special case of adding multiple objects at a time.

3.2 An incremental reduction algorithm for IDS when adding multiple objects

The traditional heuristic attribute reduction algorithm based on knowledge granularity for decision systems (THA) was discussed in detail in [57]. Based on the incremental mechanisms of knowledge granularity above, this subsection introduces an incremental attribute reduction algorithm (see Algorithm 1) using knowledge granularity when adding multiple objects to the decision system. At last, a brief comparison of time complexity between incremental reduction algorithm and traditional heuristic reduction algorithm is given.

Algorithm 1 Knowledge Granularity based Incremental Reduction algorithm when Adding Multiple objects(KGIRA-M)

Input: An incomplete decision system $IDS = \langle U, C\cup D, V, f\rangle$, the reduct RED_U on U, the incremental object set U'.
Output: A new reduct RED_{U^+} on U^+ after adding U' to IDS.

1: $B \leftarrow RED_U$;
2: Compute $\sum_{j=1}^{|U'|}|T_B^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_B^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_{C\cup D}^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^{U'}(u_j')|$ and $\sum_{j=1}^{|U'|}|T_{C\cup D}^{U'}(u_j')|$.
3: Compute $GK_{U^+}(D|B)$ and $GK_{U^+}(D|C)$(according to Proposition 2);
4: **if** $GK_{U^+}(D|B) = GK_{U^+}(D|C)$ **then**
5: go to 18;
6: **end if**
7: **while** $GK_{U^+}(D|B) \neq GK_{U^+}(D|C)$ **do**
8: **for** each $(a_i) \in (C-B)$ **do**
9: Compute $Sig_{U^+}^{outer}(a_i,B,D)$(according to Proposition 4);
10: **end for**
11: $a_0 = max\{Sig_{U^+}^{outer}(a_i,B,D), a_i \in (C-B)\}$;
12: $B \leftarrow (B\cup\{a_0\})$;
13: **end while**
14: **for** each $(a_i) \in B$ **do**
15: **if** $GK_{U^+}(D|(B-\{a\})) = GK_{U^+}(D|C)$ **then**
16: $B \leftarrow (B-\{a_i\})$;
17: **end if**
18: **end for**
19: $RED_{U^+} \leftarrow B$;
20: return RED_{U^+};

In Algorithm 1, U^+ denotes the object set of the new decision system after adding the incremental object set U' to the original object set U. The detailed execution process of Algorithm 1 is as follows. In Step 1, the reduct RED_U on U is assigned to B. In Step 2, $\sum_{j=1}^{|U'|}|T_B^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_B^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_{C\cup D}^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^{U'}(u_j')|$ and $\sum_{j=1}^{|U'|}|T_{C\cup D}^{U'}(u_j')|$ are calculated respectively. In Step 3, the knowledge granularity of B relative to D on U^+ ($GK_{U^+}(D|B)$) and the knowledge granularity of C relative to D on U^+($GK_{U^+}(D|C)$) are calculated respectively according to Proposition 2. In Step 4, $GK_{U^+}(D|B)$ is compared with $GK_{U^+}(D|C)$, if they are equal, then the algorithm flow skips to Step 19, which indicates that the reducts of U^+ and U are identical. Otherwise, the loop consisting of Steps 7–13 executes. Steps 8–10 constitute a loop to calculate the outer significance of an certain $a_i \in (C-B)$ which is $Sig_{U^+}^{outer}(a_i,B,D)$ according to Proposition 4. In Step 11, a_0 is used to store the attribute that has the maximum value of outer significance. In Step 12, a_0 is added to B. Steps 14–18 are actually used to delete attributes that satisfy the second condition in Definition 8 through a loop with $|B|$ times. In Step 15, $GK_{U^+}(D|(B-\{a\}))$ is compared with $GK_{U^+}(D|C)$, if they are equal, then a_i is deleted from B, which indicates that a_i is redundant and can not be in the reduct of U^+. In Step 19, B is assigned to RED_{U^+}. In Step 20, RED_{U^+} is returned as the result of Algorithm 1.

The time complexity of Algorithm 1 is $O(|U'||U||B|(|C|-|B|))$. According to Definition 2, the time complexity of computing $T_B^U(u_j')$ is $O(|U||B|)$. Then the time complexity of computing $\sum_{j=1}^{|U'|}|T_B^U(u_j')|$ is $O(|U'||U||B|)$. Similarly, we get the time complexity of computing $\sum_{j=1}^{|U'|}|T_B^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_C^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^U(u_j')|$, $\sum_{j=1}^{|U'|}|T_{B\cup D}^{U'}(u_j')|$, $\sum_{j=1}^{|U'|}|T_{C\cup D}^U(u_j')|$ and $\sum_{j=1}^{|U'|}|T_{C\cup D}^{U'}(u_j')|$ in Step 2 respectively. According to Proposition 2, the time complexity of Steps 2–3 is $O(|U'||U||B|+|U'||U|(|B|+1)+|U'|^2|B|+|U'|^2(|B|+1)+|U'||U||C|+|U'||U|(|C|+1)+|U'|^2|C|+|U'|^2(|C|+1))\approx O(|U'||U||C|)$. According to Proposition 4, the time complexity of computing $Sig_{U^+}^{outer}(a_i,B,D)$ is $O(|U'||U||B|)$. Then the time complexity of Steps 7–13 is $O(|U'||U||B|(|C|-|B|))$. Similarly, the time complexity of Steps 14–18 is $O(|U'||U||B|^2)$. Hence, the time complexity of Algorithm 1 is $O(|U'||U||C|+|U'||U||B|(|C|-|B|)+|U'||U||B|^2)\approx O(|U'||U||B|(|C|-|B|))$.

Two algorithms for comparison with algorithm KGIRA-M were introduced in [57]. One is the non-incremental algorithm called algorithm THA which is converted as algorithm THA-M in this paper to be suitable for incomplete decision systems under the situation of adding multiple objects at a time. Another is incremental algorithm called algorithm KGIRA which is called as algorithm KGIRA-S in this paper for incomplete decision systems under the situation of adding one object at a time. Since the number of objects is $|U^+|$ after adding the incremental object set U' to the original object set U, the time complexity of THA-M is actually $O(|U^+|^2|C|^2)$. Because $|U'|\le|U^+|$, $|U|\le|U^+|$ and $|B|\le|C|$, $O(|U'||U||B|(|C|-|B|$ is much lower than $O(|U^+|^2|C|^2)$. Hence, the time complexity of algorithm KGIRA-M is much smaller than that of algorithm THA-M. Since the number of objects is $|U^+|$ after adding the object set U' to the original object set U, the time complexity of KGIRA-S is $O(|U'||U||C|^2)$. Because $|B|\le|C|$, $O(|U'||U||B|(|C|-|B|))$ is lower than $O(|U'||U||C|^2)$. Hence, the time complexity of algorithm KGIRA-M is smaller than that of algorithm KGIRA-S. Since $|U'|\ll|U^+|$ and $|U|<|U^+|$, we can draw the conclusion that the time complexity of KGIRA-S is also much smaller than that of algorithm THA-M.

3.3 An incremental mechanism to calculate knowledge granularity for IDS when deleting multiple objects

For the sake of convenience, let $U=\{u_1,u_2,\dots,u_n\}$ be the original object set in the incomplete decision system $IDS=\langle U,C\cup D,V,f\rangle$. $U''=\{u_1'',u_2'',\dots,u_t''\}\subset U$ denotes the object set whose objects will be deleted from U. For simplicity, U^- denotes $U-U''$ in the following.

Proposition 5 *Let $IDS=\langle U,C\cup D,V,f\rangle$ be an incomplete decision system. $T_C^{U^-}(u_j'')$ and $T_C^{U''}(u_j'')$ are the tolerance classes of u_j'' with respect to attribute set C on U^- and U'', respectively. The knowledge granularity of C on U^- is*

$$GK_{U^-}(C)=\frac{1}{(n-t)^2}(n^2GK_U(C)-2\sum_{j=1}^{t}|T_C^{U^-}(u_j'')|-\sum_{j=1}^{t}|T_C^{U''}(u_j'')|) \tag{13}$$

Proof When $u_j''(1\le j\le t)$ is deleted from U, the tolerance class of $u_i(1\le i\le n-t)$ on U^- is

$$T_C^{U^-}(u_i)=\begin{cases}T_C^U(u_i)-\{u_j''\}, & (u_i,u_j'')\in T_C^U\\ T_C^U(u_i), & (u_i,u_j'')\notin T_C^U\end{cases}$$

Then we can get

$$T_C^{U^-}(u_i)=T_C^U(u_i)-T_C^{U''}(u_i),\quad 1\le i\le n-t$$

and

$$T_C^{U^-}(u_j'')=T_C^U(u_j'')-T_C^{U''}(u_j''),\quad 1\le j\le t.$$

Since $T_C^{U''}(u_i)\subset T_C^U(u_i)$, we can get

$$|T_C^{U^-}(u_i)|=|T_C^U(u_i)|-|T_C^{U''}(u_i)|,\quad 1\le i\le n-t.$$

Similarly, we can get

$$|T_C^{U^-}(u_j')|=|T_C^U(u_j'')|-|T_C^{U''}(u_j'')|,\quad 1\le j\le t.$$

If $u_j''\in T_C^{U''}(u_i)$ then $u_i\in T_C^{U^-}(u_j'')$, $1\le i\le n-t$, $1\le j\le t$ and vice versa, because tolerance relation is symmetric. Obviously, we can get $\sum_{i=1}^{|U^-|}|T_C^{U''}(u_i)|=\sum_{j=1}^{|U''|}|T_C^{U^-}(u_j'')|$. According to Definition 3, the knowledge granularity of C on U^- is described as follows:

$$\begin{aligned}&GK_{U^-}(C)\\&=\frac{1}{|U^-|^2}\sum_{i=1}^{|U^-|}|T_C^{U^-}(u_i)|\\&=\frac{1}{|U^-|^2}\left(\sum_{i=1}^{|U|}|T_C^U(u_i)|-\sum_{i=1}^{|U^-|}|T_C^{U''}(u_i)|\right.\\&\quad\left.-\sum_{j=1}^{|U''|}|T_C^{U^-}(u_j'')|-\sum_{j=1}^{|U''|}|T_C^{U''}(u_j'')|\right)\\&=\frac{1}{(n-t)^2}(n^2GK_U(C)-2\sum_{j=1}^{t}|T_C^{U^-}(u_j'')|\\&\quad-\sum_{j=1}^{t}|T_C^{U''}(u_j'')|)\end{aligned}$$

□

If $U'' = \{u''_1\}$, which denotes deleting objects one by one from the original object set U, then Eq. (13) can be written as $\frac{1}{(n-1)^2}(n^2 GK_U(C) - 2|T_C^{U^-}(u''_1)| - |T_C^{U''}(u''_1)|)$. Because $T_C^U(u''_1) = T_C^{U^-}(u''_1) \cup \{u''_1\}$ and $u''_1 \notin T_C^{U^-}(u''_1)$, we have $|T_C^{U^-}(u''_1)| = |T_C^U(u''_1)| - 1$. Then Eq. (13) can be written as $\frac{1}{(n-1)^2}(n^2 GK_U(C) - (2|T_C^U(u''_1)| - 1))$. Therefore, Eq. (13) degenerates into Eq. (13) in [57]. In other words, the case of incrementally getting the relative knowledge granularity under the circumstance of deleting objects one by one is a special case of deleting multiple objects at a time.

Proposition 6 *Let $IDS = \langle U, C \cup D, V, f \rangle$ be an incomplete decision system. u_n is the diminishing object. $T_C(u_n)$ and $T_{C\cup D}(u_n)$ are the tolerance classes of u_n with respect to attribute sets C and $C \cup D$ on U, respectively. The knowledge granularity of C relative to D on U^- is*

$$GK_{U^-}(D|C) = \frac{1}{(n-t)^2}(n^2 GK_U(D|C) - 2\sum_{j=1}^{t}(|T_C^{U^-}(u''_j)| - |T_{C\cup D}^{U^-}(u''_j)|) - \sum_{j=1}^{t}(|T_C^{U''}(u''_j)| - |T_{C\cup D}^{U''}(u''_j)|)) \tag{14}$$

Proof According to Definition 4 and Proposition 5, we have

$$\begin{aligned} &GK_{U^-}(D|C) \\ &= GK_{U^-}(C) - GK_{U^-}(C \cup D) \\ &= \frac{1}{(n-t)^2}(n^2 GK_U(C) - 2\sum_{j=1}^{|U''|}|T_C^{U^-}(u''_j)| \\ &\quad - \sum_{j=1}^{|U''|}|T_C^{U''}(u''_j)|) - \frac{1}{(n-t)^2}(n^2 GK_U(C \cup D) \\ &\quad - 2\sum_{j=1}^{|U''|}|T_{C\cup D}^{U^-}(u''_j)| - \sum_{j=1}^{|U''|}|T_{C\cup D}^{U''}(u''_j)|) \\ &= \frac{1}{(n-t)^2}(n^2 GK_U(D|C) - 2\sum_{j=1}^{t}(|T_C^{U^-}(u''_j)| \\ &\quad - |T_{C\cup D}^{U^-}(u''_j)|) - \sum_{j=1}^{t}(|T_C^{U''}(u''_j)| - |T_{C\cup D}^{U''}(u''_j)|)) \end{aligned}$$

□

If $U'' = \{u''_1\}$, then Eq. (14) can be written as

$\frac{1}{(n-1)^2}(n^2 GK_U(D|C) - 2(|T_C^{U^-}(u''_1)| - |T_{C\cup D}^{U^-}(u''_1)|) - (|T_C^{U''}(u''_1)| - |T_{C\cup D}^{U''}(u''_1)|))$.

Since $T_C^U(u''_1) = T_C^{U^-}(u''_1) \cup \{u''_1\}$ and $u''_1 \notin T_C^{U^-}(u''_1)$, we get $|T_C^{U^-}(u''_1)| = |T_C^U(u''_1)| - 1$. Similarly, $|T_{C\cup D}^{U^-}(u''_1)| = |T_{C\cup D}^U(u''_1)| - 1$. Since $T_C^{U''}(u''_1) = \{u''_1\}$, we have $|T_C^{U''}(u''_1)| = 1$. Similarly, $|T_{C\cup D}^{U''}(u''_1)| = 1$. Then Eq. (14) can be written as $\frac{1}{(n-1)^2}(n^2 GK_U(C) - 2(|T_C^U(u''_1)| - |T_{C\cup D}^U(u''_1)|))$.

Therefore, Eq. (14) degenerates into Eq. (14) in [57]. In other words, the case of incrementally getting the relative knowledge granularity under the circumstance of deleting objects one by one is a special case of deleting multiple objects at a time. To sum up, all situations of handling objects one by one in [57] are special cases of those of handling multiple objects at a time for getting new reducts in this paper.

3.4 An incremental reduction algorithm for IDS when deleting multiple objects

Based on the incremental mechanisms of knowledge granularity above, this subsection introduces an incremental attribute reduction algorithm (see Algorithm 2) when deleting multiple objects from an incomplete decision system.

Algorithm 2 Knowledge Granularity based Incremental Reduction algorithm when Deleting Multiple objects(KGIRD-M)

Input: An incomplete decision system $IDS = \langle U, C \cup D, V, f \rangle$, the reduct RED_U on U, the diminishing object set U''.
Output: A new reduct RED_{U^-} on U^- after deleting U'' from IDS.
1: $B \leftarrow RED_U$;
2: Compute $\sum_{j=1}^{|U''|}|T_C^{U^-}(u''_j)|$, $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U^-}(u''_j)|$, $\sum_{j=1}^{|U''|}|T_C^{U''}(u''_j)|$, $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U''}(u''_j)|$.
3: Compute $GP_{U^-}(D|C)$(according to Proposition 6);
4: **for** $each(a_i) \in B$ **do**
5:   **if** $GP_{U^-}(D|(B - \{a_i\})) = GP_{U^-}(D|C)$ **then**
6:     $B \leftarrow (B - \{a_i\})$;
7:   **end if**
8: **end for**
9: $RED_{U^-} \leftarrow B$;
10: return RED_{U^-};

In Algorithm 2, U^- denotes the object set of the new decision system after deleting objects in U'' from the original object set U. The detailed execution process of Algorithm 2 is as follows. In Step 1, the reduct RED_U on U is assigned to B. In Step 2, $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U^-}(u''_j)|$, $\sum_{j=1}^{|U''|}|T_C^{U^-}(u''_j)|$, $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U''}(u''_j)|$ and $\sum_{j=1}^{|U''|}(|T_C^{U''}(u''_j)|$ are calculated respectively. In Step 3, the knowledge granularity of C relative to D on $U^-(GP_{U^-}(D|C))$ is calculated according to Proposition 6. Steps 4–8 are actually used to delete attributes that satisfies the second condition in Definition 8 through a loop with $|B|$ times. In Step 5, $GP_{U^-}(D|(B-\{a_i\}))$ is compared with $GP_{U^-}(D|(B-\{a_i\}))$. If $GP_{U^-}(D|(B-\{a_i\}))$ is equal to $GP_{U^-}(D|C)$, then a_i is deleted from B, which indicates a_i is redundant and can not be in the reduct of U^-. In Step 9, B is assigned to RED_{U^-}. In Step 10, RED_{U^-} is returned as the result of Algorithm 2.

The time complexity of Algorithm 2 is $O(|U''||U^-||B|^2)$. According to Definitions 1 and 2, the time complexity of computing $\sum_{j=1}^{|U''|}|T_C^{U^-}(u''_j)|$ is $O(|U''||U^-||C|)$, the time complexity of computing $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U^-}(u''_j)|$ is $O(|U''||U^-|(|C|+1))$, the time complexity of computing $\sum_{j=1}^{|U''|}(|T_C^{U''}(u''_j)|$ is $O(|U''|^2|C|)$ and the time complexity of computing $\sum_{j=1}^{|U''|}|T_{C\cup D}^{U''}(u''_j)|$ is $O(|U''|^2(|C|+1))$. Then according to Proposition 6, the time complexity of Steps 2–3 is $O(|U''||U^-||C|+|U''||U^-|(|C|+1)+|U''|^2|C|+|U''|^2(|C|+1))\approx O(|U''||U^-||C|)$. In Step 5, the time complexity of computing $GP_{U^-}(D|(B-\{a_i\}))$ is similar to that of $GP_{U^-}(D|C)$. Since $|U''|\ll|U^-|$, the time complexity of Steps 4–8 is $O(|B|(|U''||U^-||B|+|U''||U^-|(|B|+1)+|U''|^2|B|+|U''|^2(|B|+1)))\approx O(|U''||U^-||B|^2)$. In general, if a reduct obtained is of practical value, $|C|\le|B|^2$ should be satisfied. Hence, the time complexity of Algorithm 2 is $O(|U''||U^-||C|+|U''||U^-||B|^2)\approx O(|U''||U^-||B|^2)$.

Two algorithms for comparison with algorithm KGIRD-M were introduced in [57]. One is the non-incremental algorithm called algorithm THA which is converted as algorithm THD-M in this paper to be suitable for incomplete decision systems under the situation of deleting multiple objects at a time. Another is the incremental algorithm called algorithm KGIRD which is called as algorithm KGIRD-S in this paper for incomplete decision systems under the situation of deleting one object at a time. Since the number of objects is $|U^-|$, the time complexity of THD-M is actually $O(|U^-|^2|C|^2)$ when deleting the object set U'' from the original object set U. Because $|B|\le|C|$ and $|U''|\le|U^-|$, $O(|U''||U^-||B|^2)$ is lower than $O(|U^-|^2|C|^2)$. Hence, the time complexity of algorithm KGIRD-M is much smaller than that of algorithm THD-M. Since the number of objects is $|U^-|$ after deleting the object set U'' from the original object set U, the time complexity of KGIRD-S is $O(|U''||U^-||C|^2)$. Because $|B|\le|C|$, $O(|U''||U^-||B|^2)$ is lower than $O(|U''||U^-||C|^2)$. Hence, the time complexity of algorithm KGIRD-M is smaller than that of algorithm KGIRD-S. Since $|U''|\ll|U^-|$, we can draw the conclusion that the time complexity of KGIRD-S is also much smaller than that of algorithm THD-M.

4 Empirical experiments

4.1 A description of data sets and experimental environment

The proposed incremental attribute reduction algorithms are tested on data sets available from the University of California Irvine (UCI) Repository of Machine Learning Database (http://archive.ics.uci.edu/ml) and the Kent Ridge Biomedical Data Set Repository (http://leo.ugr.es/elvira/DBCRepository/). In order to get incomplete decision systems from complete data sets, 5% of the attribute values randomly selected from each complete data set are deleted. The characteristics of the data sets are summarized in Table 3.

The proposed algorithms and comparative algorithms are implemented in Eclipse IDE for Java Developers with Neon.3 Release (4.6.3) version using JDK 1.8.0_111 version and they have been carried out on a personal computer with the following specification: Intel Core i3-3120M 2.5GHz CPU, 4.0 GB of memory and 64-bit Win7.

4.2 Performance comparison on computational time

In the comparative experiments under the condition that objects are added to decision systems, each original data set is divided into ten parts averagely according to the number of objects. Two parts of each original data set is set as the basic data set and the rest are set as the incremental data set which is going to be added to the basic data set in subsequent steps. In algorithm KGIRA-S, objects in incremental data set are added to basic data set one by one. Once all objects in one part of the original data set are added to the basic data set, the spent time is recorded, until all parts in the incremental data set are added to the basic data set. In algorithms KGIRA-M and THA-M, one part in the incremental data set is added to the basic data set at a time which makes a note of the time. In each sub-figure of Fig. 1, the x-axis represents the percentage of the number of objects in the basic data set to that of objects in the original data set and the y-axis represents the value of computational time. Asterisk marked lines denote the computational time of algorithm KGIRA-M, circle marked lines denote the computational time of algorithm KGIRA-S and plus marked lines denote the computational time of algorithm THA-M.

From Fig. 1, for each data set, the computational time of three algorithms increases with the number of objects which are added to the basic data set. The computational time of algorithms KGIRA-M and KGIRA-S is much shorter than that of the algorithm THA-M, which illustrates that incremental algorithms are more efficient than the non-incremental algorithm. Moreover, the computational time

Table 3 A description of data sets

	Data sets	Objects	Attributes	Classes
1	Wine	178	13	3
2	Lymphography	148	18	4
3	Ionosphere	351	34	2
4	Dermatology	366	33	6
5	LSVT voice	126	310	2
6	Musk1	476	166	2
7	Musk2	707	166	2
8	Mice Protein	1080	80	8
9	Multiple Features	2000	216	10
10	Colon	62	2000	2
11	AMLALL	72	7129	2

of algorithm KGIRA-M is shorter than that of algorithm KGIRA-S, which shows that processing multiple objects at a time is more efficient than processing one object at a time by the incremental approaches. It is clearly that the proposed algorithm KGIRA-M has the highest efficiency among the three comparative algorithms.

In the comparative experiments of deleting objects from the original data set, each original data set is set as basic data set and divided into ten parts (Part 1, Part2, ..., Part10) averagely according to the number of objects. In algorithm KGIRD-S, objects are deleted from the basic data set until eight parts (Part 1, Part2, ..., Part8) of the original data set are removed. In each part to be removed, objects are deleted one by one. Once one part is deleted from the basic data set, the accumulative computing time is recorded. In algorithms KGIRD-M and THD-M, objects from one part of the basic data set are deleted at a time which makes a note of the time. The x-axis represents the percentage of the number of objects retained to that of objects in the original data set and the y-axis represents the value of computational time. The accumulative time is recorded for all algorithms. Asterisk marked lines denote computational time of algorithm KGIRD-M, circle marked lines denote computational time of algorithm KGIRD-S and plus marked lines denote computational time of algorithm THD-M.

From Fig. 2, for each data set, the computational time of three algorithms increases when the number of objects deleted from decision system increases. The computational time of algorithms KGIRD-M and KGIRD-S is much smaller than that of the algorithm THD-M, which illustrates that incremental algorithms are more efficient than the non-incremental algorithm. Moreover, the computational time of algorithm KGIRD-M is smaller than that of the algorithm KGIRD-S, which shows that processing multiple objects at a time is more efficient than processing one object at a time by the incremental approaches. It should be noted that the computational time of algorithm KGIRD-M and algorithm KGIRD-S looks close for the data sets LSVT voice, Colon, and AMLALL. This is because the computational time of algorithm KGIRD-M is close to that of algorithm KGIRD-S relative to the computational time of algorithm THD-M. In fact, the computational time of algorithm KGIRD-S is several times that of algorithm KGIRD-M on the same data set. For example, on the data set AMLALL, the computational time of algorithm KGIRD-S is nearly four times that of algorithm KGIRD-M. It is clearly that the proposed algorithm KGIRD-M has the highest efficiency among the three comparative algorithms.

4.3 Performance comparison on reduct size

In general, the number of attributes in a reduct, also called the size of the reduct(SR), is an important reference index to evaluate the effects of attribute reduction algorithms. The efficiency of data processing increases with SR decreasing, because the number of attributes participating in the calculation decreases with SR decreasing. In Table 4, the reduct for each data set is obtained from the data set with all of the objects. We can find that reducts obtained by algorithm KGIRA-M or KGIRA-S are not the same as those obtained by algorithm THA-M. Moreover, SR obtained by algorithm KGIRA-M or KGIRA-S is a little more than that obtained by algorithm THA-M on most of data sets. It is because that algorithms KGIRA-M and KGIRA-S get reducts through the previous results. However, SR obtained by algorithm KGIRA-M is less than that obtained by algorithm KGIRA-S on most of data sets, because algorithm KGIRA-M get reducts based on the previous results in a way that multiple objects are processed at a time rather than one by one.

In Table 5, the reduct for each data set is obtained from the data set with remained objects. We can find that reducts obtained by algorithms KGIRD-M and KGIRD-S are obviously different from those obtained by algorithm THD-M. Furthermore, SR obtained by algorithm KGIRA-M or KGIRA-S is a little more than that obtained by algorithm THD-M. It is because, on the one hand, algorithms KGIRD-M and KGIRD-S get the reducts based on the previous results while algorithm THD-M computes the results from the beginning, on the other hand, results of attribute reduction for decision systems are not unique. However, reducts obtained by algorithm KGIRD-M are the same as those obtained by algorithm KGIRD-S on most data sets.

4.4 Performance comparison on classification accuracy

In Table 6, the classification accuracy is calculated by Naive Bayes Classifier (NB) and Decision Tree algorithm (REPTree) with fivefold cross-validation from the reducts of algorithms KGIRA-M, KGIRA-S and THA-M. From Table 6, it

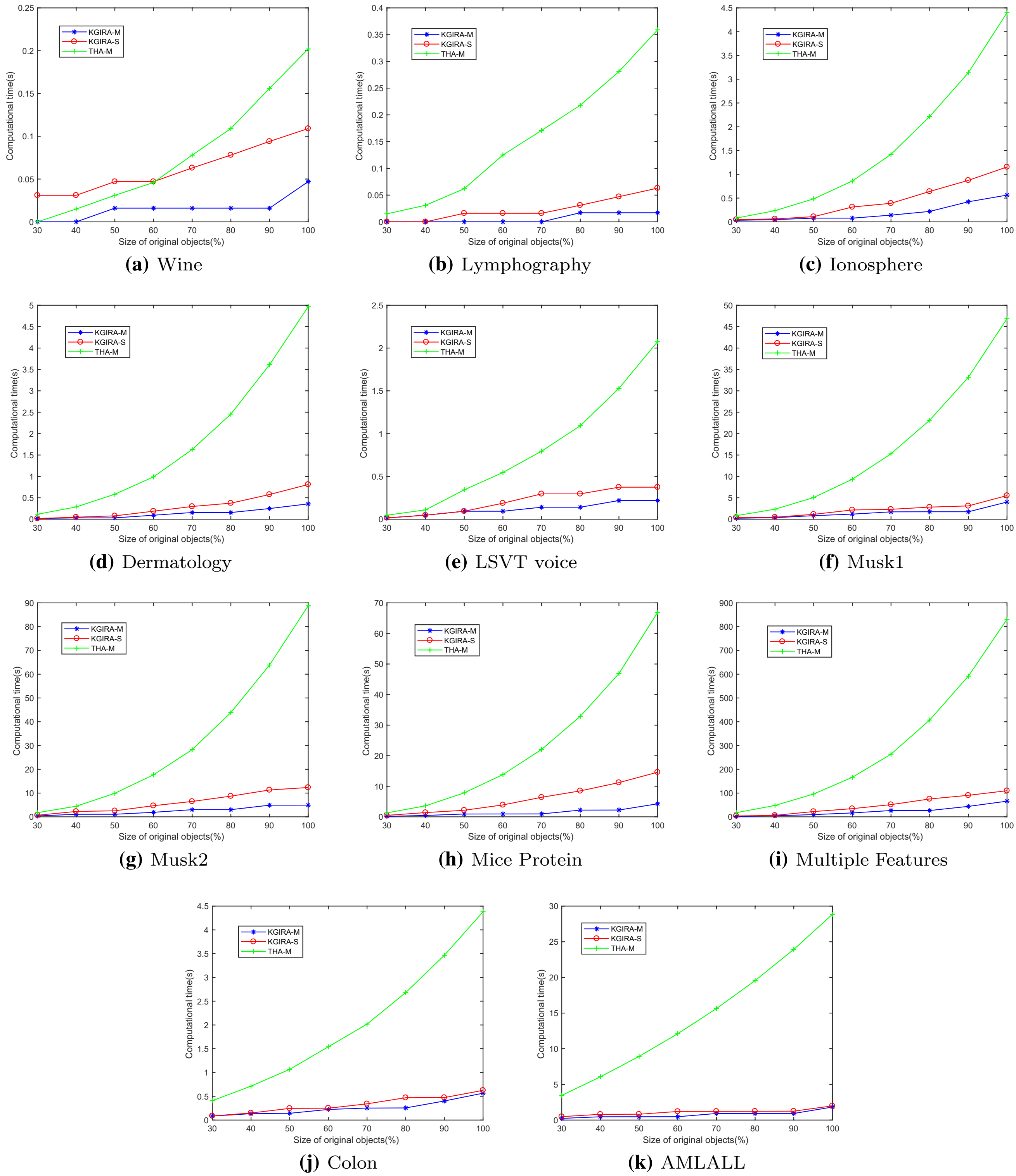

Fig. 1 Comparisons for computational time of different algorithms for adding objects

is clear that the average classification accuracy of the reducts found by incremental algorithms is better than that by non-incremental algorithm. Furthermore, the average classification accuracy of the reducts found by algorithm KGIRA-M is higher than that by algorithm KGIRA-S. The experimental results show that the proposed algorithm KGIRA-M can discover a better attribute reduct compared with algorithm KGIRA-S and THA-M from the viewpoint of classification

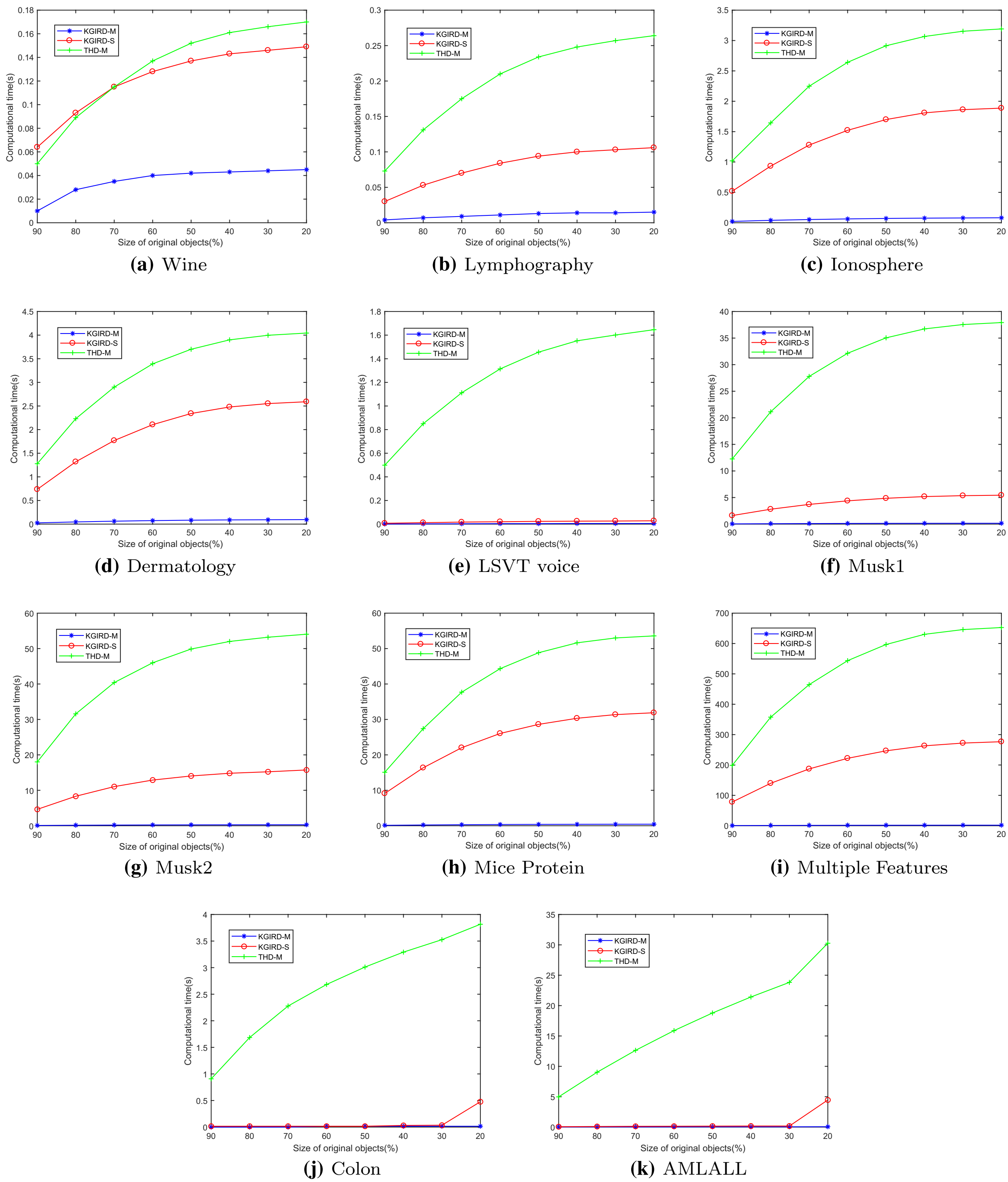

Fig. 2 Comparisons for computational time of different algorithms for deleting objects

performance. Moreover, algorithm KGIRA-M can discover a feasible attribute reduct with the shortest computational time compared with algorithms KGIRA-S and THA-M.

In Table 7, the classification performance is calculated on the reducts obtained by algorithms KGIRD-M, KGIRD-S and THD-M. The results of classification accuracy are

Table 4 Comparisons for reducts of different algorithms for adding objects

	Data sets	KGIRA-M		KGIRA-S		THA-M	
		SR	Reduct	SR	Reduct	SR	Reduct
1	Wine	5	9, 5, 0, 1, 3	5	1, 3, 0, 2, 6	5	10, 5, 0, 4, 1
2	Lymphography	7	17, 13, 12, 14, 0, 1, 10	8	17, 13, 12, 14, 0, 1, 7, 5	7	13, 17, 12, 1, 14, 15, 0
3	Ionosphere	9	15, 21, 27, 23, 4, 9, 17, 12, 28	10	21, 18, 6, 7, 16, 19, 5, 2,11, 13	9	32, 23, 4, 33, 18, 19, 6, 17, 7
4	Dermatology	11	15, 3, 18, 2, 16, 0, 1, 5, 4, 8, 17	13	15, 3, 18, 2, 14, 4, 6,1, 9, 5, 12, 17, 8	10	8, 3, 15, 18, 2, 31, 27, 1, 16, 4
5	LSVT voice	6	82, 4, 34, 41, 23, 8	7	82, 3, 1, 13, 6, 34, 8	3	86, 300, 51
6	Musk1	13	1, 13, 130, 53, 105, 3, 0, 107, 115, 5, 22, 69, 16	14	1, 13, 130, 3, 5, 7, 10, 2, 14, 0, 12,8, 22, 16	11	91, 5, 87, 107, 13, 100, 93, 150, 14,157, 1
7	Musk2	12	118, 101, 15, 124, 19, 5, 7, 26, 10, 0, 3, 35	11	101, 15, 19, 2, 10, 11, 0, 1, 3, 6, 35	9	125, 78, 91, 42, 132, 146, 7, 3, 1
8	Mice Protein	8	46, 76, 30, 48, 2, 7, 4, 0	8	46, 76, 30, 48, 2, 1, 3, 4	7	46, 76, 49, 21, 58, 64, 1
9	Multiple Features	9	92, 151, 179, 0, 4, 9, 5, 6, 1	10	92, 151, 12, 179, 2, 4, 1, 5, 0, 3	8	54, 9, 56, 151, 10, 71, 130, 4
10	Colon	5	942, 10, 0, 1, 15	6	942, 0, 4, 2, 3, 5	3	1581, 1923, 69
11	AMLALL	4	2140, 2, 1, 3	4	2140, 2, 1, 3	2	4846, 1238

Table 5 Comparisons for reducts of different algorithms for deleting objects

	Data sets	KGIRD-M		KGIRD-S		THD-M	
		SR	Reduct	SR	Reduct	SR	Reduct
1	Wine	4	10, 5, 0, 4	3	10, 5, 0	2	0, 6
2	Lymphography	4	13, 17, 12, 15	4	13, 17, 12, 1	3	13, 17, 0
3	Ionosphere	5	33, 18, 19, 6, 7	5	33, 18, 19, 6, 7	4	20, 7, 32, 5
4	Dermatology	6	3, 15, 2, 31, 27, 4	6	3, 15, 2, 31, 27, 4	6	15, 3, 27, 2, 4, 1
5	LSVT voice	3	86, 300, 51	3	86, 300, 51	2	86, 74
6	Musk1	8	5, 87, 107, 13, 93, 150, 14, 1	7	91, 107, 13, 93, 150, 14, 1	6	41, 96, 48, 87, 88, 0
7	Musk2	5	125, 91, 146, 7, 3	5	125, 91, 146, 7, 3	4	125, 128, 67, 0
8	Mice Protein	6	46, 76, 49, 21, 58, 64	6	46, 76, 49, 21, 58, 64	5	46, 76, 58, 3, 4
9	Multiple Features	7	54, 9, 56, 10, 71, 130, 4	7	54, 9, 56, 10, 71, 130, 4	5	54, 9, 106, 151, 75
10	Colon	3	1581, 1923, 69	3	1581, 1923, 69	1	627
11	AMLALL	2	4846, 1238	2	4846, 1238	1	256

calculated by Naive Bayes Classifier (NB) and Decision Tree algorithm(REPTree) with fivefold cross-validation. From Table 7, it is clear that the average classification accuracy of the reducts found by incremental algorithm KGIRD-M is slightly higher than that by algorithms KGIRD-S and THD-M. The experimental results show that the incremental algorithm KGIRD-M can discover a satisfactory attribute reduct from the viewpoint of classification performance. Moreover, algorithm KGIRD-M can discover a feasible attribute reduct with the shortest computational time compared with algorithms KGIRD-S and THD-M.

5 Conclusion

In this paper, we exploit deeper insights into incremental mechanisms for the attribute reduction process from granulation perspective. By introducing the granulation measure into incomplete decision systems, we propose two incremental attribute reduction approaches based on knowledge

Table 6 Comparison on classification accuracy of different algorithms for adding objects

	Data sets	NB			REPTree		
		KGIRA-M	KGIRA-S	THA-M	KGIRA-M	KGIRA-S	THA-M
1	Wine	91.35 ± 0.51	91.91 ± 0.63	89.44 ± 0.83	76.35 ± 2.11	84.21 ± 1.68	74.61 ± 2.94
2	Lymphography	83.65 ± 0.51	83.04 ± 0.47	80.68 ± 0.69	72.09 ± 1.86	71.22 ± 2.48	71.69 ± 2.52
3	Ionosphere	84.36 ± 0.84	86.47 ± 0.34	84.33 ± 0.36	86.38 ± 1.23	85.93 ± 1.91	88.09 ± 1.19
4	Dermatology	92.67 ± 0.43	93.44 ± 0.17	89.57 ± 0.61	73.91 ± 1.01	80.16 ± 2.06	63.36 ± 2.40
5	LSVT voice	72.22 ± 1.84	72.94 ± 2.31	75.32 ± 2.11	68.10 ± 1.80	71.67 ± 2.04	70.56 ± 1.64
6	Musk1	72.58 ± 0.70	69.50 ± 0.67	72.16 ± 0.82	69.08 ± 2.45	69.20 ± 1.95	71.28 ± 1.37
7	Musk2	83.27 ± 0.44	82.02 ± 0.68	83.86 ± 0.64	86.97 ± 0.37	86.31 ± 0.87	86.18 ± 0.61
8	Mice Protein	73.45 ± 0.44	66.08 ± 0.31	64.17 ± 0.58	66.15 ± 0.51	64.83 ± 1.17	62.94 ± 1.33
9	Multiple Features	83.53 ± 0.12	83.45 ± 0.21	78.84 ± 0.21	65.81 ± 0.65	64.66 ± 0.50	58.28 ± 0.82
10	Colon	94.73 ± 3.40	94.04 ± 2.32	94.06 ± 1.27	93.72 ± 2.34	90.83 ± 2.67	94.44 ± 0.00
11	AMLALL	70.65 ± 3.66	71.38 ± 3.13	70.00 ± 3.00	66.94 ± 3.77	63.87 ± 1.94	75.32 ± 2.70
	Average	81.94	81.08	80.17	76.89	76.49	75.86

Table 7 Comparison on classification accuracy of different algorithms for deleting objects

	Data sets	NB			REPTree		
		KGIRD-M	KGIRD-S	THD-M	KGIRD-M	KGIRD-S	THD-M
1	Wine	80.86 ± 3.39	78.57 ± 2.93	86.57 ± 1.83	85.43 ± 4.64	81.57 ± 8.83	86.57 ± 3.14
2	Lymphography	67.24 ± 3.86	78.97 ± 3.92	63.79 ± 4.43	62.62 ± 5.22	67.59 ± 3.16	54.14 ± 3.79
3	Ionosphere	89.71 ± 1.25	84.03 ± 1.62	83.57 ± 2.05	85.29 ± 3.38	85.00 ± 1.60	76.29 ± 2.32
4	Dermatology	70.00 ± 1.98	68.90 ± 2.60	70.68 ± 1.64	60.41 ± 1.67	58.49 ± 3.57	56.99 ± 1.75
5	LSVT voice	54.80 ± 7.81	51.20 ± 8.73	44.00 ± 6.69	64.00 ± 0.00	64.00 ± 0.00	64.00 ± 0.00
6	Musk1	69.21 ± 2.31	68.11 ± 2.49	57.89 ± 2.62	61.67 ± 4.90	62.41 ± 3.94	60.32 ± 2.58
7	Musk2	84.75 ± 1.39	84.11 ± 2.29	85.39 ± 1.11	81.21 ± 0.73	81.21 ± 0.73	81.21 ± 1.11
8	Mice Protein	50.37 ± 1.97	49.95 ± 1.47	45.69 ± 1.49	44.53 ± 2.56	45.23 ± 2.10	42.82 ± 2.08
9	Multiple Features	66.60 ± 1.11	66.65 ± 0.80	62.45 ± 1.04	43.45 ± 1.06	41.23 ± 1.30	45.48 ± 1.36
10	Colon	96.10 ± 3.67	95.27 ± 3.51	100.00 ± 0.00	96.83 ± 3.34	96.17 ± 3.72	100.00 ± 0.00
11	AMLALL	95.71 ± 3.50	95.00 ± 4.57	100.00 ± 0.00	95.29 ± 4.80	95.71 ± 4.18	100.00 ± 0.00
	Average	75.03	74.61	72.73	70.98	70.78	69.80

granularity under the situations of adding multiple objects and deleting multiple objects at a time, respectively. Compared with existing methods, our incremental attribute reduction approaches can deal with variation of object sets in incomplete decision systems effectively.

Attribute reduction for other variations of incomplete decision systems is not concerned, and we plan to study knowledge granularity based incremental attribute reduction solutions for incomplete decision systems under the situation of variation of attribute sets or attribute values.

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (Nos. 61976089, 61473259, 61070074, and 60703038), and the Hunan Provincial Science and Technology Project Foundation (2018TP1018, 2018RS3065).

References

1. Pawlak Z (1991) Rough sets: theoretical aspect of reasoning about data. Kluwer Academic Publishers, Dordrecht
2. Wang R, Wang X, Kwong S, Xu C (2017) Incorporating diversity and informativeness in multiple-instance active learning. IEEE Trans Fuzzy Syst 25(6):1460–1475
3. Wang X, Tsang ECC, Zhao S, Chen D, Yeung DS (2007) Learning fuzzy rules from fuzzy samples based on rough set technique. Inf Sci 177(20):4493–4514
4. Wang X, Xing H, Li Y, Hua Q, Dong C, Pedrycz W (2015) A study on relationship between generalization abilities and fuzziness of base classifiers in ensemble learning. IEEE Trans Fuzzy Syst 23(5):1638–1654
5. Dai J, Tian H, Wang W, Liu L (2013) Decision rule mining using classification consistency rate. Knowl-Based Syst 43:95–102
6. Zhao B, Ren Y, Gao D (2019) Prediction of service life of large centrifugal compressor remanufactured impeller based on

clustering rough set and fuzzy bandelet neural network. Appl Soft Comput 78:132–140
7. Hao C, Li J, Fan M, Liu W, Tsang ECC (2017) Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions. Inf Sci 415:213–232
8. Liang D, Xu Z, Liu D (2017) Three-way decisions based on decision-theoretic rough sets with dual hesitant fuzzy information. Inf Sci 396:127–143
9. Wang X, Zhai J, Lu S (2008) Induction of multiple fuzzy decision trees based on rough set technique. Inf Sci 178(16):3188–3202
10. Liu X, Qian Y, Liang J (2014) A rule-extraction framework under multigranulation rough sets. Int J Mach Learn Cybern 5(2):319–326
11. Zhang X, Mei C, Chen D, Li J (2014) Multi-confidence rule acquisition and confidence-preserved attribute reduction in interval-valued decision systems. Int J Approx Reason 55(8):1787–1804
12. Cheruku R, Edla DR, Kuppili V, Dharavath R (2018) RST-BatMiner: a fuzzy rule miner integrating rough set feature selection and bat optimization for detection of diabetes disease. Appl Soft Comput 67:764–780
13. Hamouda SKM, Wahed ME, Alez RHA, Riad K (2018) Robust breast cancer prediction system based on rough set theory at National Cancer Institute of Egypt. Comput Methods Programs Biomed 153:259–268
14. Jothi G, Inbarani HH (2016) Hybrid tolerance rough set-firefly based supervised feature selection for MRI brain tumor image classification. Appl Soft Comput 46:639–651
15. Dai J, Xu Q (2013) Attribute selection based on information gain ratio in fuzzy rough set theory with application to tumor classification. Appl Soft Comput 13(1):211–221
16. Sun L, Zhang X, Qian Y, Xu J, Zhang S, Tian Y (2019) Joint neighborhood entropy-based gene selection method with fisher score for tumor classification. Appl Intell 49(4):1245–1259
17. Dai J, Wang W, Xu Q, Tian H (2012) Uncertainty measurement for interval-valued decision systems based on extended conditional entropy. Knowl Based Syst 27:443–450
18. Dai J, Wang W, Mi J (2013) Uncertainty measurement for interval-valued information systems. Inf Sci 251:63–78
19. Dai J, Hu H, Hu Q, Huang W, Zheng N, Liu L (2018) Locally linear approximation approach for incomplete data. IEEE Trans Cybern 48(6):1720–1732
20. Wang C, Huang Y, Shao M, Chen D (2019) Uncertainty measures for general fuzzy relations. Fuzzy Sets Syst 360:82–96
21. Dai J, Wei B, Zhang X, Zhang Q (2017b) Uncertainty measurement for incomplete interval-valued information systems based on α-weak similarity. Knowl Based Syst 136:159–171
22. Ko YC, Fujita H, Li T (2017) An evidential analysis of Altman Z-score for financial predictions: case study on solar energy companies. Appl Soft Comput 52:748–759
23. Lei L (2018) Wavelet neural network prediction method of stock price trend based on rough set attribute reduction. Appl Soft Comput 62:923–932
24. Singh AK, Baranwal N, Nandi GC (2019) A rough set based reasoning approach for criminal identification. Int J Mach Learn Cybern 10(3):413–431
25. Wang X, Wang R, Xu C (2018b) Discovering the relationship between generalization and uncertainty by incorporating complexity of classification. IEEE Trans Cybern 48(2):703–715
26. Fan J, Jiang Y, Liu Y (2017) Quick attribute reduction with generalized indiscernibility models. Inf Sci 397:15–36
27. Ni P, Zhao S, Wang X, Chen H, Li C (2019) PARA: a positive-region based attribute reduction accelerator. Inf Sci 503:533–550
28. Dai J, Tian H (2013) Fuzzy rough set model for set-valued data. Fuzzy Sets Syst 229:54–68
29. Dai J (2013) Rough set approach to incomplete numerical data. Inf Sci 240:43–57
30. Konecny J, Krajca P (2018) On attribute reduction in concept lattices: experimental evaluation shows discernibility matrix based methods inefficient. Inf Sci 467:431–445
31. Wang C, He Q, Shao M, Hu Q (2018a) Feature selection based on maximal neighborhood discernibility. Int J Mach Learn Cybern 9(11):1929–1940
32. Dai J, Hu Q, Zhang J, Hu H, Zheng N (2017) Attribute selection for partially labeled categorical data by rough set approach. IEEE Trans Cybern 47(9):2460–2471
33. Dai J, Tian H (2013) Entropy measures and granularity measures for set-valued information systems. Inf Sci 240(11):72–82
34. Wang F, Liang J, Dang C (2013) Attribute reduction for dynamic data sets. Appl Soft Comput 13(1):676–689
35. Bhattacharya A, Goswami RT, Mukherjee K (2019) A feature selection technique based on rough set and improvised PSO algorithm (PSORS-FS) for permission based detection of android malwares. Int J Mach Learn Cybern 10(7):1893–1907
36. Dai J, Hu Q, Hu H, Huang D (2018a) Neighbor inconsistent pair selection for attribute reduction by rough set approach. IEEE Trans Fuzzy Syst 26:937–950
37. Dai J, Hu H, Wu WZ, Qian Y, Huang D (2018) Maximal discernibility pairs based approach to attribute reduction in fuzzy rough sets. IEEE Trans Fuzzy Syst 26(4):2174–2187
38. Li F, Jin C, Yang J (2019) Roughness measure based on description ability for attribute reduction in information system. Int J Mach Learn Cybern 10(5):925–934
39. Liu K, Yang X, Yu H, Mi J, Wang P, Chen X (2019) Rough set based semi-supervised feature selection via ensemble selector. Knowl Based Syst 165:282–296
40. Wang C, Huang Y, Shao M, Fan X (2019) Fuzzy rough set-based attribute reduction using distance measures. Knowl Based Syst 164:205–212
41. Dai J, Han H, Hu Q, Liu M (2016) Discrete particle swarm optimization approach for cost sensitive attribute reduction. Knowl Based Syst 102:116–126
42. Li Y, Jin Y, Sun X (2018) Incremental method of updating approximations in DRSA under variations of multiple objects. Int J Mach Learn Cybern 9(2):295–308
43. Liang J, Wang F, Dang C, Qian Y (2014) A group incremental approach to feature selection applying rough set technique. IEEE Trans Knowl Data Eng 26(2):294–308
44. Ma F, Ding M, Zhang T (2019) Compressed binary discernibility matrix based incremental attribute reduction algorithm for group dynamic data. Neurocomputing 294:20–27
45. Wei W, Song P, Liang J, Wu X (2019) Accelerating incremental attribute reduction algorithm by compacting a decision table. Int J Mach Learn Cybern 10(9):2355–2373
46. Yang Y, Chen D, Wang H (2017) Active sample selection based incremental algorithm for attribute reduction with rough sets. IEEE Trans Fuzzy Syst 25(4):825–838
47. Yang Y, Chen D, Wang H (2018) Incremental perspective for feature selection based on fuzzy rough sets. IEEE Trans Fuzzy Syst 26(3):1257–1273
48. Jing Y, Li T, Huang J, Zhang Y (2016a) An incremental attribute reduction approach based on knowledge granularity under the attribute generalization. Int J Approx Reason 76:80–95
49. Jing Y, Li T, Huang J, Chen H, Horng SJ (2017) A group incremental reduction algorithm with varying data values. Int J Intell Syst 32(9):900–925
50. Wei W, Wu X, Liang J, Cui J, Sun Y (2018) Discernibility matrix based incremental attribute reduction for dynamic data. Knowl Based Syst 140:142–157

51. Jing Y, Li T, Fujita H, Wang B, Cheng N (2018) An incremental attribute reduction method for dynamic data mining. Inf Sci 465:202–218
52. Yang C, Ge H, Li L, Ding J (2019) A unified incremental reduction with the variations of the object for decision tables. Soft Comput 23(15):6407–6427
53. Shu W, Shen H (2014a) Incremental feature selection based on rough set in dynamic incomplete data. Pattern Recogn 47(12):3890–3906
54. Shu W, Shen H (2014b) Updating attribute reduction in incomplete decision systems with the variation of attribute set. Int J Approx Reason 55(3):867–884
55. Xie X, Qin X (2018) A novel incremental attribute reduction approach for dynamic incomplete decision systems. Int J Approx Reason 93:443–462
56. Luo C, Li T, Yao Y (2017) Dynamic probabilistic rough sets with incomplete data. Inf Sci 417:39–54
57. Zhang C, Dai J (2019) An incremental attribute reduction approach based on knowledge granularity for incomplete decision systems. Granul Comput. https://doi.org/10.1007/s41066-019-00173-7
58. Kryszkiewicz M (1998) Rough set approach to incomplete information systems. Inf Sci 112(1):39–49
59. Jing Y, Li T, Luo C, Horng SJ, Wang G, Yu Z (2016b) An incremental approach for attribute reduction based on knowledge granularity. Knowl Based Syst 104:24–38