International Journal of Machine Learning and Cybernetics (2019) 10:2355–2373
https://doi.org/10.1007/s13042-018-0874-x

ORIGINAL ARTICLE

Accelerating incremental attribute reduction algorithm by compacting a decision table

Wei Wei[1] · Peng Song[1] · Jiye Liang[1] · Xiaoying Wu[1]

Received: 25 July 2018 / Accepted: 6 September 2018 / Published online: 19 September 2018

Abstract
The evolution of object sets over time is ubiquitous in dynamic data. To acquire reducts for this type of data, researchers have proposed many incremental attribute reduction algorithms based on discernibility matrices. Although all reducts of an updated decision table can be obtained using these algorithms, their high computation time is a critical issue. To address this issue, we first construct three new types of discernibility matrices by compacting a decision table to eliminate redundant entries in the discernibility matrices of the original decision table. We then demonstrate that the set of reducts obtained from the compacted decision table are the same as those acquired from the original decision table. Extensive experiments have demonstrated that an incremental attribute reduction algorithm based on a compacted decision table can significantly accelerate attribute reduction for dynamic data with changing object sets while the acquired reducts are identical to those obtained using existing algorithms.

Keywords Rough set · Incremental attribute reduction · Discerptiblity matrix · Compacted decision table

1 Introduction

The continuing development of data capture and storage technologies has resulted in rapidly incoming and extremely dynamic data in many real-life applications, such as transaction data and web data. Traditional non-incremental algorithms address this issue by considering updated data as completely new data without using any incremental strategy. Such approaches neglect information shared by the data before and after updating, and therefore, they are normally inefficient. As a popular mathematical tool for data analysis [19, 28, 46], rough set has been successfully applied in incremental methods for data analysis. Many rough set based incremental learning approaches have been developed to address this problem. In these approaches, intermediate results based on the original data can be used to accelerate the process of performing attribute reduction on the updated data [1, 18, 22, 27]. Attribute reduction, which is a representative feature selection method, attempts to select an attribute set that can preserve some discrimination ability in the original data [2, 5, 14, 16, 39, 41, 52]. Consequently, incremental attribute reduction has become a hot topic in rough set theory.

In the last two decades, to meet the demand of processing dynamic data, researchers have introduced several incremental algorithms for attribute reduction. We review some related work on incremental updating approaches based on rough set theory, which are categorized as follows [11, 34, 48]:

1. The object set of a data set evolves over time. Liu [26] proposed incremental arithmetic to determine the minimal reduct for an information system. To update the attribute reduct for a decision table using an incremental technique, Wang proposed an incremental attribute reduction algorithm based on Skowron's discernibility matrix in Wang and Wang [40]; however, the algorithm

✉ Peng Song
songpeng@sxu.edu.cn

Wei Wei
weiwei@sxu.edu.cn

Jiye Liang
ljy@sxu.edu.cn

Xiaoying Wu
1652851969@qq.com

[1] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China

is incapable of updating reducts for an inconsistent decision table. Some representative incremental algorithms were proposed for conducting attribute reduction based on modified discernibility matrices [49], positive region [8], 0–1 integer programming [48], and information entropies [22]. It should be noted that in Liang et al. [22], Liang et al. systemically investigated the properties of entropies for a group of objects added into a decision table and proposed an incremental attribute reduction algorithm that is much more efficient than existing ones. For a dynamic incomplete decision table with its object set changing, Shu et al. [35] proposed an incremental approach of obtaining reducts by updating the positive region, and Liu et al. [25] introduced an incremental attribute reduction approach by constructing three novel matrices: a support matrix, an accuracy matrix, and a coverage matrix. By means of computing the updated knowledge granularity, Jin et al. [9] presented the corresponding incremental algorithms for attribute reduction when some objects evolve dynamically.

2. The condition attribute set of a data set evolves over time. A number of important knowledge acquisition methods, which can effectively learn knowledge in dynamic data sets with a varying attribute set, have been presented in [6, 18, 23, 24, 32, 53]. Inspired by such literature, Wang et al. [38], for the first time, introduced an incremental attribute reduction algorithm for a dynamic data set in which new attributes are progressively added; however, the algorithm is only suitable for a complete data set. Shu et al. [34] proposed a positive region-based attribute reduction algorithm to extend the application of the incremental approach to an incomplete decision table with a varying attribute set. Li et al. [17] studied the change mechanism of P-dominating sets and P-dominated sets when some attributes are added into or deleted from a dominance-based decision table and proposed the relevant incremental approaches. By calculating the updated knowledge granularity, Jin et al. [10] proposed an incremental algorithms for attribute reduction when multiple attributes are added to a decision system. For dynamic covering decision information systems with variations of attribute sets, Lang el al. [12, 13] presented several incremental algorithms for attribute reduction.
3. The attribute values of objects evolve over time. Chen et al. [3, 4] introduced two types of incremental algorithms to update rough approximations of a target concept in the context of varying attribute values. Unfortunately, the means of implementing attribute reduction for this kind of dynamic data set was not discussed. To solve this problem, Wang et al. [37] presented three incremental attribute reduction algorithms based on three representative entropies: complementary entropy [20, 21], combination entropy [31], and Shannon's entropy [33] that can obtain an updated reduct in much lesser time. Shu et al. [35] proposed a method of modifying the positive region for an incomplete dynamic data set with attribute values evolving over time and designed two incremental algorithms to compute the updated reducts. To enhance the efficiency of the type of algorithms, Xie and Qin [45] introduced three update strategies of tolerance classes and the update mechanism of the inconsistency degree, presented a framework of the incremental attribute reduction algorithms for incomplete decision systems with variation of attribute values.

In addition, some researches on updating rough approximations in the context of multi-granulation rough set has been launched. Yang et al. [50] proposed a method of updating the multi-granulation rough approximations. Ju et al. [51] further developed the updating of the multi-granulation fuzzy rough approximations. Hu et al. [7] proposed a matrix representation of multi-granulation approximations in multi-granulation rough set, and then corresponding matrix-based algorithms for updating rough approximations. Note that, by means of the methods of updating rough approximations, it is easy to design corresponding incremental attribute reduction algorithms for dynamic data.

Among the algorithms mentioned above, discernibility matrix based incremental attribute reduction [40, 49] is one of the important algorithms in which updated reducts can be obtained by modifying a decision table's discernibility matrix based on the relationships among the newly added object and all other objects in the decision table. This type of algorithm can obtain all reducts of an updated decision table, which is its distinct advantage in comparison with other incremental algorithms. However, it should be noted that efficiency is still a critical issue for incremental attribute reduction algorithms based on discernibility matrices. Inspired by the study on compacted decision tables in Wei et al. [43], we investigate the mechanism of discernibility matrices of a compacted decision table after adding a new object into the compacted decision table to address the issue associated with efficiency. We demonstrated that the reducts derived from an updated compacted decision table are identical with those derived from its corresponding original decision table. It should be noted that incremental attribute reduction algorithms based on the discernibility matrix of a compacted decision table are more efficient than those based on the discernibility matrix of its original version because the scale of the discernibility matrix of a compacted decision table is much smaller than that of the original decision table.

The remainder of this paper is organized as follows. Section 2 reviews some preliminaries on rough sets, discernibility matrices, and incremental algorithms. In Sect. 3, we introduce the discernibility matrices of a compacted decision table and present the relationship between the discernibility matrices of a compacted decision table and those of the original decision table. In Sect. 4, we investigate the relationship between reducts obtained from an updated compacted decision table and those obtained from an updated decision table and design an incremental attribute reduction algorithm based on a compacted decision table. In Sect. 5, extensive experiments and their results are described to indicate the effectiveness and efficiency of the proposed incremental algorithms. Section 6 concludes the paper with some remarks.

2 Preliminaries

2.1 Rough set

In rough set theory, as a knowledge representation system, an information system [28–30], is a 4-tuple $S = (U, A, V, f)$, where U is a non-empty and finite set of objects, called a universe of discourse, and A is a non-empty and finite set of attributes, V_a is the domain of the attribute a, $V = \bigcup_{a \in A} V_a$ and $f_S : U \times A \rightarrow V$ is a function, $f_S(x, a) \in V_a$ $(a \in A)$. Thereafter, $S = (U, A)$ is used to represent $S = (U, A, V, f)$ for short.

For any $Y \subseteq U$, $(\overline{B}(Y), \underline{B}(Y))$ is defined as the rough set of Y with respect to B, where the lower approximation $\underline{B}(Y)$ and the upper approximation $\overline{B}(Y)$ of Y are $\underline{B}(Y) = \{x \mid [x]_B \subseteq Y\}$ and $\overline{B}(Y) = \{x \mid [x]_B \cap Y \neq \emptyset\}$, respectively. The B−boundary region of Y is defined as the set $BN_B(Y) = \overline{B}(Y) - \underline{B}(Y)$. The objects in lower approximation of Y are certainly classified as the members of Y based on information in B, while the objects in upper approximation of Y are possibly classified as the members of Y based on information in B.

When we encounter a decision problem, a decision table $DT = (U, C \cup \{d\}, V, f)$ is usually employed to represent the problem, where C is a condition attribute set and $\{d\}$ is a decision attribute, $V_d = \{v_{d_1}, v_{d_2}, ..., v_{d_l}\}$ is the domain of the attribute d. Let $B \subseteq C$, $U/\{d\} = \{Y_1, Y_2, \ldots, Y_n\}$, we define the lower and upper approximations of the decision attribute $\{d\}$ as $\underline{B}\{d\} = \{\underline{B}Y_1, \underline{B}Y_2, \ldots, \underline{B}Y_n\}$, *and* $\overline{B}\{d\} = \{\overline{B}Y_1, \overline{B}Y_2, \ldots, \overline{B}Y_n\}$. Then the positive region of $\{d\}$ with respect to B can be defined as $POS_B(\{d\}) = \bigcup_{i=1}^{n} \underline{B}Y_i$.

Note that in a decision table, the rows of the objects in the positive region of $\{d\}$ with respect to C is regarded as the consistent part of the decision table, and the other rows is its inconsistent part. If $POS_C(\{d\})$ equals U in a decision table, then we call the decision table is a consistent decision table.

2.2 Discernibility matrix

Discernibility matrix is first introduced by Skowron and Rauszer for an information system [36], by using which, it is easy to obtain reducts and make the rule acquisition easier and the extracted approximate decision rules more compact [15]. To extend the discernibility matrix to a decision table, a decision table oriented discernibility matrices was presented in Yang [49] as follows.

Definition 1 [49] Let $DT = (U, C \cup \{d\})$ be a decision table, C the condition attribute set, and d the decision attribute. The discernibility matrix in the context of the positive region is defined as $\mathbf{M}^P_{DT} = \{m^P_{ij}\}$, where

$$m^P_{ij} = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d) \text{ and } x_i, x_j \in U_1; \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2; \\ \emptyset,\ otherwise; \end{cases} \tag{1}$$

U_1 is the consistent part of the decision table DT and U_2 is the inconsistent part of the decision table DT.

From this definition, we can determine that if two objects x_i and x_j belong to U_1 of DT and their values of the decision attribute are different, and x_i is in U_1 while x_j is in U_2, then an attribute a_i in which these two objects possess different values must be in the corresponding entry (m^P_{ij}) of the discernibility matrix in the context of positive region $(\mathbf{M}^P_{DT})$. In other words, a_i can distinguish x_i and x_j. Therefore, if a_i is deleted from C, then the two objects will fall into one equivalence class if there are no other attributes that can distinguish between the two objects. Subsequently, the positive region derived from $C - \{a_i\}$ $(POS_{C-\{a_i\}}(D))$ is larger than $POS_C(D)$, which suggests that the information required for distinguishing between the two objects is preserved in a discernibility matrix. Therefore, it is possible to obtain all reducts of a decision table through its discernibility matrix. To obtain these reducts, the corresponding discernibility function in the context of positive region was defined as $\mathcal{F}(\mathbf{M}^P_{DT}) = \bigwedge\{\bigvee(m^P_{ij}) \mid \forall x, y \in U, m^P_{ij} \neq \emptyset\}$.

However, the discernibility matrix can only be used to compute reducts in the context of the positive region. To obtain the reducts in the context of Shannon entropy and complement entropy, Wei et al. [42] proposed two new discernibility matrices as follows:

Definition 2 [42] Let $DT = (U, C \cup \{d\})$ be a decision table, C the condition attribute set, and d the decision attribute. The discernibility matrix in the sense of Shannon entropy is defined as $\mathbf{M}^S_{DT} = \{m^S_{ij}\}$, where

$$m_{ij}^S = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d) \text{ and } x_i, x_j \in U_1; \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2; \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ \exists Y_k \in U/\{d\}\ such\ that\ \mu_{ik} \neq \mu_{jk}, \text{and } x_i, x_j \in U_2; \\ \emptyset,\ otherwise; \end{cases} \tag{2}$$

where U_1 is the consistent part of the decision table DT and U_2 is the inconsistent part of the decision table DT, $\mu_{ik} = \frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|}$, $\mu_{jk} = \frac{|[x_j]_C \cap Y_k|}{|[x_j]_C|}$, $[x_i]_C \in U/C$ and $[x_j]_C \in U/C$.

From this definition, we can determine that if two objects x_i and x_j belong to the consistent part U_1 of DT and their values of decision attributes are different, or x_i is in U_1 and x_j is in U_2, or there exists a decision class Y_k such that $\frac{|[x_i]_C \cap Y_k|}{|[x_i]_C|}$ and $\frac{|[x_j]_C \cap Y_k|}{|[x_j]_C|}$ when both x_i and x_j are in U_2, then an attribute a_i in which these two objects possess different values must be in the corresponding entry (m_{ij}) of the discernibility matrix in the context of Shannon entropy ($\mathbf{M}_{DT}^S$). In other words, a_i can distinguish between x_i and x_j. Therefore, if a_i is deleted from C, then the two objects will fall into one equivalence class if there is no other attribute that can distinguish between the objects. The Shannon entropy derived from $C - \{a_i\}$ ($H(D|C - \{a_i\})$) is larger than $H(D|C)$, which suggests that the information used to distinguish between the two objects is preserved in a discernibility matrix, where $H(D|C) = -\sum_{i=1}^{m} \frac{|X_i|}{|U|} \sum_{j=1}^{n} \frac{|X_i \cap Y_j|}{|X_i|} \log \frac{|X_i \cap Y_j|}{|X_i|}$, $X_i \in U/C$ and $Y_j \in U/D$ is a nonempty set. Therefore, it is possible to obtain all reducts of a decision table through its discernibility matrix. To obtain these reducts, the corresponding discernibility function in the context of Shannon entropy was defined as $\mathcal{F}(\mathbf{M}_{DT}^S) = \bigwedge \left\{ \bigvee (m_{ij}^S) \mid \forall x, y \in U, m_{ij}^S \neq \emptyset \right\}$.

Definition 3 [42] Let $DT = (U, C \cup \{d\})$ be a decision table, C the condition attribute set, and d the decision attribute. The discernibility matrix in the context of complement entropy is defined as $\mathbf{M}_{DT}^C = \{m_{ij}^C\}$, where

$$m_{ij}^C = \begin{cases} \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ f(x_i, d) \neq f(x_j, d) \text{ and } x_i, x_j \in U_1; \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i \in U_1, x_j \in U_2; \\ \{c \in C : f(x_i, c) \neq f(x_j, c)\},\ x_i, x_j \in U_2; \\ \emptyset,\ otherwise; \end{cases} \tag{3}$$

where U_1 is the consistent part of the decision table DT and U_2 is the inconsistent part of the decision table DT.

From this definition, we can determine that if two objects x_i and x_j belong to U_1 of DT and their values of decision attributes are different, or x_i is in U_1 and x_j is in U_2 of DT, or both x_i and x_j are in U_2 of DT, then an attribute a_i in which these two objects possess different values must be in the corresponding entry (m_{ij}) of the discernibility matrix in the context of complement entropy ($\mathbf{M}_{DT}^C$). In other words, a_i can distinguish x_i and x_j. Therefore, if a_i is deleted from C, then the two objects will fall into one equivalence class if there are no other attributes that can distinguish between the two objects, and then the complement entropy derived from $C - \{a_i\}$ ($E(D|C - \{a_i\})$) is larger than $E(D|C)$, which suggests that the information required for distinguishing between each object is preserved in a discernibility matrix, where $E(D|C) = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|Y_j \cap X_i|}{|U|} \frac{|Y_j^c - X_i^c|}{|U|}$, Y_j^c and X_i^c are the complements of $Y_j \in U/D$ and $X_i \in U/C$, respectively. Therefore, it is possible to obtain all reducts of a decision table through its discernibility matrix. To obtain these reducts, the corresponding discernibility function in the context of complement entropy was defined as $\mathcal{F}(\mathbf{M}_{DT}^C) = \bigwedge \left\{ \bigvee (m_{ij}^C) \mid \forall x, y \in U, m_{ij}^C \neq \emptyset \right\}$.

2.3 Compacted decision table

To eliminate information redundancy in data sets aroused by these objects in an equivalence class having the same condition attribute values, Wei et al. [43] introduced the concept of a compacted decision table, which was defined as follows:

Definition 4 [43] Given a decision table $DT = (U, C \cup \{d\})$, $U = \{x_1, x_2, \ldots, x_n\}, U/C = \{X_1, X_2, \ldots, X_m\}, V_d = \{v_{d_1}, v_{d_2}, \ldots, v_{d_l}\}$, then a compacted decision table is defined as $CDT = (CU, C \cup CD)$, where $CU = \{cx_1, cx_2, \ldots, cx_m\}$, $f_{CDT}(cx_i, C) = f_{DT}(X_i, C)$ ($f_{CDT}(cx_i, c) = f_{DT}(X_i, c)$ for $\forall c \in C$), $CD = \{cd_1, cd_2, \ldots, cd_l\}$ and $f_{CDT}(cx_j, cd_i) = |\{x \mid f_{DT}(x, d) = v_{d_i}, x \in X_j\}|$.

For a given compacted decision table $CDT = (CU, C \cup CD)$, $CU_1 = \{cx_i \mid |\{cd_k \in CD | f_{CDT}(cx_i, cd_k) \neq 0\}| = 1\}$ is regarded as its consistent part, and $CU_2 = CU - CU_1$ as its inconsistent part.

To facilitate an investigation of discernibility matrices for compacted decision tables in the context of the positive region, Shannon entropy, and complement entropy, compacted decision tables are reviewed as follows.

Definition 5 [43] Given a decision table $DT = (U, C \cup D)$ and its compacted edition $CDT = (CU, C \cup CD)$, $B \subseteq C$, $CU/B = \{X_1', X_2', \ldots, X_l'\}$, then the positive region of B with respect to CD in the compacted decision table is defined as

$$POS_B^{CU}(CD) = \{cx \in X \mid X \in CU/B \wedge X \subseteq U'_{POS} \wedge |\{cd_i \in CD \mid f(X, cd_i) \neq 0\}| = 1\}, \quad (4)$$

where $U'_{POS} = \{cx_i | [cx_i]_C \in POS_C^U(D)\}$, and $f(X, d_i) = \sum_{cx_j \in X} f(cx_j, d_i)$.

And, the Shannon condition entropy of CD with respect to B in the compacted decision table is defined as

$$H^{CU}(CD|B) = -\sum_{j=1}^{m} \frac{\sum_{i=1}^{n} f(X'_j, d_i)}{|U'|} \sum_{i=1}^{n} \frac{f(X'_j, d_i)}{\sum_{i=1}^{n} f(X'_j, d_i)} log \frac{f(X'_j, d_i)}{\sum_{i=1}^{n} f(X'_j, d_i)}, \quad (5)$$

where $f(X'_j, d_i) = \sum_{x'_k \in X'_j} f(x'_k, d_i)$.

And, the complement entropy of CD with respect to B in a compacted decision table is defined as

$$E^{CU}(CD|B) = \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{f(X'_j, d_i)}{|U|} \frac{\sum_{i=1}^{n} f(X'_j, d_i) - f(X'_j, d_i)}{|U|}, \quad (6)$$

where $f(X'_j, d_i) = \sum_{x'_k \in X'_j} f(x'_k, d_i)$.

Based on the definition reviewed above, in the following section, three new types of discernibility matrices in the context of the positive region, Shannon entropy, and complement entropy will be proposed to obtain all reducts of a compacted decision table, respectively.

3 Discernibility matrix for a compacted decision table

The discernibility matrix is an important tool for computing all reducts of a decision table. In this section, we introduce three new types of discernibility matrices to compute all reducts of a compacted decision table and demonstrate that the reducts obtained through the discernibility matrices of a compacted decision table are the same as those based on the discernibility matrices of the original decision table. Based on these new types of discernibility matrices, we will investigate the relationship between the discernibility functions of a compacted decision table and those of the original decision table using the following Theorems 1–3. The relationship between the reducts acquired from a compacted decision table and those of the original decision table can then be easily revealed.

The discernibility matrices in the context of the positive region was first investigated. Similar to the discernibility matrix in the context of the positive region proposed in Definition 1, an attribute a_i in which two objects (cx_p and cx_q) of a compacted decision table possess different values must be placed in the corresponding entry of the decision table when these two objects satisfy the following cases: (1) both cx_p and cx_q are in the consistent part (CU_1) of a compacted decision table and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$; (2) cx_p is in the consistent part (CU_1) and cx_q is in the inconsistent part (CU_2) of the compacted decision table. The reason for distinguishing between the two objects in these two cases is that if the two objects (x_i and x_j) become one object in a new compacted decision table ($S' = (CU, C' \cup CD)$, $C' \subset C$), the positive region $POS_{C'}^{CU}(CD)$ will be smaller than the positive region $POS_C^{CU}(CD)$. However, the attribute a_i has the ability to distinguish between x_i and x_j (i.e. $POS_{C' \cup \{a_i\}}^{CU}(CD) = POS_C^{CU}(CD)$), and therefore it must be placed in the entry (cm_{ij}^P) of the discernibility matrix (M_{CDT}^P). Moreover, when both cx_p and cx_q do not conform to Case 1 and Case 2 mentioned above, they do not need be distinguished, and the entry corresponding to the two objects should be an empty set.

Based on the above discussion, it is easy to propose the following discernibility matrix in the context of the positive region.

Definition 6 Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$. A discernibility matrix in the context of the positive region is defined as $\mathbf{M}_{CDT}^P = \{cm_{pq}^P\}$, where

$$cm_{pq}^P = \begin{cases} \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\}, \\ \quad \{cd_k | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k | f_{CDT}(cx_q, cd_k) \neq 0\} \text{ and} \\ \quad cx_p, cx_q \in CU_1; \\ \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\}, \ cx_p \in CU_1, cx_q \in CU_2; \\ \emptyset, \ otherwise; \end{cases} \quad (7)$$

where CU_1 is the consistent part of CDT, and CU_2 is the inconsistent part of CDT.

To demonstrate that the same reducts as those obtained by the discernibility matrix in Definition 1 can be captured by our proposed discernibility matrix in the context of positive region (shown in Definition 6), we will need to investigate the relationship between the two discernibility functions based on these two discernibility matrices, which provides the theoretical foundation for acquiring the same reducts from a compacted decision table as those from its original version. The following theorem is applied for the investigation.

Theorem 1 *Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$. The relationship between discernibility functions generated based on DT and CDT is*

$\mathcal{F}(\mathbf{M}^P_{DT}) = \mathcal{F}(\mathbf{M}^P_{CDT})$.

Proof Suppose that $U = \{x_1, x_2, \dots, x_n\}$ and $CU = \{cx_1, cx_2, \dots, cx_m\}$. From the definition of a compacted decision table, without loss of generality, we further suppose that $U/C = \{X_1, X_2, \dots, X_m\}$, and $f_{DT}(x_{p_i}, C) = f_{CDT}(cx_p, C)$ for $\forall x_{p_i} \in X_p$.

1. $cx_p, cx_q \in CU$, $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$ and $cx_p, cx_q \in CU_1$
   In this case, it is easy to obtain $cx_p, cx_q \in CU_1 \Leftrightarrow x_{p_i}, x_{q_j} \in U_1, (x_{p_i} \in X_p, x_{q_j} \in X_q)$, and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \Leftrightarrow f_{DT}(x_{p_i}, d) \neq f_{DT}(x_{q_j}, d)$. We therefore have $m^P_{p_i q_j} = cm^P_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
2. $cx_p \in CU_1, cx_q \in CU_2$
   In this case, we have $cx_p \in CU_1 \Leftrightarrow x_i \in U_1$ for $\forall x_i \in X_p$, and $cx_q \in CU_2 \Leftrightarrow x_q \in U_2$ for $\forall x_j \in X_q$. We therefore have $m^P_{p_i q_j} = cm^P_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
3. Otherwise
   In this case, it is easy to see that $m^P_{p_i q_j} = cm^P_{pq} = \emptyset$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.

   Furthermore, because of $\bigvee(m^P_{p_i q_j}) = cm^P_{pq}$, we have

$$\begin{aligned}\mathcal{F}(\mathbf{M}^P_{DT}) &= \bigwedge \left\{ \bigvee(m^P_{p_i q_j}) \mid \forall x_{p_i}, x_{q_j} \in U, m^P_{p_i q_j} \neq \emptyset \right\} \\ &= \bigwedge \left\{ \bigvee(cm^P_{pq}) \mid \forall cx_p, cx_q \in CU, cm^P_{pq} \neq \emptyset \right\} \\ &= \mathcal{F}(\mathbf{M}^P_{CDT}). \end{aligned} \tag{8}$$

□

Theorem 1 states that the discernibility function of a compacted decision table is the same as that of its original version, from which it can be observed that all reducts acquired from a decision table are the same as those acquired from its compacted version.

Next, we investigate the relationship between reducts obtained from a decision table and those from its compacted version in the context of Shannon entropy. We therefore propose a discernibility matrix in the sense of Shannon entropy for a compacted decision table. Similar to the discernibility matrix in the context of Shannon entropy proposed in Definition 2, an attribute a_i in which two objects (cx_p and cx_q) of a compacted decision table possess different values must be placed in the corresponding entry of the decision table when these two objects satisfy the following cases: (1) both cx_p and cx_q are in the consistent part (CU_1) of a compacted decision table and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$; (2) cx_p is in the consistent part (CU_1) and cx_q is in the inconsistent part (CU_2) of a compacted decision table; (3) both cx_p and cx_q are in the inconsistent part (CU_2) of a compacted decision table, and $\exists d_k \in CD$ such that $\frac{f_{CDT}(cx_p, d_k)}{\sum_{k=1}^{|CD|} f_{CDT}(cx_p, d_k)} \neq \frac{f_{CDT}(cx_q, d_k)}{\sum_{k=1}^{|CD|} f_{CDT}(cx_q, d_k)}$. The reason for distinguishing between them in these three cases is that if the two objects (x_i and x_j) become one object in a new compacted decision table ($S' = (CU, C' \cup CD)$, $C' \subset C$), the Shannon entropy $H^{CU}_{CDT}(CD|C')$ will be smaller than the Shannon entropy $H^{CU}_{CDT}(CD|C)$. However, the attribute a_i has the ability to distinguish between x_i and x_j (i.e. $H^{CU}(CD|C' \cup \{a_i\}) = H^{CU}(CD|C')$), and therefore it must be placed in the entry (cm^S_{ij}) of the discernibility matrix (M^S_{CDT}). Moreover, when both cx_p and cx_q do not conform to Case 1, or Case 2, or Case 3 mentioned above, they do not need to be distinguished, and the entry corresponding to the two objects should be an empty set.

Based on the above discussion, we can propose the following discernibility matrix in the context of Shannon entropy.

Definition 7 Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$, then a discernibility matrix in the context of Shannon entropy is defined as $\mathbf{M}^S_{CDT} = \{cm^S_{pq}\}$, where

$$cm^S_{pq} = \begin{cases} \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\}, \\ \qquad \{cd_k | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k | f_{CDT}(cx_q, cd_k) \neq 0\} \text{ and} \\ \qquad cx_p, cx_q \in CU_1; \\ \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\},\ cx_p \in CU_1, cx_q \in CU_2; \\ \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\}, \\ \qquad \exists cd_k \in CD \textit{ such that } \frac{f_{CDT}(cx_p,cd_k)}{\sum_{k=1}^{|CD|} f_{CDT}(cx_p,cd_k)} \neq \frac{f_{CDT}(cx_q,cd_k)}{\sum_{k=1}^{|CD|} f_{CDT}(cx_q,cd_k)} \text{ and} \\ \qquad cx_p, cx_q \in CU_2; \\ \emptyset,\ otherwise; \end{cases} \tag{9}$$

To demonstrate that the same reducts as those obtained by the discernibility matrix in Definition 2 can be captured by our proposed discernibility matrix in the context of Shannon entropy, we will need to investigate the relationship between the two discernibility functions based on these two discernibility matrices, which provides the theoretical foundation for getting the same reducts from a compacted decision table as those from its original version. The following theorem is applied for the investigation.

Theorem 2 *Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$, the relationship between discernibility functions generated from DT and CDT is*

$\mathcal{F}(\mathbf{M}^S_{DT}) = \mathcal{F}(\mathbf{M}^S_{CDT})$.

Proof Suppose that $U = \{x_1, x_2, \ldots, x_n\}$ and $CU = \{cx_1, cx_2, \ldots, cx_m\}$. From the definition of a compacted decision table, without loss of generality, we further assume that $U/C = \{X_1, X_2, \ldots, X_m\}$, and $f_{DT}(x_{p_i}, C) = f_{CDT}(cx_p, C)$ for $\forall x_{p_i} \in X_p$.

1. $cx_p, cx_q \in CU$, $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$ and $cx_p, cx_q \in CU_1$.
   In this case, it is easy to obtain $cx_p, cx_q \in CU_1 \Leftrightarrow x_{p_i}, x_{q_j} \in U_1, (x_{p_i} \in X_p, x_{q_j} \in X_q)$, and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\} \Leftrightarrow f_{DT}(x_{p_i}, d) \neq f_{DT}(x_{q_j}, d)$. Therefore, we have $m^S_{p_iq_j} = cm^S_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
2. $cx_p \in CU_1, cx_q \in CU_2$
   In this case, it is easy to obtain $cx_p \in CU_1 \Leftrightarrow x_p \in U_1$ for $\forall x_i \in X_p$, and $cx_q \in CU_2 \Leftrightarrow x_q \in U_2$ for $\forall x_j \in X_q$. We therefore have $m^S_{p_iq_j} = cm^S_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
3. $f(cx_i, cd_k) = f(cx_j, cd_k)$ for $\forall cd_k \in CD$, and $cx_p, cx_q \in CU_2 cx_p, cx_q \in CU_2$
   In this case, it is easy to obtain $cx_p, cx_q \in CU_2 \Leftrightarrow x_p, x_q \in U_2$. From the definition of a compacted decision table, we have $f_{CDT}(cx_p, cd_k) = X_p \cap Y_k$ and $f_{CDT}(cx_q, cd_k) = X_q \cap Y_k$, thus $f_{CDT}(cx_i, cd_k) \neq f_{CDT}(cx_j, cd_k)$ for $\forall cd_k \in CD \Leftrightarrow \mu_{pk} = \frac{|X_p \cap Y_k|}{|X_p|} \neq \frac{|X_q \cap Y_k|}{|X_q|} = \mu_{qk}$ $\mu_{pk} = \frac{|X_p \cap Y_k|}{|X_p|} \neq \frac{|X_q \cap Y_k|}{|X_q|} = \mu_{qk}$ for $\forall Y_k \in U/\{d\}$. Therefore we have $m^S_{p_iq_j} = cm^S_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
4. Otherwise
   In this case, it can be observed that $m^S_{p_iq_j} = cm^S_{pq} = \emptyset$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
   Furthermore, because of $\bigvee(m^S_{p_iq_j}) = cm^S_{pq}$, we have

$$\begin{aligned} \mathcal{F}(\mathbf{M}^S_{DT}) &= \bigwedge \left\{ \bigvee(m^S_{p_iq_j}) \mid \forall p_i, q_j \in U, m^S_{p_iq_j} \neq \emptyset \right\} \\ &= \bigwedge \left\{ \bigvee(cm^S_{pq}) \mid \forall cx_p, cx_q \in CU, cm^S_{pq} \neq \emptyset \right\} \\ &= \mathcal{F}(\mathbf{M}^S_{CDT}). \end{aligned} \tag{10}$$

□

From Theorem 2, we can observe that the discernibility function of a compacted decision table is the same as that of its original version. Therefore, it is easy to observe that all reducts acquired from a decision table are the same as those acquired from its compacted version.

Finally, we analyze the relationship between reducts obtained from a decision table and from its compacted version in the context of complement entropy. To this end, a new discernibility matrix in the context of complement entropy is introduced. Similar to the discernibility matrix in the context of complement entropy proposed in Definition 2, an attribute a_i in which two objects (cx_p and cx_q) of a compacted decision table possess different values must be placed in the corresponding entry of the decision table when these two objects satisfy the following cases: (1) both cx_p and cx_q are in the consistent part (CU_1) of a compacted decision table and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$; (2) cx_p is in consistent part (CU_1) and cx_q is in the inconsistent part (CU_2) of a compacted decision table; (3) both cx_p and cx_q are in the inconsistent part (CU_2) of a compacted decision table. The reason for distinguishing between them in these two cases is that if the two objects (x_i and x_j) become one object in a new compacted decision table ($S' = (CU, C', CD), C' \subset C$), the complement entropy $E^{CU}_{CDT}(CD|C')$ will be smaller than the complement entropy $E^{CU}_{CDT}(CD|C)$. However, the attribute a_i has the ability to distinguish between x_i and x_j (i.e. $E^{CU}(CD|C' \cup \{a_i\}) = E^{CU}(CD|C')$), and therefore it must be placed in the entry (cm^C_{ij}) of the discernibility matrix (CM^C_{CDT}). Moreover, when both cx_p and cx_q do not conform to Case 1, or Case 2, or Case 3 as mentioned above, the two objects do not need to be distinguished, and the entry corresponding to the two objects must be an empty set.

Based on the above discussion, we can propose the following discernibility matrix in the context of complement entropy.

Definition 8 Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$, a discernibility matrix in the context of complement entropy is defined as $\mathbf{M}^C_{CDT} = \{cm^C_{pq}\}$, where

$$cm^C_{pq} = \begin{cases} \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\}, \\ \qquad \{d_k | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k | f_{CDT}(cx_q, cd_k) \neq 0\} \text{ and} \\ \qquad cx_p, cx_q \in CU_1; \\ \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\},\ cx_p \in CU_1, cx_q \in CU_2; \\ \{c \in C : f_{CDT}(cx_p, c) \neq f_{CDT}(cx_q, c)\},\ cx_p, cx_q \in CU_2; \\ \emptyset,\ otherwise; \end{cases} \tag{11}$$

To demonstrate that the same reducts as those obtained by the discernibility matrix in Definition 3 can be captured by our proposed discernibility matrix in the context of complement entropy (shown in Definition 8), we will need to investigate the relationship between the two discernibility functions based on these two discernibility matrices, which provides the theoretical foundation for getting the same reducts from a compacted decision table as those from its original version. The following theorem is applied for the investigation.

Theorem 3 *Given a decision table $DT = (U, C \cup \{d\})$ and its compacted version $CDT = (CU, C \cup CD)$, the relationship between discernibility matrices generated from DT and CDT is*

$\mathcal{F}(\mathbf{M}^C_{DT}) = \mathcal{F}(\mathbf{M}^C_{CDT})$.

Proof Suppose that $U = \{x_1, x_2, \dots, x_n\}$ and $CU = \{cx_1, cx_2, \dots, cx_m\}$. From the definition of a compacted decision table, without loss of generality, we further assume that $U/C = \{X_1, X_2, \dots, X_m\}$, and $f_{DT}(x_{p_i}, C) = f_{CDT}(cx_p, C)$ for $\forall x_{p_i} \in X_p$.

1. $cx_p, cx_q \in CU$, $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{cd_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\}$ and $cx_p, cx_q \in CU_1$.
   In this case, it is easy to obtain $cx_p, cx_q \in CU_1 \Leftrightarrow x_{p_i}, x_{q_j} \in U_1, (x_{p_i} \in X_p, x_{q_j} \in X_q)$, and $\{cd_k \in CD | f_{CDT}(cx_p, cd_k) \neq 0\} \neq \{d_k \in CD | f_{CDT}(cx_q, cd_k) \neq 0\} \Leftrightarrow f_{DT}(x_{p_i}, d) \neq f_{DT}(x_{q_j}, d)$. Therefore, we have $m^C_{p_i q_j} = cm^C_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
2. $cx_p \in CU_1, cx_q \in CU_2$
   In this case, it is easy to obtain $cx_p \in CU_1 \Leftrightarrow x_i \in U_1$ for $\forall x_i \in X_p$, and $cx_q \in CU_2 \Leftrightarrow x_q \in U_2$ for $\forall x_j \in X_q$. Therefore we have $m^C_{p_i q_j} = cm^C_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
3. $cx_p, cx_q \in CU_2$
   In this case, it is easy to obtain$cx_p, cx_q \in CU_2 \Leftrightarrow cx_p, x_q \in U_2$ for $\forall x_i \in X_p$. Therefore we have $m^C_{p_i q_j} = cm^C_{pq}$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.
4. Otherwise
   In this case, it is easy to observe that $m^C_{p_i q_j} = cm^C_{pq} = \emptyset$ for $\forall x_{p_i} \in X_p, \forall x_{q_j} \in X_q$.

Furthermore, because of $\bigvee(m^C_{p_i q_j}) = cm^C_{pq}$, we have

$$\begin{aligned} \mathcal{F}(\mathbf{M}^C_{DT}) &= \bigwedge \left\{ \bigvee(m^C_{p_i q_j}) \mid \forall p_i, q_j \in U, m^C_{p_i q_j} \neq \emptyset \right\} \\ &= \bigwedge \left\{ \bigvee(cm^C_{pq}) \mid \forall cx_p, cx_q \in CU, cm^C_{pq} \neq \emptyset \right\} \\ &= \mathcal{F}(\mathbf{M}^C_{CDT}). \end{aligned} \tag{12}$$

□

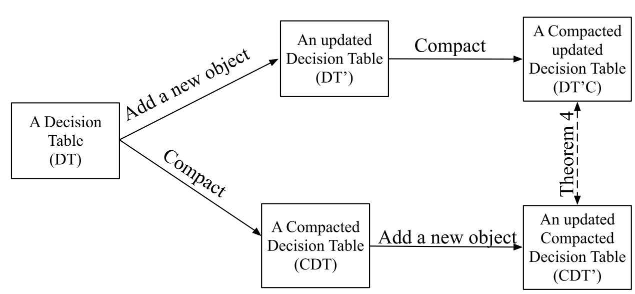


Fig. 1 The relationships among DT, DT', CDT, CDT' and $DT'C$

This theorem indicates that the discernibility function of a decision table is identical with that of its compacted version. Because of its discernibility function, all reducts of a decision table can be captured, and the reducts derived from a compacted decision table are the same as those derived from its original version.

4 Relationship between the reducts of an updated compacted decision table (CDT′) and its corresponding updated decision table (DT′)

After adding a new object into a compacted decision table, are the reducts acquired from the updated compacted table the same as those derived from the updated original table or not? If so, we can compute reducts of an updated decision table by means of its corresponding updated compacted decision table. And how to compute the new reducts of the updated compacted decision table is also becoming an important question. To answer these questions, we first investigate changes in a compacted decision table after adding a new object into it.

4.1 An updated compacted decision table

In this section, we investigate what an updated compacted decision table (CDT′) is like. In other words, we investigate the new compacted decision table generated by adding a new object into a compacted decision table (CDT), which provides the necessary context for the next section.

For facilitating the following analysis, we assume that $CDT = (CU, C \cup CD)$ is a compacted decision table, where $CU = \{cx_1, cx_2, ..., cx_u\}$ $(u = |U/C|)$ and $CD = \{cd_1, cd_2, ..., cd_s\}$ $(s = |U/\{d\}|)$. Suppose that $CDT' = (CU', C \cup CD')$ is an updated compacted decision table generated by adding a new object x_{new} into CDT, where $CU' = \{cx'_1, cx'_2, ..., cx'_v\}$, and $CD' = \{cd'_1, cd'_2, ..., cd'_t\}$. For the new object x_{new}, we use $f(x_{new}, C)$ to indicate the values of x_{new} on the condition attribute set C and $f(x_{new}, d)$ to indicate the decision value of x_{new}.

From Definition 4 and the relationships among a new object and the objects in a compacted decision table, we investigate the incremental change mechanism of a compacted decision table in the following four cases:

1. $\exists cx_p \in CU$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$, and $f(x_{new}, d) \in V_d$.
   In this case, because $\exists cx_p \in CU$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \in V_d$, it can be observed that $|CU'| = |CU| = u$ and $|CD'| = |CD| = s$. Without loss of generality, we can assume that $f_{CDT'}(cx'_i, C) = f_{CDT}(cx_i, C)\,(1 \le i \le u)$, and $f(x_{new}, d) = v_{d_r}$ $(v_{d_r} \in V_d)$. Therefore, we have $f_{CDT'}(cx'_p, cd'_r) = f_{CDT}(cx_p, cd_r) + 1$, $f_{CDT'}(cx'_p, cd'_j) = f_{CDT}(cx_p, cd_j)(1 \le j \le s, j \ne r)$, and $f_{CDT'}(cx'_i, cd'_j) = f_{CDT}(cx_i, cd_j)(1 \le i \le u, i \ne p, 1 \le j \le s)$.
2. $\exists cx_p \in CU$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$, and $f(x_{new}, d) \notin V_d$.
   In this case, because $\exists cx_p \in CU$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \notin V_d$, it can be observed that $|CU'| = |CU| = u$ and $|CD'| = |CD| + 1 = s + 1$. Without loss of generality, we can assume that $f_{CDT'}(cx'_i, C) = f_{CDT}(cx_i, C)$ $(1 \le i \le u)$ and $f_{CDT'}(cx'_i, cd'_j) = f_{CDT}(cx_i, cd_j)$ $(1 \le i \le u,\ 1 \le j \le s)$, $f_{CDT'}(cx'_p, cd'_{s+1}) = 1$, and $f_{CDT'}(cx'_i, cd'_{s+1}) = 0\,(1 \le i \le u, i \ne p)$.
3. $\forall cx_p \in CU$, $f(x_{new}, C) \ne f_{CDT}(cx_p, C)$, and $f(x_{new}, d) \in V_d$.
   In this case, because $\forall cx_p \in CU$, $f(x_{new}, C) \ne f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \in V_d$, it can be observed that $|CU'| = |CU| + 1 = u + 1$ and $|CD'| = |CD| = s$. Without loss of generality, we can assume that $f_{CDT'}(cx'_i, C) = f_{CDT}(cx_i, C)$ $(1 \le i \le u)$, $f_{CDT'}(cx'_{u+1}, C) = f(x_{new}, C)$, and $f(x_{new}, d) = v_{d_r}(v_{d_r} \in V_d)$. We therefore have $f_{CDT'}(cx'_i, cd'_j) = f_{CDT}(cx_i, cd_j)\,(1 \le i \le u, 1 \le j \le s)$, and $f_{CDT'}(cx'_{u+1}, cd'_r) = 1$ and $f_{CDT'}(cx'_{u+1}, cd'_j) = 0$ $(1 \le j \le u, j \ne r)$.

Table 1 Data sets used in experiments

ID	Datasets	Abbreviation	Samples	Attributes	Consistent	Type
1	Monk's problem	Monk	1711	7	No	Categorical
2	Wine	Wine	178	13	Yes	Numerical
3	Breast Cancer Wisconsin(Original)	BCW	683	9	Yes	Categorical
4	Banknote authentication	BA	1372	4	No	Numerical
5	Blood Transfusion Service Center	BTSC	748	4	No	Numerical
6	Image segmentation	IS	2310	18	No	Numerical
7	Page blocks classification	PBC	5473	10	No	Numerical
8	Seismic-bumps	SB	2584	11	No	Numerical
9	Wine quality	WQ	4898	11	No	Numerical

4. $\forall cx_p \in CU$, $f(x_{new}, C) \neq f_{CDT}(cx_p, C)$, and $f(x_{new}, d) \notin V_d$. In this case, because of $\forall cx_p \in CU$, $f(x_{new}, C) \neq f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \notin V_d$, it can be observed that $|CU'| = |CU| + 1 = u + 1$ and $|CD'| = |CD| + 1 = s + 1$. Without loss of generality, assume that $f_{CDT'}(cx'_i, C) = f_{CDT}(cx_i, C)$ $(1 \leq i \leq u)$, $f_{CDT'}(cx'_{u+1}, C) = f(x_{new}, C)$. Moreover, we have $f_{CDT'}(cx'_i, cd'_j) = f_{CDT}(cx_i, cd_j)$ $(1 \leq i \leq u,\ 1 \leq j \leq s)$, $f_{CDT'}(cx'_{u+1}, cd'_{s+1}) = 1$, and $f_{CDT'}(cx'_{u+1}, cd'_j) = 0\,(1 \leq j \leq s)$, and $f_{CDT'}(cx'_i, cd'_{s+1}) = 0$ $(1 \leq i \leq u)$.

The analysis mentioned above illustrates all the possible changes to a compacted decision table after a new object is added into it, which facilitates our investigation of the reducts obtained from an updated decision table after a new object is added into it.

4.2 Relationship between CDT′ and DT′C

In this section, we examine the relationship between reducts obtained from an updated compacted decision table (CDT′) and its corresponding updated decision table (DT′), which can demonstrate the effectiveness of the proposed discernibility matrices. In Sect. 3, we demonstrated that the reducts acquired from a compacted decision table (CDT) are identical to those acquired from its original version (DT). We can leverage this conclusion if an updated compacted decision table (CDT′) and a compacted updated decision table (DT′C) are proven to be the same as each other. Therefore, we first analyze the relationship between an updated compacted decision table (CDT′) and a compacted updated decision table (DT’C) (Shown in Fig. 1) using the following theorem.

Theorem 4 *Given a decision table $DT = \{U, C \cup \{d\}\}$ and its compacted version $CDT = \{CU, C \cup CD\}$, then $DT'C$ is identical to CDT', where $DT'C$ is a compacted table constructed by compacting DT', and DT' and CDT' are the decision table and the compacted decision table generated by adding the object x_{new} into DT and CDT, respectively.*

Proof Suppose that $U/C = \cup_{i=1}^{u} X_i$, $X_i = \cup_{1 \leq j \leq |X_i|}\{x_{i_j}\}$, $U/\{d\} = \cup_{i=1}^{s} Y_i$. There are two cases that should be considered as follows:

(1) $\exists x_{p_w} \in U$ such that $f(x_{new}, C) = f_{DT}(x_{p_w}, C)$.

We suppose that $U'/C = \cup_{i=1}^{u} X'_i$, $X'_i = \cup_{1 \leq j \leq |X'_i|}\{x'_{i_j}\}$, $f_{DT'}(x'_{i_j}, C) = f_{DT}(x_{i_j}, C)$ $(i \neq p, 1 \leq j \leq |X'_i| = |X_i|)$ and $f_{DT'}(x'_{p_j}, C) = f_{DT}(x_{p_w}, C) = f(x_{new}, C)$ $(1 \leq j \leq |X'_p|$, $X'_p = X_p \cup \{x_{new}\})$. From Definition 6, we can obtain a compacted updated decision table $DT'C = \{U'C, C, D'C\}$ by compacting DT' and $f_{DT'C}(x'c_i, C) = f_{DT'}(x'_{i_j}, C) = f_{DT}(x_{i_j}, C) = f_{CDT}(cx_i, C)(1 \leq i \leq u, 1 \leq j \leq |X_i|)$, where $x'c \in U'C$. From the result of case (1) in Sect. 4.1, we have $f_{DT'C}(x'c_i, C) = f_{CDT'}(cx'_i, C)$. According to the relationship between $f(x_{new}, d)$ and V_d, the following analysis is divided into two subcases:

- $f(x_{new}, d) \in V_d$. In this case, we suppose $f(x_{new}, d) = v_{d_r}$, $Y'_r = Y_r \cup \{x_{new}\}$. By Definition 6, it is easy to obtain $f_{DT'C}(cx'_p, d'c_r) = |X'_p \cap Y'_r| = |(X_p \cap Y_r) \cup x_{new}| = |X_p \cap Y_r| + 1 = f_{CDT}(cx_p, cd_r) + 1$, $f_{DT'C}(cx'_p, d'c_j) = |X'_p \cap Y'_j| = f_{CDT}(cx_p, cd_j)|X'_p \cap Y'_j| = f_{CDT}(cx_p, cd_j)(1 \leq j \leq s, j \neq r)$, and $f_{DT'C}(cx'_i, d'c_j) = |X'_i \cap Y'_j| = f_{CDT}(cx_i, cd_j)$ $(1 \leq i \leq u$, $i \neq p, 1 \leq j \leq s)$. By the result of case (1) in Sect. 4.1, we have $f_{DT'C}(cx'_i, d'c_j) = f_{CDT'}(cx_i, cd'_j)\,(1 \leq i \leq u, 1 \leq j \leq s)$.
- $f(x_{new}, d) \notin V_d$. In this case, we suppose $Y'_i = Y_i (1 \leq i \leq s)$, $Y'_{s+1} = \{x_{new}\}$. By Definition 6, it is easy to see that $f_{DT'C}(x'c_p, dc'_{s+1}) = |X'_p \cap Y'_{s+1}| = 1$, $f_{DT'C}(x'c_p, d'c_j) = |X'_p \cap Y'_j| = f_{CDT}(cx_p, cd_j)(1 \leq j \leq s)$, $f_{DT'C}(x'c_i, d'c_j) = |X'_i \cap Y'_j| = f_{CDT}(cx_i, cd_j)$ $(1 \leq i \leq u$, $i \neq p$, $1 \leq j \leq s)$, and $f_{DT'C}(x'c_i, d'c_{s+1}) = 0$ $(1 \leq i \leq u$, $i \neq p)$. By the result of case (1) in Sect. 4.1, we have $f_{DT'C}(x'c_i, d'c_j) = f_{CDT'}(cx'_i, cd'_j)\,(1 \leq i \leq u, 1 \leq j \leq s + 1)$.

(2) $\forall x_{i_j} \in U$, $f(x_{new}, C) \neq f_{DT}(x_{i_j}, C)$

Table 2 The number of objects in each data set

Data sets	*BDS*	X_1	X_2	X_3	X_4	X_5
Monk	1031	136	272	408	544	680
Wine	108	14	28	42	56	70
BCW	413	54	108	162	216	270
BA	827	109	218	327	436	545
BTSC	453	59	118	177	236	295
IS	1390	184	368	552	736	920
PBC	3288	437	874	1311	1748	2185
SB	1554	206	412	618	824	1030
WQ	2943	391	782	1173	1564	1955

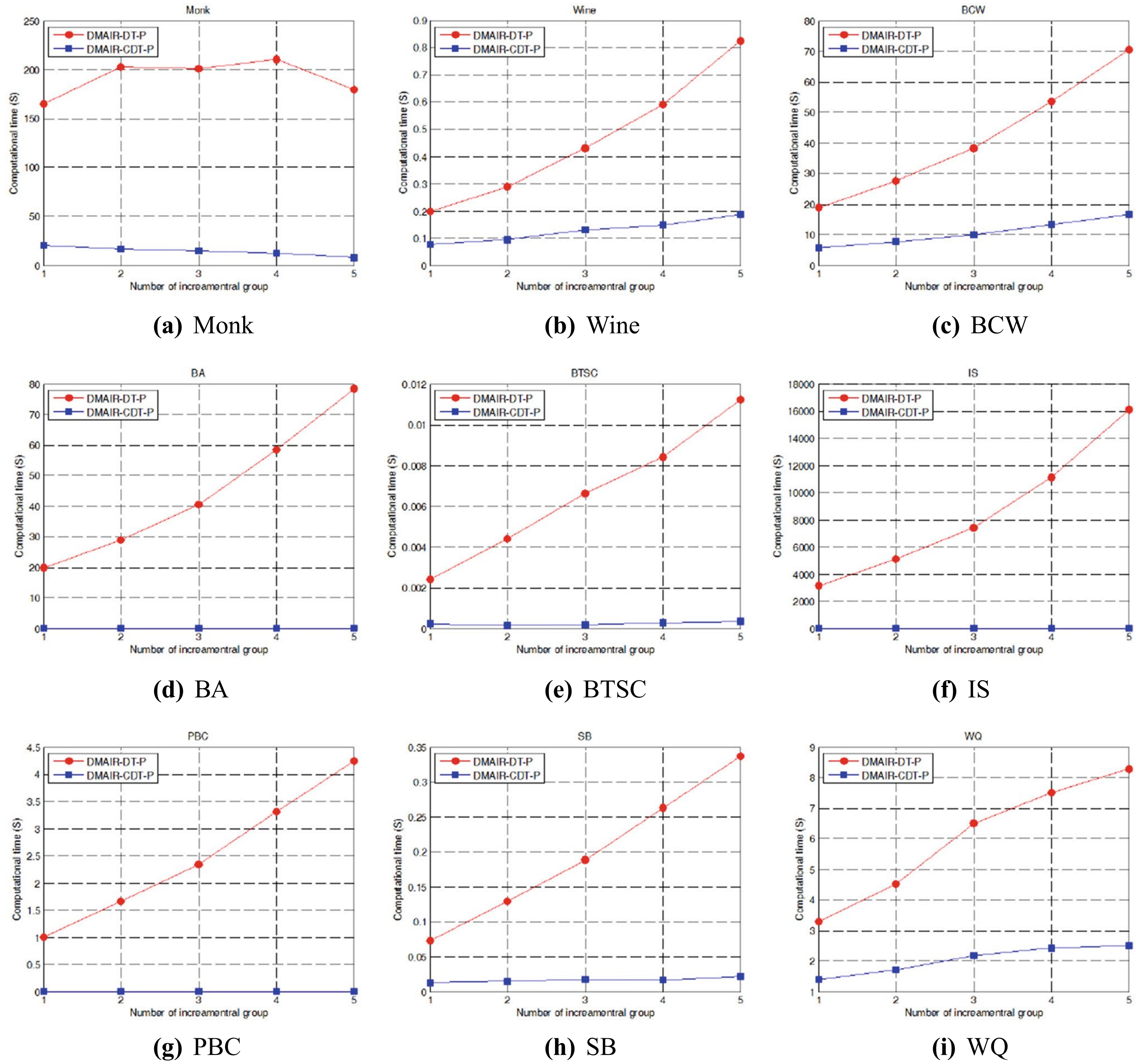


Fig. 2 A comparison of the time taken by DMIAR-DT-P and DMIAR-CDT-P **a** Monk **b** Wine **c** BCW **d** BA **d** BTSC **e** IS **f** PBC **g** SB **h** WQ

We assume that $U'/C = \cup_{i=1}^{u+1} X'_i$, $X'_i = \cup_{1\le j\le |X'_i|}\{x'_{i_j}\}$, $f_{DT'}(x'_{i_j}, C) = f_{DT}(x_{i_j}, C)$ $(1 \le i \le u, 1 \le j \le |X'_i| = |X_i|)$ and $f_{DT'}(x'_{(u+1)_j}, C) = f(x_{new}, C)$ $(j = |X'_{u+1}| = 1)$. By Definition 6, we have $f_{DT'C}(x'c_i, C) = f_{DT'}(x'_{i_j}, C) = f_{DT}(x_{i_j}, C) = f_{CDT}(cx_i, C)(1 \le i \le u, 1 \le j \le |X_i|)$ and $f_{DT'C}(x'c_{(u+1)_j}, C) = f(x_{new}, C)$. By the result of case (1) in Sect. 4.1, we have $f_{DT'C}(x'c_i, C) = f_{CDT'}(cx'_i, C)$. According to the relationship between $f(x_{new}, d)$ and $f_{DT}(x_{p_w}, d)$, the following analysis is divided into two subcases:

- $f(x_{new}, d) \in V_d$. Similar to Case (1), based on the results of Case (3) in Sect. 4.1, we have $f_{DT'C}(x'c_i, d'c_j) = f_{CDT'}(cx'_i, cd'_j)$ $(1 \le i \le u+1, 1 \le j \le s)$.
- $f(x_{new}, d) \notin V_d$. Similar to case (1), based on the results of Case (4) in Sect. 4.1, we have $f_{DT'C}(x'c_i, d'c_j) = f_{CDT'}(cx'_i, cd'_j)$ $(1 \le i \le u+1, 1 \le j \le s+1)$.

□

From Theorem 4, we can observe that an updated compacted decision table (CDT′) is identical to a compacted updated decision table (DT′C), it is easy to be seen in Fig. 1.

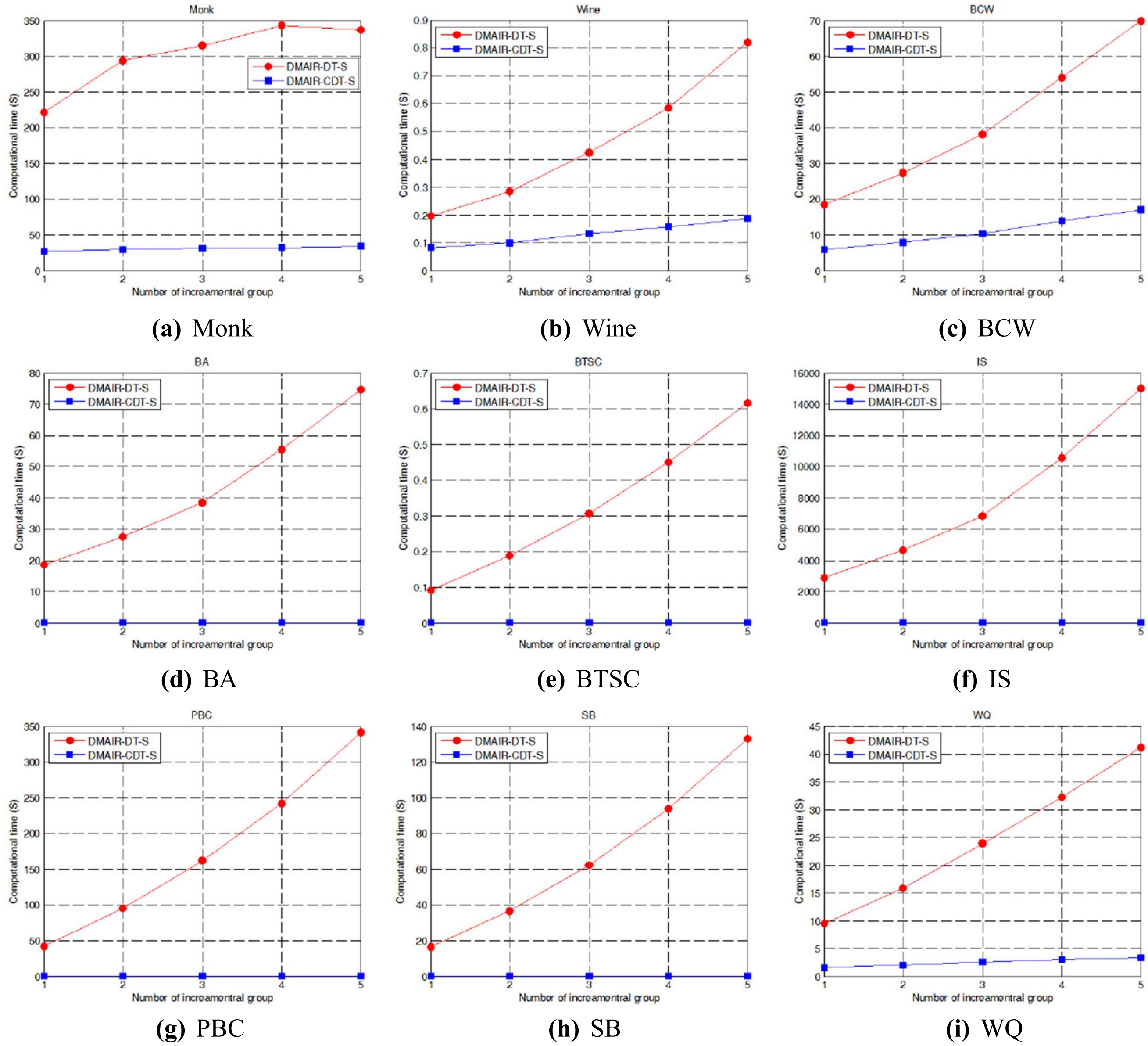


Fig. 3 A comparison of the time taken by DMIAR-DT-S and DMIAR-CDT-S **a** Monk **b** Wine **c** BCW **d** BA **e** BTSC **f** IS **g** PBC **h** SB **i** WQ

4.3 Relationship between reducts obtained by a CDT′ and its corresponding DT′

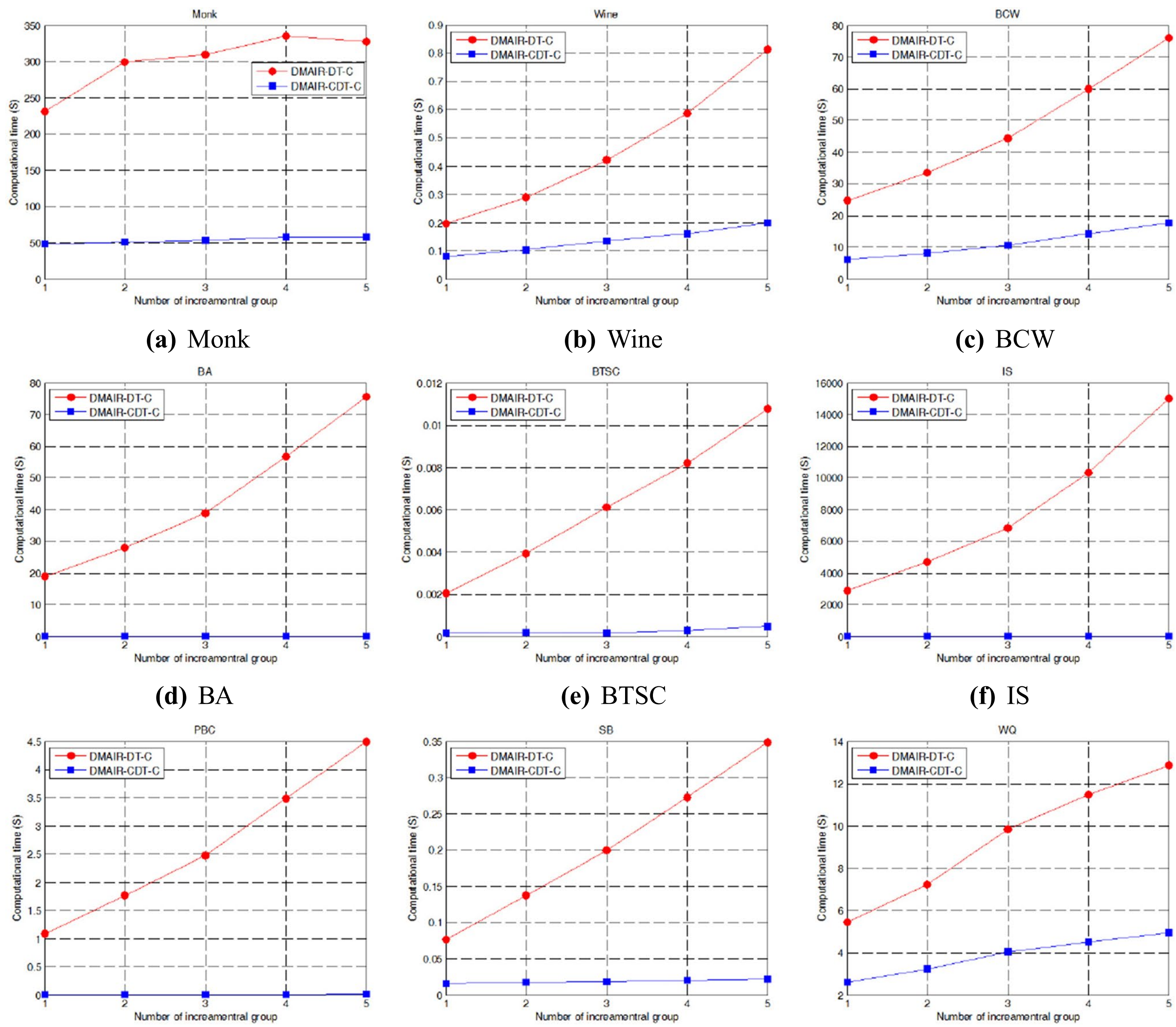
According to Theorem 4, it is easy to infer the following two corollaries to indicate the relationship between reducts obtained by a CDT′ and its corresponding DT′.

Corollary 1 *Given a decision table* $DT = \{U, C \cup \{d\}\}$ *and its compacted version* $CDT = \{CU, C \cup CD\}$, *if* DT' *is a decision table generated by adding a new object* x_{new} *into* DT, *and* CDT' *is a compacted decision table generated by adding the object* x_{new} *into* CDT, *then*

$$\mathcal{F}(\mathbf{M}^{P}_{DT'}) = \mathcal{F}(\mathbf{M}^{P}_{CDT'}).$$

Proof From the results of Theorem 4, it can be observed that $DT'C$ is the same as CDT'. Furthermore, we can conclude $\mathcal{F}(\mathbf{M}^{P}_{DT'}) = \mathcal{F}(\mathbf{M}^{P}_{CDT'})$ using Theorem 1. □

Corollary 1 demonstrates that when a new object is added into a decision table and its compacted version, the discernibility function in the context of the positive region derived from an updated decision table is identical to that derived from the updated compacted table, which indicates that one can obtain the same reducts through the updated compacted version of a decision table as through the updated



(a) Monk **(b)** Wine **(c)** BCW **(d)** BA **(e)** BTSC **(f)** IS **(g)** PBC **(h)** SB **(i)** WQ

Fig. 4 A comparison of the time taken DMIAR-DT-C and DMIAR-CDT-C **a** Monk **b** Wine **c** BCW **d** BA **e** BTSC **f** IS **g** PBC **h** SB **i** WQ

decision table itself. We can also prove the equality of the discernibility functions of an updated decision table and an updated compacted decision table in the context of Shannon entropy and complement entropy based on the following two corollaries.

Corollary 2 *Given a decision table* $DT = \{U, C \cup \{d\}\}$ *and its compacted version* $CDT = \{CU, C \cup CD\}$, *if* DT' *is a decision table generated by adding a new object* x_{new} *into* DT, *and* CDT' *is a compacted decision table generated by adding the object* x_{new} *into* CDT, *then*

$$\mathcal{F}(\mathbf{M}^S_{DT'}) = \mathcal{F}(\mathbf{M}^S_{CDT'}),\ \mathcal{F}(\mathbf{M}^C_{DT'}) = \mathcal{F}(\mathbf{M}^C_{CDT'}).$$

The proof of this corollary is similar to that of Corollary 1 and so the proof has been omitted.

5 Algorithms and experimental analysis

Several experiments were performed to demonstrate the effectiveness of our proposed algorithm for dynamic data sets with dynamic object sets. In the experiments, nine data sets downloaded from the UCI Machine Learning Database Repository and a synthetic data set were chosen as benchmark data sets for performance tests. To ensure that the data set is suitable for our algorithms, which aim to analyze

Table 3 Comparison of non-empty entries in the discernibility matrix in the context of positive region

Data set	Algorithm	$BDS+X_1$	$BDS+X_2$	$BDS+X_3$	$BDS+X_4$	$BDS+X_5$
Monk	DMIAR-DT-P	154741	169841	168791	172255	158448
	DMIAR-CDT-P	55851	52098	49411	46791	41944
Wine	DMIAR-DT-P	4826	6008	7324	8756	10299
	DMIAR-CDT-P	2499	2990	3624	4201	4715
BCW	DMIAR-DT-P	51799	63832	76467	90744	104784
	DMIAR-CDT-P	28704	33432	38430	44550	49629
BA	DMIAR-DT-P	54514	66905	79289	95695	110483
	DMIAR-CDT-P	330	331	331	331	349
BTSC	DMIAR-DT-P	43	44	44	44	44
	DMIAR-CDT-P	24	24	24	24	25
IS	DMIAR-DT-P	597264	729082	853493	1024383	1211070
	DMIAR-CDT-P	29827	32786	35180	38833	41229
PBC	DMIAR-DT-P	11318	13970	16106	19062	21580
	DMIAR-CDT-P	741	857	898	907	985
SB	DMIAR-DT-P	1912	2130	2317	2539	2700
	DMIAR-CDT-P	1293	1355	1386	1457	1491
WQ	DMIAR-DT-P	26568	30083	35002	37354	38870
	DMIAR-CDT-P	15126	17213	19685	21079	22388

Table 4 Comparison of non-empty entries in the discernibility matrix in the context of Shannon entropy

Data set	Algorithm	$BDS+X_1$	$BDS+X_2$	$BDS+X_3$	$BDS+X_4$	$BDS+X_5$
Monk	DMIAR-DT-S	178288	203596	209235	217706	214393
	DMIAR-CDT-S	63393	65807	67778	69114	71246
Wine	DMIAR-DT-S	4826	6008	7324	8756	10299
	DMIAR-CDT-S	2499	6008	3624	4201	4715
BCW	DMIAR-DT-S	51799	63832	76467	90744	104784
	DMIAR-CDT-S	28704	33432	38430	44550	49629
BA	DMIAR-DT-S	54469	66860	79244	95650	110438
	DMIAR-CDT-S	314	314	314	314	332
BTSC	DMIAR-DT-S	23	23	23	23	23
	DMIAR-CDT-S	15	15	15	15	15
IS	DMIAR-DT-S	596991	728785	853145	1024008	1210693
	DMIAR-CDT-S	29770	32699	35042	38668	41063
PBC	DMIAR-DT-S	11213	13865	16001	18942	21460
	DMIAR-CDT-S	716	831	871	865	943
SB	DMIAR-DT-S	1686	1903	2089	2287	2447
	DMIAR-CDT-S	1226	1286	1316	1362	1393
WQ	DMIAR-DT-S	20869	23592	28015	29506	30372
	DMIAR-CDT-S	14232	15534	17532	18070	18735

categorical data sets, we discretize each numerical attribute into three equilength intervals and assign each interval a symbol to ensure that the attributes of the numerical datasets in Table 1 hold three categorical values. Experimental results were then collected on a personal computer equipped with an Intel Core i7 processor and 8 GB memory, and all code was programmed in *C*#.

5.1 Incremental attribute reduction algorithms

To computing all reducts of a decision table, we design an incremental attribute reduction algorithm based on discernibility matrices in Wei et al. [42].

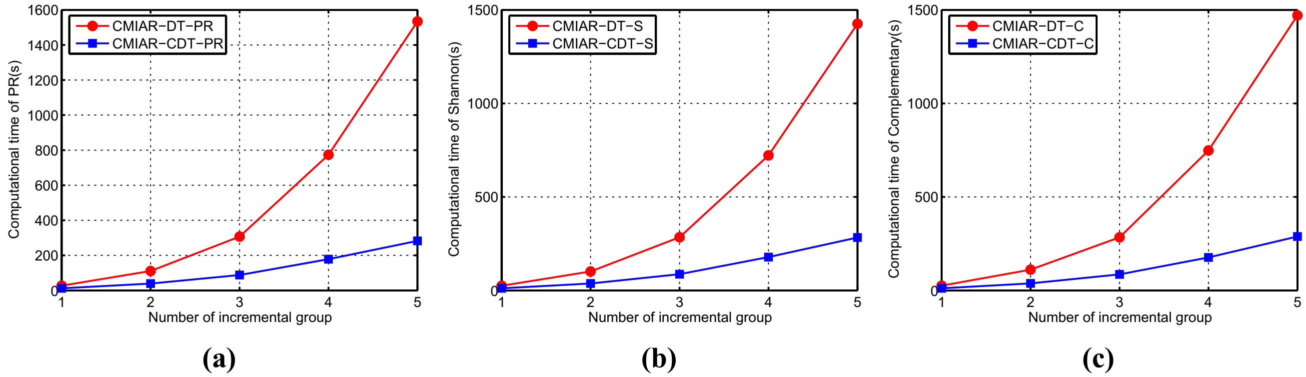


Fig. 5 Comparison of elapsed times on a synthetic data set

Table 5 Comparison of non-empty entries in the discernibility matrix in the context of complement entropy

Data set	Algorithm	$BDS+X_1$	$BDS+X_2$	$BDS+X_3$	$BDS+X_4$	$BDS+X_5$
Monk	DMIAR-DT-C	182236	205086	209261	21770	214393
	DMIAR-CDT-C	82680	85065	86955	88260	90372
Wine	DMIAR-DT-C	4826	6008	7324	8756	10299
	DMIAR-CDT-C	2499	2990	3624	4201	4715
BCW	DMIAR-DT-C	51799	63832	76467	90744	104784
	DMIAR-CDT-C	28704	33432	38430	44550	49629
BA	DMIAR-DT-C	54514	66905	79289	95695	110483
	DMIAR-CDT-C	359	359	359	359	377
BTSC	DMIAR-DT-C	44	44	44	44	44
	DMIAR-CDT-C	36	36	36	36	36
IS	DMIAR-DT-C	597267	729085	853496	1024386	1211071
	DMIAR-CDT-C	30046	32999	35393	39046	41441
PBC	DMIAR-DT-C	11318	13970	16106	19062	21580
	DMIAR-CDT-C	821	936	976	985	1063
SB	DMIAR-DT-C	1917	2134	2320	2540	2700
	DMIAR-CDT-C	1457	1517	1547	1615	1646
WQ	DMIAR-DT-C	26647	30147	35036	37381	38887
	DMIAR-CDT-C	20010	22089	24553	25945	27250

Table 6 Comparison of reducts for $BDS+X_5$ acquired by DMIAR-DT and DMIAR-CDT

Data sets	DMIAR-DT-P	DMIAR-CDT-P	DMIAR-DT-S	DMIAR-CDT-S	DMIAR-DT-C	DMIAR-CDT-C
Monk ($BDS+X_5$)	1	1	1	1	1	1
Wine ($BDS+X_5$)	64	64	64	64	64	64
BCW ($BDS+X_5$)	20	20	20	20	20	20
BA ($BDS+X_5$)	1	1	1	1	1	1
BTSC ($BDS+X_5$)	2	2	2	2	2	2
IS ($BDS+X_5$)	1	1	1	1	1	1
PBC ($BDS+X_5$)	1	1	1	1	1	1
SB ($BDS+X_5$)	2	2	2	2	2	2
WQ ($BDS+X_5$)	1	1	1	1	1	1

Algorithm 1 An incremental attribute reduction algorithm for a decision table (DMIAR-DT)

Input: Decision table $DT = (U, C \cup d)$, the discernibility matrix $\mathbf{M}_{DT}^{\Delta}$ of DT and a new object x_{new};
Output: The set of all reducts RED of the updated decision table $DT' = (U \cup \{x_{new}\}, C \cup \{d\})$.
Step 1: Judge the relationships among the new object x_{new} and the objects in DT:
If x_{new} satisfies case (1) $\exists x_{p_w}$ such that $f(x_{new}, C) = f_{DT}(x_{p_w}, C)$ and $f(x_{new}, D) = f_{DT}(x_{p_w}, D)$, then we must insert a new row $\mathbf{m}_{n+1}^{\Delta} = \mathbf{m}_{p_w}^{\Delta}$, a new column $\mathbf{m}_{\Delta}^{n+1} = \mathbf{m}_{\Delta}^{p_w}$, and a new entry $m_{(n+1)(n+1)} = \emptyset$ into $\mathbf{M}_{DT}^{\Delta}$. //The updated discernibility matrix is a $(n+1) \times (n+1)$ one;
If x_{new} satisfies case (2) $\exists x_{p_w}$ such that $f(x_{new}, C) = f_{DT}(x_{p_w}, C)$ and $f(x_{new}, D) \neq f_{DT}(x_{p_w}, D)$, then we must modify the rows $\mathbf{m}_i^{\Delta}$ and the columns $\mathbf{m}_{\Delta}^{i}$ $(x_i \in X_{p_w})$ according to Definitions 1–3, and then add a new row $\mathbf{m}_{n+1}^{\Delta} = \mathbf{m}_{p_w}^{\Delta}$, a new column $\mathbf{m}_{\Delta}^{n+1} = \mathbf{m}_{\Delta}^{p_w}$, and a new entry $m_{(n+1)(n+1)} = \emptyset$ into $\mathbf{M}_{DT}^{\Delta}$. //The updated discernibility matrix is a $(n+1) \times (n+1)$ one;
If x_{new} satisfies case (3) $\forall x_{p_w}$ such that $f(x_{new}, C) \neq f_{DT}(x_{p_w}, C)$, then we must add a new row $\mathbf{m}_{n+1}^{\Delta}$ and a new column $\mathbf{m}_{\Delta}^{n+1}$ according to Definitions 1–3, and then add a new entry $m_{(n+1)(n+1)} = \emptyset$ into $\mathbf{M}_{DT}^{\Delta}$; //The updated discernibility matrix is a $(n+1) \times (n+1)$ one;
Step 2: Compute the new discernibility function $\mathcal{F}(\mathbf{M}_{DT}^{\Delta})$ by means of the updated discernibility matrix $\mathbf{M}_{DT}^{\Delta}$;
Step 3: Compute RED by using $\mathcal{F}(\mathbf{M}_{DT}^{\Delta})$;
Step 4: Return RED and end.

where $\mathbf{m}_{n+1}^{\Delta}$ and $\mathbf{m}_{p_w}^{\Delta}$ are two n-dimension row vectors, $\mathbf{m}_{\Delta}^{p_w}$ and $\mathbf{m}_{\Delta}^{n+1}$ are two n-dimension column vectors, $\Delta \in \{P, S, C\}$. For convenience, we denote DMIAR-DT-P, DMIAR-DT-S, and DMIAR-DT-C as different versions of algorithm DMIAR-DT based on the positive region, Shannon entropy, and complement entropy, respectively.

Time complexity of Algorithm 1 is as follows: When a new object x_{new} is added to a decision table DT, the number of possible items that vary with the change in a discernibility matrix is $2(|[x_p]_C| + 1) \times |U| \times |C| + 2|U| \times |C|$. Thus, the complexity of updating a discernibility matrix is $O(|U|^2 \times |C|)$. It is common to know that the complexity of obtaining all reducts from a discernibility matrix is $O(2^{|C|})$. Therefore, Algorithm 1's time complexity is $O(|U|^2 \times |C| + 2^{|C|})$.

To accelerate the incremental attribute reduction algorithm, an incremental attribute reduction algorithm are designed based on a compacted decision table. A detailed description of the algorithm is given below.

Algorithm 2 An incremental attribute reduction algorithm for a compacted decision table (DMIAR-CDT)

Input: A compacted decision table $CDT = (CU, C \cup CD)$, the discernibility matrix $\mathbf{M}_{CDT}^{\Delta}$, and a new object x_{new};
Output: The set of all reducts RED of the new decision table $CDT' = (U \cup \{x_{new}\}, C \cup CD)$.
Step 1: Scan CDT to determine the relationship between objects in CDT and the new objects x_{new}:
If the new object agrees with case (1) $\exists cx_p$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \in \{v_{d_k} | f_{CDT}(cx_p, d_k) \neq 0\}$, then the discernibility matrix does not require modification. //the size of the updated matrix is $m \times m$;
If the new object agrees with case (2) $\exists cx_p$ such that $f(x_{new}, C) = f_{CDT}(cx_p, C)$ and $f(x_{new}, d) \notin \{v_{d_k} | f_{CDT}(cx_p, d_k) \neq 0\}$, the row $\mathbf{cm}_p^{\Delta}$ and the column $\mathbf{cm}_{\Delta}^{p}$ must be modified according to Definition 1. //the size of the updated matrix is $m \times m$;
If the new object agrees with case (3) $\forall cx_p$ such that $f(x_{new}, C) \neq f_{CDT}(cx_p, C)$, then we need to add a new row $\mathbf{cm}_{n+1}^{\Delta}$ and a new column $\mathbf{cm}_{\Delta}^{n+1}$ with values specified by Definitions 1–3, and then add a new entry $cm_{(m+1)(m+1)} = \emptyset$. //the size of the updated matrix is $(m+1) \times (m+1)$;
Step 2: Compute a new discernibility function $\mathcal{F}(\mathbf{M}_{CDT}^{\Delta})$ from the updated discernibility matrix $\mathbf{M}_{CDT}^{\Delta}$;
Step 3: Reduce the discernibility function and obtain the set of all the reducts (RED) of the updated CDT;
Step 4: Return RED and end;

where $\mathbf{cm}_{n+1}^{\Delta}$ and $\mathbf{cm}_{p}^{\Delta}$ are two n-dimension row vectors, $\mathbf{cm}_{\Delta}^{p}$ and $\mathbf{cm}_{\Delta}^{n+1}$ are two n-dimension column vectors, $\Delta = \{P, S, C\}$. For convenience, we denote DMIAR-CDT-P, DMIAR-CDT-S, and DMIAR-CDT-C as different versions of algorithm DMIAR-CDT based on the positive region, Shannon entropy, and complement entropy, respectively.

Time complexity of Algorithm 2 is as follows: When a new object x_{new} is added to a decision table CDT, the number of possible items that vary with the change in a discernibility matrix is $2|CU| \times |C| + 2|CU| \times |C|$. Thus, the complexity of updating a discernibility matrix is $O(|CU| \times |C|)$. It is well known that the complexity of obtaining all reducts by using a discernibility matrix is $O(2^{|C|})$. Therefore, Algorithm 2's time complexity is $O\big(|CU| \times |C| + 2^{|C|}\big)$.

Note that, according to the results of Corollaries 1–2, it is easy to know the same reducts as those of an updated decision table can be captured by its compacted version. Thus, we can obtain the same reducts by Algorithm 1 and by Algorithm 2.

5.2 Efficiency analysis

To illustrate the efficiency of the algorithms based on compacted decision tables, for each data set in Table 1, we randomly chose 60% of the objects in the data set as the basic data set (*BDS*), and we divided other objects into five equal parts, denoted by x_i $(i \in \{1, 2, \ldots, 5\})$. Let $X_i = \cup_{j=1}^{i} x_i$ represent the object sets added into the basic data set. The concrete partition of each data set is shown in Table 2.

When each object in X_i was added to the basic data set, the two types of incremental reduction algorithms (DMIAR-DT-Δ and DMIAR-CDT-Δ) were employed to compute all the reducts of a decision table and a compacted decision table, respectively. The runtimes of the algorithms are used to evaluate their performance from the perspective of efficiency, and Figs. 2, 3, 4 illustrate these runtimes of DMIAR-DT-P and DMIAR-CDT-P, DMIAR-DT-S and DMIAR-CDT-S, and DMIAR-DT-C and DMIAR-CDT-C. From Fig. 2, we can observe that the runtime of DMIAR-CDT-P is much smaller than that of DMIAR-DT-P on all the UCI data sets in Table 1. The experimental results illustrate that our proposed algorithm DMIAR-CDT-P is more efficient than DMIAR-DT-P. Figures 3 and 4 demonstrate similar results to those in Fig. 2 and indicate that DMIAR-CDT-S and DMIAR-CDT-C are also faster than DMIAR-DT-S and DMIAR-DT-C, respectively. It is worth noting that the performance of these proposed algorithms improves as the number of added objects increases. In other words, the larger the new data set that is input into the basic data set, the more efficient the algorithms DMIAR-CDT-P, DMIAR-CDT-S, and DMIAR-CDT-C are.

However, it can be observed that the advantages of DMIAR-CDT over DMIAR-DT are insignificant on the data set Harberman's Survival (HS), which raises the question of how can these updated reducts be computed more quickly by DMIAR-CDT-P than DMIAR-DT-P. To answer this question, Table 3 is employed, from which we can determine that for each dataset in Table 1, the number of non-empty set entries in the discernibility matrix generated by Algorithm DMIAR-CDT-P is much smaller than those generated by Algorithm DMIAR-DT-P. Tables 4 and 5 illustrate the existence of similar relationships between the numbers of non-empty entries in the discernibility matrices of DMIAR-CDT-S and of DMIAR-DT-S and between the numbers of non-empty entries in the discernibility matrices of DMIAR-CDT-C and of DMIAR-DT-C. Moreover, as we all know, all the reducts of a decision table or of a compacted decision table can be acquired through its discernibility matrix [42]. Therefore, it is obvious that the smaller the number of non-empty set entries in a discernibility matrix generated by an incremental attribute reduction algorithm, the more efficient is the algorithm.

From the experimental analysis mentioned above, we learn that the non-empty entries in a discernibility matrix seem to be a key factor for a discernibility matrix based attribute reduction algorithm. In fact, for a compacted table, the number of non-empty entries is closely related to the compaction ratio (the ratio of the number of objects in a compacted decision table to the number of objects in its original version). To better illustrate that the performance of our proposed algorithm evolves with the compaction ratio, we conducted experiments on a synthetic dataset. For the synthetic dataset, the basic data set comprised of 200 objects, and the additional first to fifth parts include 200, 400, 600, 800, and 1000 objects, respectively. After each of these parts was added into the basic dataset of a compacted decision table and its original decision table, their corresponding compaction ratios were 80%, 75%, 70%, 65% and 60%, respectively. From Fig. 5, we can observe that the smaller the compaction ratio, the more significant advantages our proposed incremental attribute reduction algorithms (DMIAR-CDT-P, DMIAR-CDT-S and DMIAR-CDT-C) possess.

5.3 Effectiveness analysis

To verify the effectiveness of our proposed algorithms, we first conducted experiments to compare the reducts obtained by DMIAR-DT-Δ and DMIAR-CDT-Δ on all the data sets in Table 1. Table 6 shows the number of attributes in each reduct derived from each data set through DMIAR-DT-P and DMIAR-CDT-P, DMIAR-DT-S and DMIAR-CDT-S, and DMIAR-DT-C and DMIAR-CDT-C. It can be observed from Table 6 that for each data set in Table 1, the reducts acquired through DMIAR-CDT-Δ are identical to those acquired through DMIAR-DT-Δ, which is consistent with the theoretical results in Sect. 4. Instead of listing each reduct acquired from each data set in Table 1, only the number of attributes in each reduct derived from each data set is shown in Table 6 because of the word count limitations of this paper. In fact, we have compared all reducts derived from each data set and its compacted version in this experiment and determined that the reducts are identical to each other.

6 Conclusion

In this paper, to mitigate the problems associated with the efficiency of the incremental attribute reduction algorithms based on the discernibility matrices of a decision table, we introduced three new types of discernibility matrices for a compacted decision table. We theoretically demonstrated that all the reducts obtained by the algorithms based on these proposed discernibility matrices of a compacted decision table are identical to those based on its original version. Extensive experiments were then conducted to illustrate that the algorithms based on a compacted decision table are much more efficient than those based on its original version while the same reducts can be captured by these two kinds of incremental attribute reduction algorithms. Experiments were also performed to reveal the real reason that enables the proposed algorithms to much more efficiently capture all the reducts of a dynamic data set. In the future, it would be promising to investigate the incremental attribute reduction algorithms for some complex decision information system, such as multi-source information systems [47] and multi-scale information systems [44].

Acknowledgements The research was supported by the National Natural Science Foundation of China (Nos. 61772323, 71301090, 61303008, 61432011 and U1435212) and the National Key Basic Research and Development Program of China (973) (No. 2013CB329404).

References

1. Bang WC, Zeungnam B (1999) New incremental learning algorithm in the framework of rough set theory. Int J Fuzzy Syst 1(1):25–36
2. Chan C (1998) A rough set approach to attribute generalization in data mining. Inf Sci 107(1–4):177–194
3. Chen HM, Li TR, Qiao SJ, Ruan D (2010) A rough set based dynamic maintenance approach for approximations in coarsening and refining attribute values. Int J Intell Syst 25:1005–1026
4. Chen HM, Li TR, Ruan D (2013) A rough-set based incremental approach for updating approximations under dynamic maintenance environments. IEEE Trans Knowl Data Eng 25(2):274–284
5. Deng TQ, Yang CD, Wang XF (2012) A reduct derived from feature selection. Pattern Recogn Lett 33:1638–1646
6. Guan S, Li S (2001) Incremental learning with respect to new incoming input attributes. Neural Process Lett 14(3):241–260
7. Chengxiang Hu, Liu Shixi, Liu Guoxiu (2017) Matrix-based approaches for dynamic updating approximations in multigranulation rough sets. Knowl Based Syst 122:51–63
8. Hu F, Wang GY, Huang H, Wu Y (2005) Incremental attribute reduction based on elementary sets. In: Proceeding of the 10th international conference on rough sets, fuzzy sets, data mining and granular computing, Regina, Canada, in: LNCS, vol 3641, pp 185–193
9. Jing YG, Li TR, Luo C, Horng SJ, Wang GY, Yu Z (2016) An incremental approach for attribute reduction based on knowledge granularity. Knowl Based Syst 104:24–38
10. Jing YG, Li TR, Huang JF, Zhang YY (2016) An incremental attribute reduction approach based on knowledge granularity under the attribute generalization. Int J Approx Reason 76:80–95
11. Jing YG, Li TR, Fujita H, Yu Z, Wang B (2017) An incremental attribute reduction approach based on knowledge granularity with a multi-granulation view. Inf Sci 411:23–38
12. Lang GM, Miao DQ, Cai MJ, Zhang ZF (2017) Incremental approaches for updating reducts in dynamic covering information systems. Knowl Based Syst 134:85–104
13. Lang GM, Cai MJ, Fujita H, Xiao QM (2018) Related families-based attribute reduction of dynamic covering decision information systems. Knowl Based Syst. https://doi.org/10.1016/j.knosys.2018.05.019
14. Li JH, Mei CL, Lv YJ (2011) Knowledge reduction in decision formal contexts. Knowl Based Syst 24:709–715
15. Li JH, Mei CL, Lv YJ (2013) Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction. Int J Approx Reason 54(1):149–165
16. Li JH, Kumar CA, Mei CL, Wang Xizhao (2017) Comparison of reduction in formal decision contexts. Int J Approx Reason 80:100–122
17. Li SY, Li TR, Liu D (2013) Incremental updating approximations in dominance-based rough sets approach under the variation of the attribute set. Knowl Based Syst 40:17–26
18. Li TR, Ruan D, Geert W, Song J, Xu Y (2007) A rough sets based characteristic relation approach for dynamic attribute generalization in data mining. Knowl Based Syst 20(5):485–494
19. Li WT, Xu WH (2015) Double-quantitative decision-theoretic rough set. Inf Sci 316:54–67
20. Liang JY, Chin KS, Dang CY, Yam Richid CM (2002) A new method for measuring uncertainty and fuzziness in rough set theory. Int J Gen Syst 31(4):331–342
21. Liang JY, Shi ZZ, Li DY, Wierman MJ (2006) The information entropy, rough entropy and knowledge granulation in incomplete information table. Int J Gen Syst 35(6):641–654
22. Liang JY, Wang F, Dang CY, Qian YH (2014) A group incremental approach to feature selection applying rough set technique. IEEE Trans Knowl Data Eng 26(2):294–308
23. Liu D, Li TR, Ruan D, Zou WL (2009) An incremental approach for inducing knowledge from dynamic information systems. Fundam Inf 94(2):245–260
24. Liu D, Li TR, Ruan D, Zhang JB (2011) Incremental learning optimization on knowledge discovery in dynamic business intelligent systems. J Glob Optim 51:325–344
25. Liu D, Li TR, Zhang JB (2014) A rough set-based incremental approach for learning knowledge in dynamic incomplete information systems. Int J Approx Reason 55:1764–1786
26. Liu ZT (1999) An incremental arithmetic for the smallest reduction of attributes. Chin J Electron 27(11):96–98
27. Orlowska M, Orlowski M (1992) Maintenance of knowledge in dynamic information systems. In: Slowinski R (ed) Intelligent decision support: handbook of applications and advances of the rough set theory. Kluwer Academic Publishers, South Holland, pp 315–330
28. Pawlak Z (1991) Rough sets: theoretical aspects of reasoning about data. Kluwer Academic Publishers, Boston
29. Pawlak Z, Skowron A (2007) Rudiments of rough sets. Inf Sci 177(1):3–27
30. Pawlak Z, Skowron A (2007) Rough sets and boolean reasoning. Inf Sci 177(1):41–73
31. Qian YH, Liang JY (2008) Combination entropy and combination granulation in rough set theory. Int J Uncertain Fuzziness Knowl Based Syst 16(2):179–193
32. Shan N, Ziarko W (1995) Data-based acquisition and incremental modification of classification rules. Comput Intell 11(2):357–370
33. Shannon CE (1948) The mathematical theory of communication. Bell Syst Tech J 27(3, 4):373–423

34. Shu WH, Shen H (2014) Updating attribute reduction in incomplete decision systems with the variation of attribute set. Int J Approx Reason 55:867–884
35. Shu WH, Shen H (2014) Incremental feature selection based on rough set in dynamic incomplete data. Pattern Recogn 47:3890–3906
36. Skowron A, Rauszer C (1992) The discernibility matrices and functions in information systems. In: Slowinski R (ed) Intelligent decision support: handbook of applications and advances of the rough sets theory. Kluwer, Dordrecht
37. Wang F, Liang JY, Dang CY (2013) Attribute reduction for dynamic data sets. Appl Soft Comput 13:676–689
38. Wang F, Liang JY, Qian YH (2013) Attribute reduction: a dimension incremental strategy. Knowl Based Syst 39:95–108
39. Wang GY, Yu H, Yang DC (2002) Decision table reduction based on conditional information entropy. Chin J Comput 25(7):759–766
40. Wang J, Wang J (2001) Reduction algorithms based on discernibility matrix: the ordered attributes method. J Comput Sci Technol 16(6):489–504
41. Wei W, Liang JY, Qian YH (2012) A comparative study of rough sets for hybrid data. Inf Sci 190:1–16
42. Wei W, Liang JY, Wang JH, Qian YH (2013) Decision-relative discernibility matrixes in the sense of entropies. Int J Gen Syst 42(7):721–738
43. Wei W, Wang JH, Liang JY, Mi X, Dang CY (2015) Compacted decision tables based attribute reduction. Knowl Based Syst 86:261–277
44. Wu WZ, Leung Y (2013) Optimal scale selection for multi-scale decision tables. Int J Approx Reason 54(8):1107–1129
45. Xie XJ, Qin XL (2018) A novel incremental attribute reduction approach for dynamic incomplete decision systems. Int J Approx Reason 93:443–462
46. Xu WH, Guo YT (2016) Generalized multigranulation double-quantitative decision-theoretic rough set. Knowl Based Syst 105:190–205
47. Xu WH, Yu JH (2017) A novel approach to information fusion in multi-source datasets: a granular computing viewpoint. Inf Sci 378:410–423
48. Xu YT, Wang LS, Zhang RY (2011) A dynamic attribute reduction algorithm based on 0–1 integer programming. Knowl Based Syst 24:1341–1347
49. Yang M (2006) An incremental updating algorithm of the computation of a core based on the improved discernibility matrix. Chin J Comput 29(3):407–413
50. Yang XB, Qi Y, Yu HL, Song XN, Yang JY (2014) Updating multigranulation rough approximations with increasing of granular structures. Knowl Based Syst 64:59–69
51. Ju HR, Yang XB, Song XN, Qi YS (2014) Dynamic updating multigranulation fuzzy rough set: approximations and reducts. Int J Mach Learn Cybern 5:981–990
52. Yao YY, Zhao Y (2008) Attribute reduction in decision-theoretic rough set models. Inf Sci 178(17):3356–3373
53. Zhang JB, Li TR, Ruan D, Liu D (2012) Rough sets based matrix approaches with dynamic attribute variation in set-valued information systems. Int J Approx Reason 53:620–635