Dynamic composite decision-theoretic rough set under the change of attributes

Linna Wang [1,3], Xin Yang [2,3]*, Yong Chen [2], Ling Liu [1], Shiyong An [2], Pan Zhuo [2]

[1]*School of Electronic and Information Engineering, Sichuan Technology and Business University, Chengdu, 611745, P.R.China*
E-mail: vanilla-823@163.com, 20548950@qq.com

[2]*Key Laboratory of Cloud Computing and Intelligent Information Processing, Sichuan Technology and Business University, Chengdu, 611745, P.R.China*
E-mail: yangxin2041@163.com, 23688671@qq.com, anshiyong163@163.com, 26229644@qq.com

[3]*Department of Computer Science, University of Regina, Regina, S4S 0A2, Canada*

Abstract

In practical decision-making, we prefer to characterize the uncertain problems with the hybrid data, which consists of various types of data, e.g., categorical data, numerical dada, interval-valued data and set-valued data. The extended rough sets can deal with single type of data based on specific binary relation, including the equivalence relation, neighborhood relation, partial order relation, tolerance relation, etc. However, the fusion of these relations is a significant challenge task in such composite information table. To tackle this issue, this paper proposes the intersection and union composite relation, and further introduces a quantitative composite decision-theoretic rough set model. Subsequently, we present a novel matrix-based approach to compute the upper and lower approximations in proposed model. Moreover, we propose the incremental updating mechanisms and algorithms under the addition and deletion of attributes. Finally, experimental valuations are conducted to illustrate the efficiency of proposed method and algorithms.

Keywords: Composite information table, Decision-theoretic rough set, Quantitative composite relation, Matrix, Incremental updating

1. Introduction

The information table may involve various types of data, e.g., categorical data, numerical data, interval-valued data, set-valued data, etc. As a useful tool to describe the uncertain problems, the theory of rough sets can be utilized to tackle the different types of data by different binary relations. For instance, the traditional rough set model was proposed by Pawlak based on the equivalence relation to address the categorical dada [23]. Hu et al. presented the neighborhood relation to characterize the similarity of two

objects with numerical data [5]. Guan et al. discussed the set-valued information systems with the tolerance relation [3]. Qian et al. proposed the interval ordered information systems for attribute reduction and ordering rules extraction [24]. However, most existing studies focus on single type of data with a simple binary relation under the static information table.

Decision-theoretic rough set (DTRS) is a general probabilistic rough set model [32]. By considering the misclassification cost, DTRS model provides a mathematical interpretation of thresholds based on Bayesian decision procedure. Recently, there are

---

* Corresponding author.

many interesting works in decision-making model [1,22,9,10,31,14,13,17,15,16,40,41,18]. The original DTRS model only can handle categorical data. Recently, many extended DTRS models are proposed to solve different types of data. Li et al. presented a neighborhood based on DTRS model with numerical data and discussed the minimum cost attribute reduction in proposed model [12]. To directly deal with real-valued and interval-valued data, Zhao et al. introduced fuzzy and interval-valued fuzzy DTRS model [39]. Yang et al. studied weighted mean, optimistic and pessimistic multigranulation DTRS in incomplete information table [27]. Qian et al. proposed multigranulation DTRS [25] for the fusion of different relations. However, these studies don't consider the composite information table in DTRS. Actually, it is significant to characterize the objects in practical problem solving with the hybrid data. Moveover, multiple types of data may be changed in dynamic information environment, e.g., the addition and deletion of attributes or objects.

Recently, the incremental updating strategies have been widely researched in rough sets [7,6,19,21,20,30,4,26,38,8]. To efficiently obtain the useful acknowledge under the change of information system, we can propose the incremental updating methods to reduce the computational time in the theory of rough sets. Li et al. introduced the novel model to incrementally update the lower and upper approximations based on the characteristic relation under the change of attributes [11]. Zheng et al. developed a rough set and rule tree based incremental knowledge acquisition algorithm[42]. Yang et al. presented a unified framework of dynamic probabilistic rough sets, which can incrementally update three regions under fifteen situations of change [29], further they proposed a unified model of sequential three-way decisions and multilevel incremental processing [28]. Furthermore, the hybrid data should be considered in real-world applications and it may vary in an information table. Zhang et al. investigated the definition of composite information table and proposed a composite rough set model to deal with the different types of data simultaneously [36]. Then, they further provided a parallel matrix-based approach for computing composite rough set approximations [37]. Chen et al. proposed the distribution attribute reduction method under probabilistic composite rough set [2]. However, the composite relation defined by the intersection operation of relations is too strict for classification problems. In this paper, we define a novel quantitative composite relation w.r.t. multiple types of data. We provide the matrix-based method to compute lower and upper approximations. Furthermore, we propose the incremental approach for updating approximations when the attributes are added or deleted in composite information table. Experiments on four datasets show that the incremental algorithms can efficiently improve the performance of approximations update.

The rest of this paper is organized as follows. Section 2 briefly reviews some basic notions and concepts of DTRS model. In Section 3, we propose a quantitative composite DTRS model based on the quantitative composite relation and further introduce a novel matrix-based approach for the calculation of approximations. Section 4 presents the incremental updating mechanisms and algorithms with composite DTRS model when the attributes are added or deleted in composite information table. Finally, experiments are conducted in Section 5 and Section 6 concludes the paper and elaborates on further works.

2. Decision-theoretic rough sets

Based on well-known Bayesian decision theory, DTRS model provides a mathematical approach for computing thresholds in probabilistic rough sets under the minimum decision risk or cost. In this section, we briefly review the basic concepts and notions of DTRS model [32,34].

Definition 1. Let $SS = (U, A = C \bigcup D, V, f)$ be a single information table, where U is a nonempty finite set of objects; A is also a nonempty finite set, called the attributes of objects, C denotes the condition attributes set, which consists of a single type data, D denotes decision attribute set, $C \bigcap D = \varnothing$; V is a domain of the attributes, $V = \bigcup_{a \in A} V_a$, $f = U \times A \rightarrow V$ is an information function; $f(x_i, a_l)$ denotes the attribute value of object x_i under a_l, $i = 1, 2, \cdots, |U|$, $l = 1, 2, \cdots, |A|$.

In a single information table, the type of data

with respect to the condition attributes is same. The traditional rough sets commonly deal with such data. Given an approximation space (U, R). R is an equivalence relation on U, $B \subseteq C$, $R_B = \{(x,y) \in U \times U | \forall b \in B, f(x,b) = f(y,b)\}$. U/R denotes a partition of U induced by the equivalence relation R. In Pawlak rough sets model, for $X \in U$, the lower and upper approximations can be denoted as:

$$\underline{apr}(X) = \{x \in U \mid [x] \subseteq X\},$$
$$\overline{apr}(X) = \{x \in U \mid [x] \cap X \neq \emptyset\}, \quad (1)$$

where $[x]$ denotes the equivalence class, $[x] = \{y \in U | (x,y) \in R\}$. There are another representation and interpretation of approximations. According to the theory of three-way decisions [33], one of the mainly task is dividing the universe U into three pair-wise disjoint regions. We can obtain positive region (acceptance decisions), boundary region (non-commitment decisions), negative region (rejection decisions) as follows:

$$\text{POS}(X) = \underline{apr}(X),$$
$$\text{BND}(X) = \overline{apr}(X) - \underline{apr}(X), \quad (2)$$
$$\text{NEG}(X) = U - \overline{apr}(X).$$

In order to accept the tolerable errors in rough sets, a series of probabilistic rough sets are proposed in the past two decades. Particularly, DTRS presented by Yao [34] is a general probabilistic model, which can obtain a reasonable pair of thresholds. In the following, the decision procedure of DTRS model is described as follows:

Based on bayesian decisions rules, we usually make the optimized decisions with the minimum risk. For simplicity, we consider two states and three actions in a binary classification problem. The set of states is given by $\Omega = \{X, X^C\}$ indicating that an element is in X and not in X, respectively. Under two opposite states, the set of actions is given by $AC = \{a_P, a_B, a_N\}$, where a_P, a_B, a_N represent the three actions in classifying an object, deciding $POS(X)$, deciding $BND(X)$ and deciding $NEG(X)$. For cost-sensitive learning, the loss function contained six parameters is presented as the matrix $L_{3 \times 2}$ shown in Table 1.

Table 1. The loss function

	$X(P)$	$X^C(N)$
a_P	λ_{PP}	λ_{PN}
a_B	λ_{BP}	λ_{BN}
a_N	λ_{NP}	λ_{NN}

In the matrix, λ_{PP}, λ_{BP} and λ_{NP} denote the losses incurred for taking actions of a_P, a_B and a_N, respectively, when an object belongs to X. Similarly, λ_{PN}, λ_{BN} and λ_{NN} denote the losses incurred for taking the same actions when the object belongs to X^C. The expected losses associated with taking the three actions can be expressed as:

$$R(a_P|[x]) = \lambda_{PP} Pr(X|[x]) + \lambda_{PN} Pr(X^C|[x]),$$
$$R(a_B|[x]) = \lambda_{BP} Pr(X|[x]) + \lambda_{BN} Pr(X^C|[x]), \quad (3)$$
$$R(a_N|[x]) = \lambda_{NP} Pr(X|[x]) + \lambda_{NN} Pr(X^C|[x]),$$

where $Pr(X|[x]) = \frac{|X \cap [x]|}{|[x]|}$ is the condition probability, $|*|$ denotes the cardinality of a set.

Based on minimum-cost decision rules in the Bayesian decision procedure, the (α, β) lower and upper approximations of DTRS model can be defined by:

$$\underline{R}_{(\alpha, \beta)}(X) = \{x \in U \mid Pr(X|[x]) \geqslant \alpha\},$$
$$\overline{R}_{(\alpha, \beta)}(X) = \{x \in U \mid Pr(X|[x]) > \beta\}, \quad (4)$$

where the parameters α, β satisfied $0 \leqslant \beta < \alpha \leqslant 1$ can be calculated as:

$$\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})},$$
$$\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}. \quad (5)$$

The detailed derivation of thresholds can reference literature [32,34]. The original three-way decisions derive from DTRS model [33]. Compared with two way decisions, we consider the delay option in three-way decisions when the information is insufficient for supporting definite decisions. The (α, β) probabilistic three regions are given as follows:

$$\text{POS}(X) = \{x \in U \mid Pr(X|[x]) \geqslant \alpha\},$$
$$\text{BND}(X) = \{x \in U \mid \beta < Pr(X|[x]) < \alpha\}, \quad (6)$$
$$\text{NEG}(X) = \{x \in U \mid Pr(X|[x]) \leqslant \beta\}.$$

Three regions, namely, positive region, boundary region, negative region, provide three direction of decisions: the first is the acceptance decisions when $Pr(X|[x]) \geqslant \alpha$; the second is the non-commitment decisions when $\beta \leqslant Pr(X|[x]) \leqslant \alpha$; the third is the rejection decisions when $Pr(X|[x]) \leqslant \beta$.

3. Composite decision-theoretic rough sets and matrix representation of approximations

3.1. Quantitative composite DTRS model

In general, the previous DTRS model can deal with the single data type in a single information table. However, hybrid data types usually appear in a composite information table, which contains categorical data, numerical data , set-valued data, interval-valued data, etc. We can utilize different binary relations to handle different types of data. In what follows, we introduce the definition of composite information table, and present the composite binary relation by fusion strategy in DTRS model.

Definition 2. Let $CS = (U, A = C\bigcup D, V, f)$ be a composite information table, where U is a nonempty finite set of objects; A is also a nonempty finite set, called the attributes of objects, C denotes the condition attributes set consisted of hybrid type data, $C = \bigcup C_k, k = 1, 2, \ldots, m$, where C_k is a subset of C with the same data type and m denotes the number of data types, D denotes decision attribute set, $C \bigcap D = \varnothing$; V is a domain of the attributes, $V = \bigcup_{a \in A} V_a$, $f = U \times A \to V$ is an information function; $f(x_i, a_l)$ denotes the attribute value of object x_i under $a_l, i = 1, 2, \cdots, |U|, l = 1, 2, \cdots, |A|$.

Example 1. In Table 2, there are four types of data in a composite information table $CS = (U, A = C\bigcup D, V, f)$. Let $B = C = \cup B_k, k = 1, 2, 3, 4$, where $B_1 = \{b_{categorical}\}, B_2 = \{b_{numerical}\}, B_3 = \{b_{interval-valued}\}, B_4 = \{b_{set-valued}\}, D = \{d_{categorical}\}$, denote categorical data, numerical data, interval-valued data, set-valued data, categorical data, respectively.

For classification and decisions in such information table, the key issue is the fusion of different binary relations. Based on existing studies [36], we in-

troduce three approaches to define the composite relation, namely, the intersect composite relation, the union composite relation, and the quantitative composite relation.

Definition 3. Let $CS = (U, A = C\bigcup D, V, f)$ be a composite information table, where $U = \{x_1, x_2, \ldots, x_n\}$, X be a subset of U. $B = \cup B_k \subseteq C$, $B_k \subseteq C_k$. The characteristic matrix $E_{n \times 1}^X = [e_1, e_2, \ldots, e_n]^T$ is defined as follows:

$$e_i = \begin{cases} 1, & x_i \in X; \\ 0, & x_i \notin X. \end{cases} \tag{7}$$

where $i = 1, 2, \ldots, n$ and T denotes the transpose operation.

Definition 4. Let $CS = (U, A = C\bigcup D, V, f)$ be a composite information table, where $U = \{x_1, x_2, \ldots, x_n\}$, $B = \cup B_k \subseteq C$, $B_k \subseteq C_k$. The relation matrix $M_{R_{B_k}} = (m_{ij})_{n \times n}$ can be defined as follows:

$$M_{R_{B_k}} = (m_{ij})_{n \times n} = \begin{cases} 1, & (x_i, x_j) \in R_{B_k}; \\ 0, & (x_i, x_j) \notin R_{B_k}. \end{cases} \tag{8}$$

Definition 5. [36] (The intersect composite relation) Given a composite information table $CS = (U, A = C\bigcup D, V, f)$. Let $x, y \in U$ and $B = \cup B_k \subseteq C, B_k \subseteq C_k$, the intersect composite relation $CR_B^{\cap}$ is defined as:

$$CR_B^{\cap} = \left\{ (x, y) \mid (x, y) \in \bigcap_{B_k \subseteq B} R_{B_k} \right\}, \tag{9}$$

where $R_{B_k} \subseteq U \times U$ is a binary relation defined for one type of data on the attribute set B_k. For simplicity, we give a intuitive presentation as follows:

$$CR_B^{\cap} = R_{B_1} \wedge R_{B_2} \wedge \ldots \wedge R_{B_m}. \tag{10}$$

Furthermore, the relation matrix $M_{CR_B^{\cap}} = (m_{ij}^{CR_B^{\cap}})_{n \times n}$ based on the intersect composite relation $CR_B^{\cap}$ can be given by:

$$M_{CR_B^{\cap}} = M_{R_{B_1}} \wedge M_{R_{B_2}} \wedge \ldots \wedge M_{R_{B_m}}, \tag{11}$$

where $m_{ij}^{CR_B^{\cap}} = m_{ij}^{R_{B_1}} \wedge m_{ij}^{R_{B_2}} \wedge \ldots \wedge m_{ij}^{R_{B_m}}$.

Similarly, we can define the union composite relation as follows.

Table 2. A composite information table

U	$b_{categorical}$	$b_{numerical}$	$b_{interval-valued}$	$b_{set-valued}$	$d_{categorical}$
x_1	1	0.4	[1.2 1.5]	{1, 2}	1
x_2	1	0.2	[1.4 1.8]	{0, 1, 2}	1
x_3	2	0.2	[1.5 2.2]	{1}	0
x_4	1	0.5	[1.6 2.1]	{1, 3 }	1
x_5	2	0.2	[1.4 1.7]	{1, 2, 3}	0
x_6	1	0.4	[2.2 2.5]	{3}	1

Definition 6. (The union composite relation) Given a composite information table $CS = (U, A = C \bigcup D, V, f)$. Let $x, y \in U$ and $B = \cup B_k \subseteq C, B_k \subseteq C_k$, the union composite relation $CR_B^{\cup}$ is defined as:

$$CR_B^{\cup} = \left\{ (x,y) \mid (x,y) \in \bigcup_{B_k \subseteq B} R_{B_k} \right\}, \quad (12)$$

where $R_{B_k} \subseteq U \times U$ is a binary relation defined for one type of data on the attribute set B_k. For simplicity, we give a intuitive presentation as follows:

$$CR_B^{\cup} = R_{B_1} \vee R_{B_2} \vee \dots \vee R_{B_m}. \quad (13)$$

Furthermore, the relation matrix $M_{CR_B^{\cup}} = (m_{ij}^{CR_B^{\cup}})_{n \times n}$ based on the union composite relation $CR_B^{\cup}$ can be given by:

$$M_{CR_B^{\cup}} = M_{R_{B_1}} \vee M_{R_{B_2}} \vee \dots \vee M_{R_{B_m}}, \quad (14)$$

where $m_{ij}^{CR_B^{\cup}} = m_{ij}^{R_{B_1}} \vee m_{ij}^{R_{B_2}} \vee \dots \vee m_{ij}^{R_{B_m}}$.

In a composite information table, we present the intersection composite relation $CR_B^{\cap}$ and the union composite relation $CR_B^{\cup}$. The former is the strict relation which needs to satisfy every binary relation with respect to hybrid type data between two objects. It leads to obtain the finer granules. The latter is the relaxed relation which needs to satisfy at least one binary relation with respect to hybrid type data between two objects. Conversely, it leads to the coarser granules. Obviously, two are not our best choice. Therefore, we propose the quantitative composite relation as follows.

Definition 7. (The quantitative composite relation) Given a composite information table $CS = (U, A =$

$C \bigcup D, V, f)$. Let $x, y \in U$ and $B = \cup B_k \subseteq C, B_k \subseteq C_k, k = 1, 2, \dots, m$. Suppose threshold θ satisfied $0 \leqslant \theta < 1$, the quantitative composite relation QCR_B is defined as:

$$QCR_B = \left\{ (x,y) \mid \frac{|\{R_{B_k} : (x,y) \in R_{B_k}\}|}{m} > \theta \right\}, \quad (15)$$

where m is the total number of binary relations, $|\{R_{B_k} : (x,y) \in R_{B_k}\}|$ is the number of the satisfied binary relations. The quantitative composite relation matrix $M_{QCR_B} = (m_{ij}^{QCR_B})_{n \times n}$ can be given by:

$$m_{ij}^{QCR_B} = \begin{cases} 1, & \dfrac{\sum_{k=1}^{m} M_{R_{B_k}}}{m} > \theta; \\ 0, & \text{else}. \end{cases} \quad (16)$$

The quantitative composite relation QCR_B is reflexive, but not symmetric and transitive. The intersection composite relation and the union composite relation can drive from the quantitative composite relation. When $\frac{m-1}{m} \leqslant \theta < 1$, the quantitative composite relation becomes the intersection composite relation; when $0 \leqslant \theta < \frac{1}{m}$, the quantitative composite relation becomes the union composite relation. Hence, we can control the threshold θ to establish different model to deal with the hybrid data.

Definition 8. Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information table, QCR_B be a quantitative composite relation. Suppose $X \subseteq U$, the (α, β) lower and upper approximations of concept X in composite DTRS model can be defined by:

$$\underline{R}_{(\alpha, \beta)}^{QCR_B}(X) = \{x \in U \mid Pr(X|QCR_B(x)) \geqslant \alpha\},$$
$$\overline{R}_{(\alpha, \beta)}^{QCR_B}(X) = \{x \in U \mid Pr(X|QCR_B(x)) > \beta\}. \quad (17)$$

The (α,β) probabilistic three regions are given as follows:

$$POS(X) = \{x \in U \mid Pr(X|QCR_B(x)) \geqslant \alpha\},$$
$$BND(X) = \{x \in U \mid \beta < Pr(X|QCR_B(x)) < \alpha\}, (18)$$
$$NEG(X) = \{x \in U \mid Pr(X|QCR_B(x)) \leqslant \beta\}.$$

Definition 9. Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information table, QCR_B be a quantitative composite relation. Let $U/D = \{D_1, D_2, \ldots, D_i, \ldots, D_s\}$ be a partition of U based on the decision attribute d. Then the (α,β) lower and upper approximations of the decision class D_i in composite DTRS model can be defined by:

$$\underline{R}^{QCR_B}_{(\alpha,\beta)}(D_i) = \{x \in U \mid Pr(D_i|QCR_B(x)) \geqslant \alpha\},$$
$$\overline{R}^{QCR_B}_{(\alpha,\beta)}(D_i) = \{x \in U \mid Pr(D_i|QCR_B(x)) > \beta\}. \quad (19)$$

Example 2. (Continuation of Example 1) In table 2, to deal with four types of data, four binary relations, namely, the equivalence relation, the neighborhood relation, the partial order relation, and the tolerance relation are given as follows:

1. The equivalence relation (see definition in Section 2)

2. The neighborhood relation

$$R^N_{B_k} = \left\{ (x,y) \mid y \in U, b^i_k \in B_k, \Delta_B(x,y) \leqslant \delta \right\}, (20)$$

where

$$\Delta_B(x,y) = \left(\sum_{i=1}^{|B_k|} |f(x, b^i_k) - f(y, b^i_k)|^2 \right)^{1/2}.$$

3. The partial order relation

$$R^P_{B_k} = \{(x,y) \mid y \in U, b^i_k \in B_k, f^L(y, b^i_k) \geqslant f^L(x, b^i_k), f^U(y, b^i_k) \geqslant f^U(x, b^i_k)\}. \quad (21)$$

4. The tolerance relation

$$R^T_{B_k} = \{(x,y) \mid y \in U, \forall b^i_k \in B_k, f(x, b^i_k) \cap f(y, b^i_k) \neq \emptyset\}. \quad (22)$$

Suppose neighborhood threshold $\delta = 0.1$. According to four definitions of binary relations, we can calculate four relation matrices with respect to four types of data respectively on U as follows:

$$M_{R^E_{B_1}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix},$$

$$M_{R^N_{B_2}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix},$$

$$M_{R^P_{B_3}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$M_{R^T_{B_4}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

According to Definition 5 and Definition 6, the intersection composite relation matrix and the union composite relation matrix can be calculated as:

$$M_{CR^\cap_B} = M_{R^E_{B_1}} \wedge M_{R^N_{B_2}} \wedge M_{R^P_{B_3}} \wedge M_{R^T_{B_4}}$$
$$= \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$M_{CR_B^{\cup}} = M_{R_{B_1}^E} \vee M_{R_{B_2}^N} \vee M_{R_{B_3}^P} \vee M_{R_{B_4}^T}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Suppose the quantitative threshold $\theta = 0.7$, then the quantitative composite relation matrix can be calculated as:

$$M_{QCR_B} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

3.2. Matrix representation of approximations

In order to propose a fast matrix-based approach for updating approximations under the change of attributes, another representation of approximations [35] in DTRS model is given as follows.

Definition 10. Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information table, QCR_B be a quantitative composite relation. Suppose $X \subseteq U$, the (α', β') lower and upper approximations of concept X in composite DTRS model can be defined by:

$$\underline{R}_{(\alpha, \beta)}^{QCR_B}(X) = \left\{ x \in U \mid \frac{|X \cap QCR_B(x)|}{|X^C \cap QCR_B(x)|} \geqslant \alpha' \right\},$$

$$\overline{R}_{(\alpha, \beta)}^{QCR_B}(X) = \left\{ x \in U \mid \frac{|X \cap QCR_B(x)|}{|X^C \cap QCR_B(x)|} > \beta' \right\}, \quad (23)$$

where a pair of thresholds (α, β) can be calculated as:

$$\alpha' = \frac{\alpha}{1-\alpha} = \frac{\lambda_{PN} - \lambda_{BN}}{\lambda_{BP} - \lambda_{PP}},$$

$$\beta' = \frac{\beta}{1-\beta} = \frac{\lambda_{BN} - \lambda_{NN}}{\lambda_{NP} - \lambda_{BP}}. \quad (24)$$

When $|X^C \bigcap QCR_B(x)| = 0$, we define that the value of $\frac{|X \bigcap QCR_B(x)|}{|X^C \bigcap QCR_B(x)|}$ equals to infinity, denoted as the symbol ∞.

Definition 11. Let $M_{n \times 1}$ be a matrix. The cut matrices of $M^{\downarrow}$ and $M^{\uparrow}$ are defined respectively as:

$$M^{\downarrow} = \begin{cases} 1, & M(i) \geqslant \alpha'; \\ 0, & else. \end{cases} \quad (25)$$

$$M^{\uparrow} = \begin{cases} 1, & M(i) > \beta'; \\ 0, & else. \end{cases} \quad (26)$$

where $i = 1, 2, \ldots, n$.

Definition 12. Let $CS = (U, AT = C \bigcup D, V, f)$ be a composite information table, QCR_B be a quantitative composite relation, where $U = \{x_1, x_2, \ldots, x_n\}$. $E_{n \times n}^X$ is the characteristic matrix, M_{QCR_B} is the quantitative composite relation matrix. Then the characteristic matrices of lower and upper approximations in composite DTRS model can be defined respectively as follows:

$$E_{\underline{R}_{(\alpha, \beta)}^{QCR_B}(X)} = ((M_{QCR_B} * E)./(M_{QCR_B} * (\sim E)))^{\downarrow},$$

$$E_{\overline{R}_{(\alpha, \beta)}^{QCR_B}(X)} = ((M_{QCR_B} * E)./(M_{QCR_B} * (\sim E)))^{\uparrow}. \quad (27)$$

The matrix-based algorithm for computing approximations in composite DTRS model is outlined in Algorithm 1. Step 1 is to construct the characteristic matrix, whose time complexity is $O(|U|)$. Step 2 is to compute the equivalence relation matrix, the neighborhood relation matrix, the partial order relation matrix, and the tolerance relation matrix, whose time complexity is $O(|U||B_1|)$, $O(|U||log|U|B_2|)$, $O(|U||B_3|^2)$, $O(|U||B_4|^2)$. Step 3 is to compute the quantitative composite relation matrix, whose time complexity is $O(|U|^2)$. Step 4 is to compute the intersection matrix and non-intersection matrix, whose time complexity is $O(|U|^2)$. Step 5 and Step 6 are to compute the characteristic matrix of lower and upper approximations, whose time complexity is $O(|U|)$. In the following, we illustrate the proposed matrix-based method with an numerical example.

Algorithm 1 (NCDTRS) The matrix-based algorithm for computing approximations in composite DTRS model.

Input: A composite information table $CS = (U, A = C \bigcup D, V, f)$, the concept $X (X \subseteq U)$, and the loss function $L_{3 \times 2}$.

Output: The characteristic matrices of lower and upper approximations $E_{\underline{R}^{QCR_B}_{(\alpha, \beta)}(X)}, E_{\overline{R}^{QCR_B}_{(\alpha, \beta)}(X)}$.

Step 1: Construct the characteristic matrix $E^X_{n \times 1}$ with respect to the concept X;

Step 2: Compute the equivalence relation matrix $M_{R^E_{B_2}}$, the neighborhood relation matrix $M_{R^N_{B_2}}$, the partial order relation matrix $M_{R^P_{B_2}}$, and the tolerance relation matrix $M_{R^T_{B_2}}$, respectively;

Step 3: Compute the quantitative composite relation matrix M_{QCR_B} according to Definition 7;

Step 4: Compute the intersection matrix $W = (w_i)_{n \times 1} = M_{QCR_B} * E$ and the non-intersection matrix $W' = (w'_i)_{n \times 1} = M_{QCR_B} * (\sim E)$

Step 5: Compute the characteristic matrix of lower approximations $E_{\underline{R}^{QCR_B}_{(\alpha, \beta)}}(X) = (W./W')^{\downarrow}$;

Step 6: Compute the characteristic matrix of upper approximations $E_{\overline{R}^{QCR_B}_{(\alpha, \beta)}}(X) = (W./W')^{\uparrow}$;

Example 3. (Continuation of Example 2) Suppose the loss function $L = [0\ 6; 1\ 3; 5\ 0]$. Then we have $\alpha' = 3, \beta' = 0.75$. The lower approximation of composite DTRS model based on the matrix approach can be calculated as:

$$E_{\underline{R}^{QCR_B}_{(\alpha, \beta)}(X)} = ((M_{QCR_B} * E)./(M_{QCR_B} * (\sim E)))^{\downarrow}$$

$$= \left\{ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} ./ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}^{\downarrow}$$

$$= \left\{ \begin{bmatrix} 4 \\ 2 \\ 0 \\ 3 \\ 1 \\ 2 \end{bmatrix} ./ \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 2 \\ 0 \end{bmatrix} \right\}^{\downarrow} = \left\{ \begin{bmatrix} \infty \\ 2 \\ 0 \\ \infty \\ 0.5 \\ \infty \end{bmatrix} \right\}^{\downarrow} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Similarity, the upper approximation of composite DTRS model based on the matrix approach can be calculated as:

$$E_{\overline{R}^{QCR_B}_{(\alpha, \beta)}(X)} = ((M_{QCR_B} * E)./(M_{QCR_B} * (\sim E)))^{\uparrow},$$

$$= \left\{ \begin{bmatrix} \infty \\ 2 \\ 0 \\ \infty \\ 0.5 \\ \infty \end{bmatrix} \right\}^{\uparrow} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Then we can obtain the lower and upper approximations of composite DTRS model as follows:

$$E_{\underline{R}^{QCR_B}_{(\alpha, \beta)}(X)} = \{x_1, x_4, x_6\},$$
$$E_{\overline{R}^{QCR_B}_{(\alpha, \beta)}(X)} = \{x_1, x_2, x_4, x_6\}.$$

4. The approach for incremental updating approximations under the change of attributes

In an dynamic composite information table, the change of attributes have two situations. One is the addition of attributes and another is the deletion of attributes. To achieve a fast calculate process, incremental updating the quantitative composite relation matrix and the characteristic matrices of approximations are two importance tasks. In this section, we introduce the incremental updating of matrix-based strategy in composite DTRS model.

4.1. updating the quantitative composite relation matrix and the characteristic matrices of approximations when adding attributes

In this subsection, the incremental update of approximations is considered as from time t to $t+1$. Let $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$

be the composite information table at time t, where $U = \{x_1, x_2, \ldots, x_n\}$. The composite information table $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$ at time t only contains four types of data, where $C^t = B_1^t \bigcup B_2^t \bigcup B_3^t \bigcup B_4^t$, B_1^t denotes categorical data, B_2^t denotes numerical data, B_3^t denotes set-valued data, B_4^t denotes interval-valued data. Let $CS^{t+1} = (U^{t+1}, A^{t+1} = C^{t+1} \bigcup D^{t+1}, V^{t+1}, f^{t+1})$ be the composite information table at time $t+1$, where $U = \{x_1, x_2, \ldots, x_n\}$. At time $t+1$, a set of attributes ΔC are added into $CS^{t+1} = (U^{t+1}, A^{t+1} = C^{t+1} \bigcup D^{t+1}, V^{t+1}, f^{t+1})$, $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t \bigcup \Delta C$.

From time t to $t+1$, there are two steps to update the quantitative composite relation matrix. Step 1 is updating the equivalence relation matrix $M_{R_{B_2}^E}$, the neighborhood relation matrix $M_{R_{B_2}^N}$, the partial order relation matrix $M_{R_{B_2}^P}$, and the tolerance relation matrix $M_{R_{B_2}^T}$, respectively. Step 2 is updating the quantitative composite relation matrix.

Proposition 1. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R_{B_1}^E}$ *is the equivalence relation matrix on U. Suppose* $\Delta C = \Delta B_1$, $C^{t+1} = C^t \bigcup \Delta C$, *the equivalence relation matrix* $M_{R_{B_1}^E}$ *by adding* ΔC *to C from time t to t + 1 can be updated as:*

$$(m_{ij}^{R_{B_1}})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \vee (x_i, x_j)^{t+1} \notin R_{\Delta C}^E; \\ 1, & \text{else.} \end{cases} \quad (28)$$

Proposition 2. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R_{B_2}^N}$ *is the neighborhood relation matrix on U. Suppose* $\Delta C = \Delta B_2$, $C^{t+1} = C^t \bigcup \Delta C$, *the neighborhood relation matrix* $M_{R_{B_2}^N}$ *by adding* ΔC *to C from time t to t + 1 can be updated as:*

$$(m_{ij}^{R_{B_2}})_{n \times n}^{t+1} = \begin{cases} 1, & m_{ij}^t = 1 \wedge (x_i, x_j)^{t+1} \in R_{C^{t+1}}^N; \\ 0, & \text{else.} \end{cases} \quad (29)$$

Proposition 3. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U =$

$\{x_1, x_2, \ldots, x_n\}$. $M_{R_{B_3}^P}$ *is the interval-valued relation matrix on U. Suppose* $\Delta C = \Delta B_3$, $C^{t+1} = C^t \bigcup \Delta C$, *the interval-valued relation matrix* $M_{R_{B_3}^P}$ *by adding* ΔC *to C from time t to t + 1 can be updated as:*

$$(m_{ij}^{R_{B_3}})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \vee (x_i, x_j)^{t+1} \notin R_{\Delta C}^P; \\ 1, & \text{else.} \end{cases} \quad (30)$$

Proposition 4. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R_{B_4}^T}$ *is the set-valued relation matrix on U. Suppose* $\Delta C = \Delta B_4$, $C^{t+1} = C^t \bigcup \Delta C$, *the set-valued relation matrix* $M_{R_{B_4}^T}$ *by adding* ΔC *to C from time t to t + 1 can be updated as:*

$$(m_{ij}^{R_{B_4}})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \vee (x_i, x_j)^{t+1} \notin R_{\Delta C}^T; \\ 1, & \text{else.} \end{cases} \quad (31)$$

Proposition 5. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U = \{x_1, x_2, \ldots, x_n\}$. M_{QRC_B} *is the quantitative composite relation matrix on U. Suppose* $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t \bigcup \Delta C$, *the quantitative composite relation matrix* M_{QRC_B} *by adding* ΔC *to C from time t to t + 1 can be updated as:*

$$(m_{ij}^{QRC_B})_{n \times n}^{t+1} = \begin{cases} 1, & m_{ij}^t = 1 \wedge (x_i, x_j)^{t+1} \in QCR_{C^{t+1}}; \\ 0, & \text{else.} \end{cases} \quad (32)$$

Proposition 6. *Let* $CS = (U, A = C \bigcup D, V, f)$ *be a composite information system, where* $U = \{x_1, x_2, \ldots, x_n\}$. M_{QRC_B} *is the quantitative composite relation matrix on U.* $W = (w_i)_{n \times 1} = M_{QCR_B} * E$ *is the intersection matrix, and* $W' = (w_i')_{n \times 1} = M_{QCR_B} * (\sim E)$ *is the non-intersection matrix,* $i = 1, 2, \ldots, n$. *Suppose* $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t \bigcup \Delta C$. *The intersection matrix* $W = (w_i)_{n \times 1} = M_{QCR_B} * E$ *can be updated as:*

$$(w_i)_{n \times 1}^{t+1} = \begin{cases} w_i^t - 1, & m_{ij}^t = 1 \wedge m_{ij}^{t+1} = 0 \wedge e_j^t = 1; \\ w_i^t, & \text{else.} \end{cases} \quad (33)$$

Similarity, the non-intersection matrix $W' = (w'_i)_{n \times 1} = M_{QCR_B} * (\sim E)$ can be updated as the same way.

The incremental algorithm for computing approximations in composite DTRS model based on matrix when adding attributes is outlined in Algorithm 2. Step 1 is to update four relation matrices according to Proposition 1-4, whose time complexity is $O(|U|^2)$. Step 2 is to updating quantitative relation matrix according to Proposition 5, whose time complexity is $O(|U|^2)$. Step 3 is to update W and W' according to Proposition 6, whose time complexity is $O(|U|^2)$. Step 4 and Step 5 are to compute the characteristic matrices of lower and upper approximations, whose time complexity is $O(|U|)$.

Algorithm 2 (ICDTRS-AA) The incremental algorithm for computing approximations in composite DTRS model based on matrix when adding attributes.

Input: At time t, a composite information table $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$, the loss function $L_{3 \times 2}$, the characteristic matrix $E^X_{n \times 1}$, the equivalence relation matrix $M_{R^E_{B_2}}$, the neighborhood relation matrix $M_{R^N_{B_2}}$, the partial order relation matrix $M_{R^P_{B_2}}$, and the tolerance relation matrix $M_{R^T_{B_2}}$, the quantitative composite relation matrix M_{QCR_B}, the intersection matrix W, the non-intersection matrix W'. From time t to $t+1$, ΔC is the addition of attributes, $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t \bigcup \Delta C$.

Output: The characteristic matrices of lower and upper approximations $E^{t+1}_{\underline{R}^{QCR_B}_{(\alpha, \beta)}(X)}, E^{t+1}_{\overline{R}^{QCR_B}_{(\alpha, \beta)}(X)}$ at time $t+1$.

Step 1: Updating $M_{R^E_{B_2}}, M_{R^N_{B_2}}, M_{R^P_{B_2}}, M_{R^T_{B_2}}$ according to Proposition 1-4;

Step 2: Updating M_{QCR_B} according to Proposition 5;

Step 3: Updating W and W' according to Proposition 6;

Step 4: Compute the characteristic matrix of lower approximations $E_{\underline{R}^{QCR_B}_{(\alpha, \beta)}}(X) = (W./W')^{\downarrow}$;

Step 5: Compute the characteristic matrix of upper approximations $E_{\overline{R}^{QCR_B}_{(\alpha, \beta)}}(X) = (W./W')^{\uparrow}$;

4.2. updating the quantitative composite relation matrix and the characteristic matrices of approximations when deleting attributes

In this subsection, the incremental update of approximations is considered as from time t

to $t+1$. Let $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$ be the composite information table at time t, where $U = \{x_1, x_2, \ldots, x_n\}$. The composite information table $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$ at time t only contains four types of data, where $C^t = B^t_1 \bigcup B^t_2 \bigcup B^t_3 \bigcup B^t_4$, B^t_1 denotes categorical data, B^t_2 denotes numerical data, B^t_3 denotes set-valued data, B^t_4 denotes interval-valued data. Let $CS^{t+1} = (U^{t+1}, A^{t+1} = C^{t+1} \bigcup D^{t+1}, V^{t+1}, f^{t+1})$ be the composite information table at time $t+1$, where $U = \{x_1, x_2, \ldots, x_n\}$. At time $t+1$, a set of attributes ΔC are deleted from $CS^{t+1} = (U^{t+1}, A^{t+1} = C^{t+1} \bigcup D^{t+1}, V^{t+1}, f^{t+1})$, $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t - \Delta C$.

Proposition 7. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R^E_{B_1}}$ is the equivalence relation matrix on U. Suppose $\Delta C = \Delta B_1$, $C^{t+1} = C^t - \Delta C$, the equivalence relation matrix $M_{R^E_{B_1}}$ from time t to $t+1$ can be updated as:*

$$(m^{R_{B_1}}_{ij})^{t+1}_{n \times n} = \begin{cases} 0, & m^t_{ij} = 0 \wedge (x_i, x_j)^{t+1} \notin R^E_{\Delta C}; \\ 1, & \text{else.} \end{cases} \tag{34}$$

Proposition 8. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R^N_{B_2}}$ is the neighborhood relation matrix on U. Suppose $\Delta C = \Delta B_2$, $C^{t+1} = C^t - \Delta C$, the neighborhood relation matrix $M_{R^N_{B_2}}$ by adding ΔC to C from time t to $t+1$ can be updated as:*

$$(m^{R_{B_2}}_{ij})^{t+1}_{n \times n} = \begin{cases} 0, & m^t_{ij} = 0 \vee (x_i, x_j)^{t+1} \in R^N_{C^{t+1}}; \\ 1, & \text{else.} \end{cases} \tag{35}$$

Proposition 9. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R^P_{B_3}}$ is the interval-valued relation matrix on U. Suppose $\Delta C = \Delta B_3$, $C^{t+1} = C^t - \Delta C$, the interval-valued relation matrix $M_{R^P_{B_3}}$ by adding*

ΔC to C from time t to $t+1$ can be updated as:

$$(m_{ij}^{R_{B_3}})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \land (x_i, x_j)^{t+1} \notin R_{\Delta C}^P; \\ 1, & \text{else.} \end{cases} \tag{36}$$

Proposition 10. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. $M_{R_{B_4}^T}$ is the set-valued relation matrix on U. Suppose $\Delta C = \Delta B_4$, $C^{t+1} = C^t - \Delta C$, the set-valued relation matrix $M_{R_{B_4}^T}$ by adding ΔC to C from time t to $t+1$ can be updated as:*

$$(m_{ij}^{R_{B_4}})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \land (x_i, x_j)^{t+1} \notin R_{\Delta C}^T; \\ 1, & \text{else.} \end{cases} \tag{37}$$

Proposition 11. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. M_{QRC_B} is the quantitative composite relation matrix on U. Suppose $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t - \Delta C$, the quantitative composite relation matrix M_{QRC_B} by adding ΔC to C from time t to $t+1$ can be updated as:*

$$(m_{ij}^{QRC_B})_{n \times n}^{t+1} = \begin{cases} 0, & m_{ij}^t = 0 \land (x_i, x_j)^{t+1} \notin QCR_{C^{t+1}}; \\ 1, & \text{else.} \end{cases} \tag{38}$$

Proposition 12. *Let $CS = (U, A = C \bigcup D, V, f)$ be a composite information system, where $U = \{x_1, x_2, \ldots, x_n\}$. M_{QRC_B} is the quantitative composite relation matrix on U. $W = (w_i)_{n \times 1} = M_{QRC_B} * E$ is the intersection matrix, and $W' = (w_i')_{n \times 1} = M_{QRC_B} * (\sim E)$ is the non-intersection matrix, $i = 1, 2, \ldots, n$. Suppose $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t - \Delta C$. The intersection matrix $W = (w_i)_{n \times 1} = M_{QRC_B} * E$ can be updated as:*

$$(w_i)_{n \times 1}^{t+1} = \begin{cases} w_i^t - 1, & m_{ij}^t = 1 \land m_{ij}^{t+1} = 0 \land e_j^t = 1; \\ w_i^t, & \text{else.} \end{cases} \tag{39}$$

Similarity, the non-intersection matrix $W' = (w_i')_{n \times 1} = M_{QRC_B} * (\sim E)$ can be updated as the same way.

The incremental algorithm for computing approximations in composite DTRS model based on matrix when deleting attributes is outlined in Algorithm 3. Step 1 is to update four relation matrices according to Proposition 5-8, whose time complexity is $O(|U|^2)$. Step 2 is to updating quantitative relation matrix according to Proposition 9, whose time complexity is $O(|U|^2)$. Step 3 is to update W and W' according to Proposition 10, whose time complexity is $O(|U|^2)$. Step 4 and Step 5 are to compute the characteristic matrices of lower and upper approximations, whose time complexity is $O(|U|)$.

Algorithm 3 (ICDTRS-DA) The incremental algorithm for computing approximations in composite DTRS model based on matrix when deleting attributes.

Input: At time t, a composite information table $CS^t = (U^t, A^t = C^t \bigcup D^t, V^t, f^t)$, the loss function $L_{3 \times 2}$, the characteristic matrix $E_{n \times 1}^X$, the equivalence relation matrix $M_{R_{B_2}^E}$, the neighborhood relation matrix $M_{R_{B_2}^N}$, the partial order relation matrix $M_{R_{B_2}^P}$, and the tolerance relation matrix $M_{R_{B_2}^T}$, the quantitative composite relation matrix M_{QCR_B}, the intersection matrix W, the non-intersection matrix W'. From time t to $t+1$, ΔC is the addition of attributes, $\Delta C = \Delta B_1 \bigcup \Delta B_2 \bigcup \Delta B_3 \bigcup \Delta B_4$, $C^{t+1} = C^t - \Delta C$.

Output: The characteristic matrices of lower and upper approximations $E_{\underline{R}_{(\alpha, \beta)}^{QCR_B}(X)}^{t+1}, E_{\overline{R}_{(\alpha, \beta)}^{QCR_B}(X)}^{t+1}$ at time $t+1$.

Step 1: Updating $M_{R_{B_2}^E}, M_{R_{B_2}^N}, M_{R_{B_2}^P}, M_{R_{B_2}^T}$ according to Proposition 5-8;

Step 2: Updating M_{QCR_B} according to Proposition 9;

Step 3: Updating W and W' according to Proposition 10;

Step 4: Compute the characteristic matrix of lower approximations $E_{\underline{R}_{(\alpha, \beta)}^{QCR_B}}(X) = (W./W')^{\downarrow}$;

Step 5: Compute the characteristic matrix of upper approximations $E_{\overline{R}_{(\alpha, \beta)}^{QCR_B}}(X) = (W./W')^{\uparrow}$;

---

5. Experimental evaluations

In this section, we conduct the comparative experiments to verify the performance of the proposed algorithms for incremental updating approximations in composite DTRS model when the attributes are changed. We elect two datasets from the machine learning data repository, University of California

Table 3. The description of datasets

Datesets	Objects	Class	Attributes				
			Categorical	Numerical	Set-valued	Interval-valued	Total
Chess (King-Rock vs.King-Pawn)	3196	2	36	0	0	0	36
Musk (Version 2)	6598	2	0	168	0	0	168
CDATA1	3000	2	10	10	10	10	40
CDATA2	6000	2	20	20	20	20	80
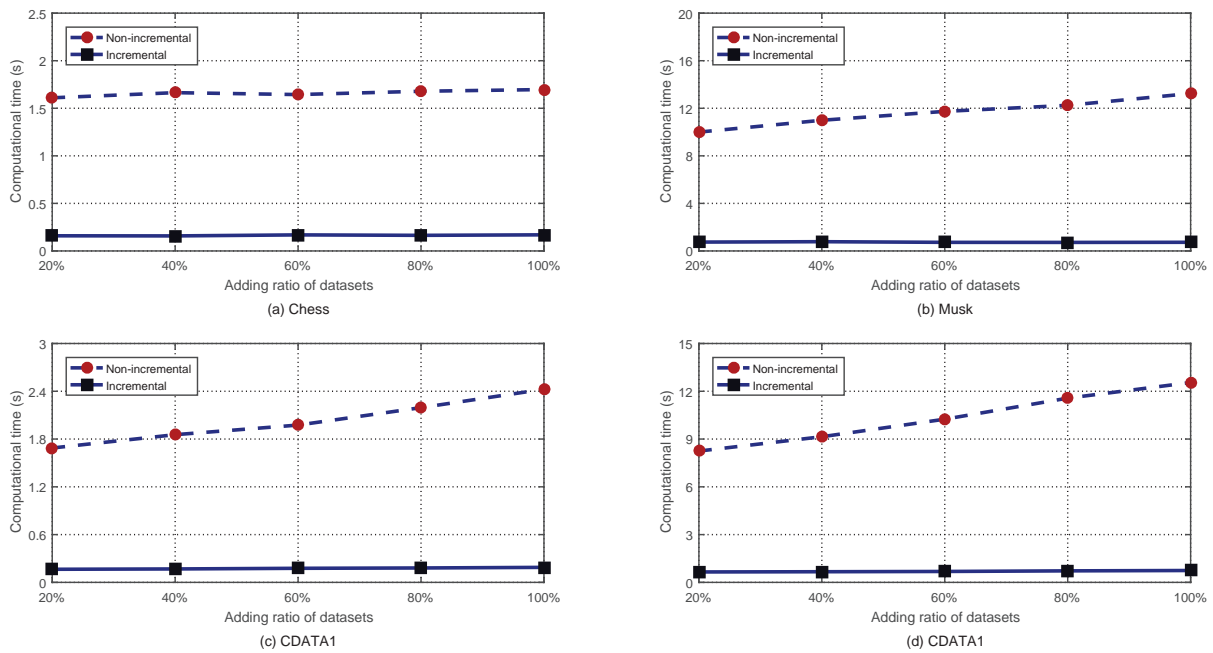
Fig. 1. Comparison of non-incremental and incremental algorithms versus adding different ratios of attributes.

Table 4. The incremental speedup versus adding the different radios of attributes

Datesets	The added ratios of attributes				
	20%	40%	60%	80%	100%
Chess (King-Rock vs.King-Pawn)	10.1067	10.5190	9.7336	10.2247	10.0106
Musk (Version 2	13.4140	14.1118	16.2100	17.0289	18.1688
CDATA1	10.2584	10.9947	11.1742	12.1243	12.8600
CDATA2	12.5924	13.6767	14.9453	15.9917	16.8302

at Irvine (UCI) (`http://archive.ics.uci.edu/ml/`). One is the categorical data and another is the numerical data. Moreover, we generate two composite datasets, which contain four data types, namely, categorical data, numerical data, set-valued data, interval-valued data. The detailed datasets are

shown in Table . All experiments were performed on a computer with Microsoft Windows 10, Inter (R) Core (TM) i5-4210U CUP @ 2.40 GHz and 12.0 GB of memory and the programming language is MATLAB R2016a.

5.1. A comparison of experiments when adding attributes

We divide four datasets into ten equal size of subsets respectively according to the number of attributes. At each datasets, the first five subsets is the original dataset, and the rest five subsets is the added datasets. We set five ratios for adding datasets, namely, $20\%, 40\%, 60\%, 80\%, 100\%$. The comparison of experimental results between Algorithm NCDTRS and Algorithm ICDTRS-AA are shown in Figure 1.

In Figure 1, we can observe that the computational time with respect to Algorithm NCDTRS and Algorithm ICDTRS-AA all increase with addition of attributes. However, it is easy to see that the computational time of incremental algorithm is less than the one of non-incremental algorithm in each

sub-figure of figure 1. Furthermore, the bigger the datasets, more efficient the performance of incremental algorithm will be.

To further show the advantage of the incremental algorithm ICDTRS-AA, we calculate the incremental speedup, which denotes the ratios between the computational time of non-incremental algorithm and the one of incremental algorithm in Table 4. It is easy to see that the incremental speedup of four datasets in Table 3 is greater than one.

5.2. A comparison of experiments when deleting attributes

Similarly to the experimental methods in subsection 5.1, we also divide each dataset into ten equal size of subsets respectively according to the number of attributes. At each datasets, all ten subsets is the original dataset. We delete one subsets step by step from the original datasets. We set five ratios for deleting datasets, namely, $10\%, 20\%, 30\%, 40\%, 50\%$. The comparison of experimental results between Algorithm NCDTRS and Algorithm ICDTRS-DA are shown in Figure 2.
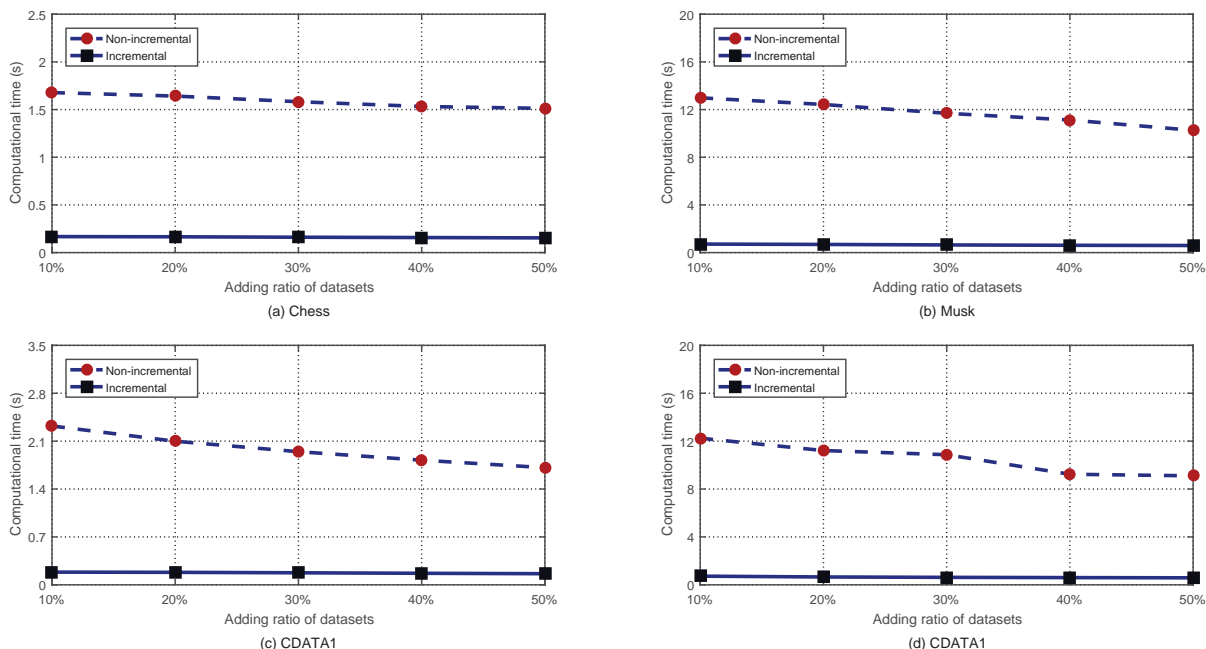

Fig. 2. Comparison of non-incremental and incremental algorithms versus deleting different ratios of attributes.

Table 5. The incremental speedup versus deleting the different radios of attributes

Datesets	The deleted ratios of attributes				
	10%	20%	30%	40%	50%
Chess (King-Rock vs.King-Pawn)	10.0012	9.9005	9.8218	9.7104	9.8074
Musk (Version 2	18.1741	18.1528	18.1931	18.1667	17.2593
CDATA1	12.5081	11.5049	11.0677	10.8165	10.5514
CDATA1	16.9251	17.1266	17.4265	15.3336	15.5728

In Figure 2, it is observed that the computational time with respect to Algorithm NCDTRS and Algorithm ICDTRS-DA all decrease with deletion of attributes. However, we can find that the computational time of incremental algorithm is less than the one of non-incremental algorithm in each sub-figure of figure 2. To further show the advantage of the incremental algorithm ICDTRS-DA, we calculate the incremental speedup, which denotes the ratios between the computational time of non-incremental algorithm and the one of incremental algorithm in Table 5. It is easy to see that the incremental speedup of four datasets in Table 3 is greater than one.

6. Conclusions

In this paper, we investigated the composite information table, which contains various types of data. We proposed the quantitative composite relation for fusion of multiple binary relations. Based on such composite relation, we introduced a quantitative composite DTRS model and provided a novel matrix-based approach to compute approximations. Moreover, to reduce running time w.r.t. the computation of the upper and lower approximations, the increase learning methods based on matrix updating strategy are presented in composite DTRS model versus the addition and deletion of attributes respectively. Experiment results show that the incremental algorithms are more efficient and effective to update approximations in composite DTRS model. Our future work will focus on the incremental updating mechanisms when objects are added or deleted in the composite information system.

References

1. H. M. Chen, T. R. Li, C. Luo, S. J. Horng, and G. Y. Wang. A decision-theoretic rough set approach for dynamic data mining. *IEEE Transactions on Fuzzy Systems*, 23(6):1958–1970, Dec 2015.
2. H. M. Chen, T. R. Li, J. B. Zhang, C. Luo, and X. J. Li. Probabilistic composite rough set and attribute reduction. In *Knowledge Engineering and Management*, pages 189–197. Springer, 2014.
3. Y. Y. Guan and H. K. Wang. Set-valued information systems. *Information Sciences*, 176(17):2507–2525, 2006.
4. J. Hu, T. R. Li, C. Luo, H. Fujita, and S. Y. Li. Incremental fuzzy probabilistic rough sets over two universes. *International Journal of Approximate Reasoning*, 81:28–48, 2017.
5. Q. H. Hu, D. R. Yu, J. F. Liu, and C. X. Wu. Neighborhood rough set based heterogeneous feature subset selection. *Information Sciences*, 178(18):3577–3594, 2008.
6. Y. Y. Huang, T. R. Li, C. Luo, H. Fujita, and S. J. Horng. Dynamic variable precision rough set approach for probabilistic set-valued information systems. *Knowledge-Based Systems*, 122:131–147, 2017.
7. Y. Y. Huang, T. R. Li, C. Luo, H. Fujita, and S. J. Horng. Matrix-based dynamic updating rough fuzzy approximations for data mining. *Knowledge-Based Systems*, 119:273–283, 2017.
8. Y. G. Jing, T. R. Li, H. Fujita, Z. Yu, and B. Wang. An incremental attribute reduction approach based on knowledge granularity with a multi-granulation view.

Information Sciences, 411:23–38, 2017.

9. H. X. Li, L. B. Zhang, B. Huang, and X. Z. Zhou. Sequential three-way decision and granulation for cost-sensitive face recognition. *Knowledge-Based Systems*, 91:241–251, 2016.

10. H. X. Li, L. B. Zhang, X. Z. Zhou, and B. Huang. Cost-sensitive sequential three-way decision modeling using a deep neural network. *International Journal of Approximate Reasoning*, 85:68–78, 2017.

11. T. R. Li, D. Ruan, W. Geert, J. Song, and Y. Xu. A rough sets based characteristic relation approach for dynamic attribute generalization in data mining. *Knowledge-Based Systems*, 20(5):485–494, 2007.

12. W. W. Li, Z. Q. Huang, X. Y. Jia, and X. Y. Cai. Neighborhood based decision-theoretic rough set models. *International Journal of Approximate Reasoning*, 69:1–17, 2016.

13. D. C. Liang, D. Liu, and A. Kobina. Three-way group decisions with decision-theoretic rough sets. *Information Sciences*, 345:46–64, 2016.

14. D. C. Liang, Z. S. Xu, and D. Liu. Three-way decisions with intuitionistic fuzzy decision-theoretic rough sets based on point operators. *Information Sciences*, 375:183–201, 2017.

15. D Liu, T. R. Li, and D. C. Liang. Incorporating logistic regression to decision-theoretic rough sets for classifications. *International Journal of Approximate Reasoning*, 55(1):197–210, 2014.

16. D. Liu and D. C. Liang. Three-way decisions in ordered decision system. *Knowledge-Based Systems*, 137:182–195, 2017.

17. D. Liu, D. C. Liang, and C. C. Wang. A novel three-way decision model based on incomplete information system. *Knowledge-Based Systems*, 91:32–45, 2016.

18. W. Liu, Z. Zheng, and K. Y. Cai. Bi-level programming based real-time path planning for unmanned aerial vehicles. *Knowledge-Based Systems*, 44:34–47, 2013.

19. C. Luo, T. R. Li, H. M. Chen, H. Fujita, and Y. Zhang. Efficient updating of probabilistic approximations with incremental objects. *Knowledge-Based Systems*, 109:71–83, 2016.

20. C. Luo, T. R. Li, H. M. Chen, H. Fujita, and Y. Zhang. Incremental rough set approach for hierarchical multicriteria classification. *Information Sciences*, 417:39–54, 2018.

21. C. Luo, T. R. Li, and Y. Y. Yao. Dynamic probabilistic rough sets with incomplete data. *Information Sciences*, 417:39–54, 2017.

22. C. Luo, T. R. Li, Y. Zhang, and H. Fujita. Matrix approach to decision-theoretic rough sets for evolving data. *Knowledge-Based Systems*, 99:123–134, 2016.

23. Z. Pawlak. Rough sets. *International Journal of Computer & Information Sciences*, 11(5):341–356, 1982.

24. Y. H. Qian, J. Y. Liang, and C. Y. Dang. Interval ordered information systems. *Computers & Mathematics with Applications*, 56(8):1994–2009, 2008.

25. Y. H. Qian, H. Zhang, Y. L. Sang, and J. Y. Liang. Multigranulation decision-theoretic rough sets. *International Journal of Approximate Reasoning*, 55(1, Part 2):225–237, 2014.

26. S. Wang, T. R. Li, C. Luo, and H. Fujita. Efficient updating rough approximations with multi-dimensional variation of ordered data. *Information Sciences*, 372:690–708, 2016.

27. H. L. Yang and Z. L. Guo. Multigranulation decision-theoretic rough sets in incomplete information systems. *International Journal of Machine Learning and Cybernetics*, 6(6):1005–1018, 2015.

28. X. Yang, T. R. Li, H. Fujita, D. Liu, and Y. Y. Yao. A unified model of sequential three-way decisions and multilevel incremental processing. *Knowledge-Based Systems*, 134:172–188, 2017.

29. X. Yang, T. R. Li, D. Liu, H. M. Chen, and C. Luo. A unified framework of dynamic three-way probabilistic rough sets. *Information Sciences*, 420:126–147, 2017.

30. X. B. Yang, Y. Qi, H. L. Yu, X. N. Song, and J. Y. Yang. Updating multigranulation rough approximations with increasing of granular structures. *Knowledge-Based Systems*, 64:59–69, 2014.

31. X. B. Yang, Y. S. Qi, X. N. Song, and J. Y. Yang. Test cost sensitive multigranulation rough set: model and minimal cost selection. *Information Sciences*, 250:184–199, 2013.

32. Y. Y. Yao. Decision-theoretic rough set models. In *International Conference on Rough Sets and Knowledge Technology*, pages 1–12. Springer, 2007.

33. Y. Y. Yao. Three-way decisions with probabilistic rough sets. *Information Sciences*, 180(3):341–353, 2010.

34. Y. Y. Yao and S. K. M. Wong. A decision theoretic framework for approximating concepts. *International Journal of Man-Machine Studies*, 37(6):793–809, 1992.

35. Y. Y. Yao and B. Zhou. Naive bayesian rough sets. In *International Conference on Rough Sets and Knowledge Technology*, pages 719–726. Springer, 2010.

36. J. B. Zhang, T. R. Li, and H. M. Chen. Composite rough sets for dynamic data mining. *Information Sciences*, 257:81–100, 2014.

37. J. B. Zhang, Y. Zhu, Y. Pan, and T. R. Li. Efficient parallel boolean matrix based algorithms for computing composite rough set approximations. *Information Sciences*, 329:287–302, 2016.

38. Y. Y. Zhang, T. R. Li, C. Luo, J. B. Zhang, and H. M. Chen. Incremental updating of rough approximations in interval-valued information systems under attribute generalization. *Information Sciences*, 373:461–475, 2016.

39. X. R. Zhao and B. Q. Hu. Fuzzy and interval-valued

fuzzy decision-theoretic rough set approaches based on fuzzy probability measure. *Information Sciences*, 298:534–554, 2015.

40. Y. Zhao, S. K. M. Wong, and Y. Y. Yao. A note on attribute reduction in the decision-theoretic rough set model. *Transactions on Rough Sets XIII, LNCS*, 13:260–275, 2011.

41. Z. Zheng, J. Lu, G. Q. Zhang, and Q. He. Rule sets based bilevel decision model and algorithm. *Expert Systems with Applications*, 36(1):18–26, 2009.

42. Z. Zheng and G. Y. Wang. A rough set and rule tree based incremental knowledge acquisition algorithm. *Fundamenta Informaticae*, 59:299–313, 2004.