Incremental approximation computation in incomplete ordered decision systems

Guanglei Gou [1,2] **, Guoyin Wang** [3,*]

[1] *School of Information Science and Technology, Southwest Jiaotong University,*
Chengdu 610031, P.R. China

[2] *School of Computer Science and Engineering Chongqing University of Technology,*
Chongqing 400065, P.R. China
E-mail: ggl@cqut.edu.cn

[3] *Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications,*
Chongqing 400065, P.R. China
E-mail: wanggy@cqupt.edu.cn

Received 24 October 2016

Accepted 17 December 2016

Abstract

Approximation computation is a critical step in rough sets theory used in knowledge discovery and other related tasks. In practical applications, an information system often evolves over time by the variation of attributes or objects. Effectively computing approximations is vital in data mining. Dominance-based rough set approach can handle information with preference-ordered attribute domain, but it is not able to handle the situation of data missing. Confidential Dominance-based Rough Set Approach (CDRSA) is introduced to process Incomplete Ordered Decision System (IODS). This paper focuses on incremental updating approximations under dynamic environment in IODS. With the CDRSA, the principles of incremental updating approximations are discussed while the variation of attribute sets or the union of subsets of objects and the corresponding incremental algorithms are developed. Comparative experiments on data sets of UCI and results show that the proposed incremental approaches can improve the performance of updating approximations effectively by a significant shortening of the computational time.

Keywords: Incomplete Ordered Decision Systems, Confidential dominance relation, Approximations, Incremental updating

1. Introduction

Rough set theory is proposed by Pawlak[1] to deal with inconsistency problems, which is useful in fields such as knowledge discovery[2,3,4], decision analysis[5,6], data mining[7,8], etc. Classic Rough Set (CRS) is based on equivalence relation, which is used to deal with discrete attributes values. However, CRS can not be able to discover uncertainty from attributes with preference order domains, which is important to the multi-criteria decision analysis, e.g., risk evaluation, pollution rating. In order to solve this problem, Greco, Matarazzo and Słowinski proposed Dominance-

---

[*] Corresponding author.

based Rough Set Approach (DRSA), which is based on dominance relation instead of equivalence relation[9,10]. DRSA attracted much attention in recent years, and has been applied to multi-criteria classification[11], attribute reduction[12,13], customer behavior prediction[14], bankrupt risk prediction[15], water quality evaluation[16], etc.

DRSA has been introduced to handle the crisp ordered information system, however, Real life applications face data missing due to various causes. Greco et al. extended dominance relation, which requires the referent object has no missing data, to deal with missing data in rough set analysis of multi-attribute and multi-criteria decision problem[17,18]. Błaszczynski et al. discussed different ways of handling missing values in sorting problems with monotonicity constraints in DRSA[19]. He et al. investigated an extend dominance relation to discover knowledge from approximations[20]. Shao and Zhang applied extended DRSA to reasoning, rule extracting and knowledge reduction in incomplete ordered information system[21]. In order to avoid the comparison of two objects have no common known attribute value, Hu and Liu discussed a limited extended dominance[22] and generalized extended dominance relation[23]. Generalized extended dominance relation contains limited extended dominance by chosen the parameter. Chen et al. studied *k*-degree extended dominance characteristic (*k*-degree EDCR), which is considering tow case of missing data that are "do not known" and " do not care"[24]. Du et al. proposed the characteristic-based dominance relation to deal with incomplete ordered information[25]. Yang et al. proposed a similarity dominance relation and studied knowledge reduction under incomplete ordered information system[26]. Luo et al. combined the limited extended dominance relation with similarity dominance relation and proposed a limited extended dominance relation considering maximum and minimum values in dominating or dominated relation when comparing their lost[27]. It is too strict to be practical because maximum and minimum attribute values are hard to known. Yang et al. introduced a valued dominance relation considering the probability that an object is dominating another one by using the information given unknown values[28]. In order to avoid semantic contradiction on the ordered relation, Gou et al. proposed confidential dominance relation to extend DRSA to deal with missing data in Incomplete Ordered Decision System (IODS)[29].

The information system commonly evolves with time, typically, the set of attributes, the set of objects and attribute values may dynamically change. The approximations may alter by the variation of the information system. The traditional methodologies to update approximations are inefficient because they need to recomputing from scratch. Incremental update is a feasible and effective in processing dynamic information since previous data structures and knowledge are optimized for updating approximations. In the case of variation of attributes, Chan proposed an incremental learning method for maintaining approximations in CRS by added into or deleted from one attribute[30]. Li et al. proposed an incremental approach for updating approximations under rough set based the characteristic relation when adding or removing some attributes in incomplete information system[31]. Zhang et al. presented matrix approaches for dynamic attribute variation in set-valued information system[32]. Li et al. studies the principle of updating dominating sets and dominated sets and proposed incremental approaches for updating approximations[33]. Zhang et al. investigated incremental updating of rough approximations in interval-valued information systems[34]. Liu et al. presented approaches for incremental updating approximations in probabilistic rough sets under the variation of attributes[35]. Yang et al. designed fast algorithms for updating the multigranulation rough approximations with increasing of the granular structures[36]. In the case of variation of objects, Shan et al. introduced an incremental modification for classification rules[37]. Zheng and Wang developed a rough set and rule tree based incremental knowledge acquisition algorithm, which updates rules by adding and pruning the rule tree incrementally[38]. Zhang et al. designed parallel algorithms based on MapReduce for calculating approximations of rough sets[39,40]. Luo et al. proposed a matrix approach to decision-theoretic rough sets for evolving data[41]. Luo et al. presented efficient updating of probabilistic approximations with incremental objects[42]. Li et al. developed incremental algorithms for updating approximations of composite rough sets when adding or removing some objects[43]. Li et al. studied the dynamic maintenance of approximations in DRSA when one object is added or removed[44], furthermore, parallel algorithms are developed to speed up computing approximations in DRSA[45]. Błaszczynski and Słowinski proposed an incremental algorithm for induction rules based on the Apriori under the variable consistency DRSA[46]. Wang et al. presented efficient updating rough approximations with multi-dimensional variation of ordered data[47]. In the case of variation of attributes values, Zeng et al. developed dynamical updating fuzzy rough approximations for hybrid data under the variation of attribute values[48]. Chen et al. proposed an incremental method for updating approximations while objects dynamically alter and attributes values vary in variable precision rough set model[49]. And Chen et al. discussed the principles of maintenance of approximations and algorithms for incremental updating approximations are designed when attribute's values coarsening or refining in IODS[24]. However, incremental approaches for updating approximations under IODS when the variation of attributes or objects have not been taken into account until now. In real-life application, such as monitoring domain, attributes and objects are always dynamic changed instead of attribute values. Thus, this paper focuses on incremental methods for updating approximations of confidential dominance-based rough set under the variation of attributes or objects in IODS.

The paper is organized as follows. In Section 2, basic concepts of IODS and dominance relation are reviewed and confidential dominance-based rough set approach is introduced. In Section 3, we discuss the principles of incremental updating approximations and propose algorithms when some attributes are added into or deleted from. In Section 4, incremental approach for updating approximations is given when subsets of objects are merged. In Section 5, some experiments are conducted to evaluate the performance of proposed approaches. The paper ends with conclusion and future work in Section 6.

2. Confidential Dominance-based Rough Set Approach

In this section, some basic concepts of decision systems, including IODS, confidential dominate relation are introduced.

Definition 1 [24] *A decision system is a 4-tuple* $IODS = (U,A,V,f)$. *U is a finite non-empty set of objects, called the universe. A is a non-empty finite set of attributes,* $A = C \cup d$ *where C and d denote the sets of condition attributes and decision attribute. V is a domain of attributes. The domain of C and d is ranked according to an increasing or decreasing preference. The attributes are criteria.* $f : U \times A \to V$ *is an information function.* $f = \{f(x_i,a)|f(x_i,a) : x_i \to v_a, a \in C, x_i \in U, 1 \leqslant i \leqslant |U|\}$, *where* $|\cdot|$ *denotes the cardinality of a set.* $f(x_i,a) = v_a$ ($1 \leqslant i \leqslant |U|$) *denotes the attribute value of object x_i under a. If all attributes? value are know in the decision system, it is a complete ordered decision system. If exists any missing data, it is an incomplete ordered decision system (IODS). All the missing values are denoted by "∗".*

Definition 2 [9,10] *Let* $x,y \in U$, $P \subseteq C$. *If* $\forall a \in P$, $f(y,a) \succeq f(x,a)$ *then we denote* yD_Px. *The relation is called a dominance relation. Here,* $f(y,a) \succeq f(x,a)$ *means y at least as good as (outranks) x with respect to criterion a, which is defined by* $f(y,a) \geqslant f(x,a)$ *according to increasing preference ordering, and* $f(y,a) \leqslant f(x,a)$ *for decreasing ordering.*

Definition 3 *Let* $IODS = (U,A,V,f)$, $P \subseteq C$, $B_P(x) = \{b|b \in P \wedge f(x,b) \neq *\}$. *A confidential dominance relation (CDR) is defined as follows.*

$$\begin{aligned} CDR(P) =& \{(x,y) \in U^2 | (B_P(x) \subseteq B_P(y)) \\ & \wedge \forall_{q \in P}((f(x,q) = * \wedge f(y,q) = *) \vee \\ & (f(x,q) = * \wedge f(y,q) \neq *) \vee f(y,a) \succeq f(x,a)\} \end{aligned} \tag{1}$$

$yD_P^{CDR}x$ denotes that y is dominates x in CDR, which means the object y is more detailed than referent x, and y is weakly preferred over x with respect to each criterion q for which the evaluation of referent x is known.

Definition 4 *For* $P \subseteq C$, $x \in U$, $D_P^{CDR+}(x) = \{y \in U|yD_P^{CDR}x\}$ *is P confidential dominating set of x.*

$D_P^{CDR-}(x)=\{y\in U|xD_P^{CDR}y\}$ *is P confidential dominated set of x.*

The set of decision attributes d partitions U into a finite number of classes. Let $Cl=\{Cl_t|t\in\{1,2,...,n\}\}$, where $Cl_n\succ Cl_{n-1}\succ...\succ Cl_1$. An upward union and downward union of classes are defined respectively as follows: $Cl_t^{\geqslant}=\bigcup_{s\geqslant t}Cl_s$, $Cl_t^{\leqslant}=\bigcup_{s\leqslant t}Cl_s$, where $t,s\in\{1,2,...,n\}$. $x\in Cl_t^{\geqslant}$ means x at least belongs to class Cl_t, while $x\in Cl_t^{\leqslant}$ means x at most belongs to class Cl_t

Definition 5 *Let $IODS=(U,A,V,f)$, $P\subseteq C$, $x\in U$. The P upper and lower approximations of $Cl_t^{\geqslant}$ and $Cl_t^{\leqslant}$ under the CDR are defined respectively as follows.*

$$\underline{P}(Cl_t^{\geqslant})=\{x\in U|D_P^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}$$

$$\overline{P}(Cl_t^{\geqslant})=\bigcup_{x\in Cl_t^{\geqslant}}D_P^{CDR+}(x)=\{x|D_P^{CDR-}(x)\cap Cl_t^{\geqslant}\neq\emptyset\}$$

$$\underline{P}(Cl_t^{\leqslant})=\left\{x\in U|D_P^{CDR-}(x)\subseteq Cl_t^{\leqslant}\right\}$$

$$\overline{P}(Cl_t^{\leqslant})=\bigcup_{x\in Cl_t^{\leqslant}}D_P^{CDR-}(x)=\left\{x|D_P^{CDR+}(x)\cap Cl_t^{\leqslant}\neq\emptyset\right\}$$

The boundary region is defined as follows.

$$Bn_P\left(Cl_t^{\geqslant}\right)=\overline{P}(Cl_t^{\geqslant})-\underline{P}(Cl_t^{\geqslant})$$

$$Bn_P\left(Cl_t^{\leqslant}\right)=\overline{P}(Cl_t^{\leqslant})-\underline{P}(Cl_t^{\leqslant})$$

The objects in the boundary region are inconsistent, which means an object x is confidential dominating another object y , however, x is assigned to a class worse than y.

Definition 6 *For any $x\in U$ in an $IODS=(U,A,V,f)$, $\delta(x)=<l_P(x),u_P(x)>$ is called P generalized decision of the object x, where $l_P(x)=\min\{t|D_P^{CDR+}(x)\cap Cl_t\neq\emptyset\}$ and $u_P(x)=\max\{t|D_P^{CDR-}(x)\cap Cl_t\neq\emptyset\}$.*

Property 1 *Let $IODS=(U,A,V,f)$, $P\subseteq C$. For any $x\in U$, the following items hold.*
(1) if $l_P(x)\geqslant Cl_t$, then $x\in\underline{P}(Cl_t^{\geqslant})$
(2)if $u_P(x)\leqslant Cl_t$, then $x\in\underline{P}(Cl_t^{\leqslant})$
(3)if $l_P(x)\leqslant Cl_t$, then $x\in\overline{P}(Cl_t^{\leqslant})$
(4)if $u_P(x)\geqslant Cl_t$, then $x\in\overline{P}(Cl_t^{\geqslant})$

Proof. (1) U is partitioned by d into decision classes $Cl=\{Cl_t|t\in\{1,2,...,n\}\}$. Since $l_P(x)=\min\{t\in\{1,2...,n\}:D_P^{CDR+}(x)\cap Cl_t\}$, for $\forall y\in D_P^{CDR+}(x)$ we have $f(y,d)\geqslant Cl_t$ if $l_P(x)\geqslant t$. Hence, $D_P^{CDR+}(x)\subseteq Cl_t^{\geqslant}\Rightarrow x\in\underline{P}(Cl_t^{\geqslant})$.
(2)-(4) are similar to (1). □

According to Property 1, approximations of $Cl_t^{\geqslant}$ and $Cl_t^{\leqslant}$ under the CDR are defined as follows.

$$\underline{P}(Cl_t^{\geqslant})=\{x\in U|l_P(x)\geqslant Cl_t\}$$

$$\underline{P}(Cl_t^{\leqslant})=\{x\in U|u_P(x)\leqslant Cl_t\}$$

$$\overline{P}(Cl_t^{\geqslant})=\{x\in U|u_P(x)\geqslant Cl_t\}$$

$$\overline{P}(Cl_t^{\leqslant})=\{x\in U|l_P(x)\leqslant Cl_t\}$$

Example 1 Let us consider an incomplete ordered decision system $IODS=(U,A,V,f)$ presented in Table 1, where $U=\{x_1,x_2,\cdots,x_{10}\}$, $A=C\cup d$, $C=\{a_1,a_2,a_3\}$, $V_d=\{1,2,3,4\}$.

Table 1. An IODS

	a_1	a_2	a_3	d
x_1	4	4	3	4
x_2	3	2	3	3
x_3	4	*	2	3
x_4	2	2	2	2
x_5	2	1	2	2
x_6	3	1	2	1
x_7	*	2	2	2
x_8	4	1	2	2
x_9	3	*	2	3
x_{10}	4	3	3	3

From Table 1, upward union of decision is calculated as follows.

$Cl_4^{\geqslant}=\{x_1\}$, $Cl_3^{\geqslant}=\{x_1,x_2,x_3,x_9,x_{10}\}$, $Cl_2^{\geqslant}=\{x_1,x_2,x_3,x_4,x_5,x_7,x_8,x_9,x_{10}\}$, $Cl_1^{\geqslant}=U$.

C confidential dominating sets are calculated as follows.

$D_C^{CDR+}(x_1)=\{x_1\}$, $D_C^{CDR+}(x_2)=\{x_1,x_2,x_{10}\}$, $D_C^{CDR+}(x_3)=\{x_1,x_3,x_8,x_{10}\}$, $D_C^{CDR+}(x_4)=\{x_1,x_2,x_4,x_{10}\}$,
$D_C^{CDR+}(x_5)=\{x_1,x_2,x_4,x_5,x_6,x_8,x_{10}\}$, $D_C^{CDR+}(x_6)=$

$\{x_1,x_2,x_6,x_8,x_{10}\}$, $D_C^{CDR+}(x_7)=\{x_1,x_2,x_4,x_7,x_{10}\}$, $D_C^{CDR+}(x_8)=\{x_1,x_8,x_{10}\}$, $D_C^{CDR+}(x_9)=\{x_1,x_2,x_3,x_6,x_8,x_9,x_{10}\}$, $D_C^{CDR+}(x_{10})=\{x_1,x_{10}\}$.

Thus, approximations of upward union $Cl_t^{\geqslant}$ with respect to C confidential dominating sets are computed as follows.

$\underline{C}(Cl_4^{\geqslant})=\{x_1\}$, $\underline{C}(Cl_3^{\geqslant})=\{x_1,x_2,x_{10}\}$, $\underline{C}(Cl_2^{\geqslant})=\{x_1,x_2,x_3,x_4,x_7,x_8,x_{10}\}$, $\underline{C}(Cl_1^{\geqslant})=U$.

$\overline{C}(Cl_4^{\geqslant})=\{x_1\}$, $\overline{C}(Cl_3^{\geqslant})=\{x_1,x_2,x_3,x_6,x_8,x_9,x_{10}\}$, $\overline{C}(Cl_2^{\geqslant})=U$, $\overline{C}(Cl_1^{\geqslant})=U$.

Analogously, approximations of C confidential dominated set with respect to upward union $Cl_t^{\leqslant}$ can be calculated. For simplicity and clarity, the following examples discuss CDRSA only with the increasing preference order.

3. Incremental updating approximations under the variation of attribute sets

Approximations computing is a critical step for applying CDRSA in knowledge discovery. Traditional approach of approximations computing is time-consuming when attributes are changed because it should recompute from scratch. In this section, approaches of incremental updating approximations under the variation of attributes are discussed.

3.1. Incremental updating approximations when new attributes added

Lemma 1 *Let $IODS=(U,A,V,f)$, $P\subset C$ and $Q\subseteq C-P$. For any $x\in U$, we have*

(1)$D_{P\cup Q}^{CDR+}(x)=D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)$

(2)$D_{P\cup Q}^{CDR-}(x)=D_P^{CDR-}(x)\cap D_Q^{CDR-}(x)$

Proof. (1) Assume $y\in D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)$, namely $y\in D_P^{CDR+}(x)$ and $y\in D_Q^{CDR+}(x)$, which means $(x,y)\subseteq CDR(P)$ and $(x,y)\subseteq CDR(Q)$. It follows $y\in D_{P\cup Q}^{CDR+}(x)$. $D_{P\cup Q}^{CDR+}(x)=D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)$ holds.
(2) It is similar to (1). □

By Lemma 1, Confidential dominating/dominated set of x can be easily obtained when attribute set Q is added into P.

Theorem 1 *Let $IODS=(U,A,V,f)$, $P\subset C$ and $Q\subseteq C-P$. For each $Cl_t^{\geqslant}$, we have*

(1)$\underline{P\cup Q}(Cl_t^{\geqslant})=\underline{P}(Cl_t^{\geqslant})\cup T(Cl_t^{\geqslant})$,

where $T(Cl_t^{\geqslant})=\{x|l_{P\cup Q}(D_{P\cup Q}^{CDR+}(x))\geqslant Cl_t : x\in Bn_P(Cl_t^{\geqslant})\}$

(2)$\overline{P\cup Q}(Cl_t^{\geqslant})=\overline{P}(Cl_t^{\geqslant})-K(Cl_t^{\geqslant})$,

where $K(Cl_t^{\geqslant})=\{x|u_{P\cup Q}(D_{P\cup Q}^{CDR-}(x))<Cl_t : x\in \overline{P}(Cl_t^{\geqslant})\}$

Proof. (1) $\underline{P\cup Q}(Cl_t^{\geqslant})=\{x\in U|D_{P\cup Q}^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}$.
Since $D_{P\cup Q}^{CDR+}(x)=D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)$, we have $\underline{P\cup Q}(Cl_t^{\geqslant})=\{x\in U|D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}=\{x\in U|D_P^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}\cup\{x\in U-\underline{P}(Cl_t^{\geqslant})|D_{P\cup Q}^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}$. And $\{x\in U-\underline{P}(Cl_t^{\geqslant})\}=\{x\in(Bn_P(Cl_t^{\geqslant})\cup(U-\overline{P}(Cl_t^{\geqslant})))\}$, it follows $\{x\in U-\underline{P}(Cl_t^{\geqslant})|D_{P\cup Q}^{CDR+}(x)\subseteq Cl_t^{\geqslant}\}=\{x\in Bn_P(Cl_t^{\geqslant})|l_{P\cup Q}(D_{P\cup Q}^{CDR+}(x))\geqslant Cl_t\}$. Let $T(Cl_t^{\geqslant})=\{x|l_{P\cup Q}(D_{P\cup Q}^{CDR+}(x))\geqslant Cl_t : x\in Bn_P(Cl_t^{\geqslant})\}$. $\underline{P\cup Q}(Cl_t^{\geqslant})=\underline{P}(Cl_t^{\geqslant})\cup T(Cl_t^{\geqslant})$ holds.

(2) Since $\overline{P\cup Q}(Cl_t^{\geqslant})=\bigcup_{x\in Cl_t^{\geqslant}}D_{P\cup Q}^{CDR+}(x)=\bigcup_{x\in Cl_t^{\geqslant}}(D_P^{CDR+}(x)\cap D_Q^{CDR+}(x))$
and $D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)=D_P^{CDR+}(x)-(U-D_{P\cup Q}^{CDR+}(x))$,
we have $\bigcup_{x\in Cl_t^{\geqslant}}(D_P^{CDR+}(x)\cap D_Q^{CDR+}(x))=\bigcup_{x\in Cl_t^{\geqslant}}(D_P^{CDR+}(x)-(U-D_{P\cup Q}^{CDR+}(x)))$
$=\bigcup_{x\in Cl_t^{\geqslant}}D_P^{CDR+}(x)-\bigcup_{x\in Cl_t^{\geqslant}}(U-D_{P\cup Q}^{CDR+}(x)))$. According to Property 1.(4), we have $\overline{P\cup Q}(Cl_t^{\geqslant})=\overline{P}(Cl_t^{\geqslant})-\{u_{P\cup Q}(D_{P\cup Q}^{CDR-}(x))<Cl_t : x\in\overline{P}(Cl_t^{\geqslant})\}$. Let $K(Cl_t^{\geqslant})=\{x|u_{P\cup Q}(D_{P\cup Q}^{CDR-}(x))<Cl_t : x\in\overline{P}(Cl_t^{\geqslant})\}$.
$\overline{P\cup Q}(Cl_t^{\geqslant})=\overline{P}(Cl_t^{\geqslant})-K(Cl_t^{\geqslant})$ holds. □

Theorem 2 *Let $IODS=(U,A,V,f)$, $P\subset C$ and $Q\subseteq C-P$. For each $Cl_t^{\leqslant}$, we have*

(1)$\underline{P\cup Q}(Cl_t^{\leqslant})=\underline{P}(Cl_t^{\leqslant})\cup Y(Cl_t^{\leqslant})$,

where $Y(Cl_t^{\leqslant})=\{x|u_{P\cup Q}(D_{P\cup Q}^{CDR-}(x))\leqslant Cl_t : x\in Bn_P(Cl_t^{\leqslant})\}$

(2)$\overline{P\cup Q}(Cl_t^{\leqslant})=\overline{P}(Cl_t^{\leqslant})-Z(Cl_t^{\leqslant})$,

where $Z(Cl_t^{\leqslant})=\{x|l_{P\cup Q}(D_{P\cup Q}^{CDR+}(x))>Cl_t : x\in\overline{P}(Cl_t^{\leqslant})\}$

Proof. It is similar to proof of Theorem 1. □

Based on Theorems 1-2, an algorithm of updating approximation of CDRSA is proposed when new attributes are added into the IODS (See algorithm 1 **IUA**).

Example 2 (Continued from Example 1) Let $P = \{a_1, a_3\}$, $Q = \{a_2\}$, $C = P \cup Q$.

$D_P^{CDR+}(x_1) = D_P^{CDR+}(x_{10}) = \{x_1, x_{10}\}$,
$D_P^{CDR+}(x_2) = \{x_1, x_2, x_{10}\}$,
$D_P^{CDR+}(x_3) = D_P^{CDR+}(x_8) = \{x_1, x_3, x_8, x_{10}\}$,
$D_P^{CDR+}(x_4) = D_P^{CDR+}(x_5) = U - \{x_7\}$,
$D_P^{CDR+}(x_6) = D_P^{CDR+}(x_9) = \{x_1, x_2, x_3, x_6, x_8, x_9, x_{10}\}$,
$D_P^{CDR+}(x_7) = U$.

$D_Q^{CDR+}(x_1) = \{x_1\}$,
$D_Q^{CDR+}(x_2) = D_Q^{CDR+}(x_4) = D_Q^{CDR+}(x_7) = \{x_1, x_2, x_4, x_7, x_{10}\}$,
$D_Q^{CDR+}(x_3) = D_Q^{CDR+}(x_9) = U$,
$D_Q^{CDR+}(x_5) = D_Q^{CDR+}(x_6) = D_Q^{CDR+}(x_8) = U - \{x_3, x_9\}$.

Approximation of $Cl_t^{\geqslant}$ with respect to P confidential dominating sets are computing as follows.

$\underline{P}(Cl_4^{\geqslant}) = \emptyset$, $\underline{P}(Cl_3^{\geqslant}) = \{x_1, x_2, x_{10}\}$,
$\underline{P}(Cl_2^{\geqslant}) = \{x_1, x_2, x_3, x_8, x_{10}\}$, $\underline{P}(Cl_1^{\geqslant}) = U$.

$\overline{P}(Cl_4^{\geqslant}) = \{x_1, x_{10}\}$, $\overline{P}(Cl_3^{\geqslant}) = \{x_1, x_2, x_3, x_6, x_8, x_9, x_{10}\}$,
$\overline{P}(Cl_2^{\geqslant}) = U$, $\overline{P}(Cl_1^{\geqslant}) = U$.

According to Lemma 1, $C = P \cup Q$ confidential dominating sets are easily calculated as follows. Take x_1 and x_4 for example.
$D_C^{CDR+}(x_1) = D_{P\cup Q}^{CDR+}(x_1) = D_P^{CDR+}(x_1) \cap D_Q^{CDR+}(x_1) = \{x_1\}$.
$D_{P\cup Q}^{CDR+}(x_4) = D_P^{CDR+}(x_4) \cap D_Q^{CDR+}(x_4) = \{x_1, x_2, x_4, x_{10}\}$.

Based on Theorem 2, approximations of $Cl_t^{\geqslant}$ with respect to $P \cup Q$ confidential dominating sets are computing as follows.

$T(Cl_4^{\geqslant}) = \{x | l_{P\cup Q}(D_{P\cup Q}^{CDR+}(x)) \geqslant Cl_4 : x \in Bn_P(Cl_4^{\geqslant})\} = \{x_1\}$,
$T(Cl_3^{\geqslant}) = \emptyset$, $T(Cl_2^{\geqslant}) = \{x_4, x_7\}$, $T(Cl_1^{\geqslant}) = \emptyset$.

$K(Cl_4^{\geqslant}) = \{x | u_{P\cup Q}(D_{P\cup Q}^{CDR-}(x)) < Cl_4 : x \in \overline{P}(Cl_4^{\geqslant})\} = \{x_{10}\}$,
$K(Cl_3^{\geqslant}) = K(Cl_2^{\geqslant}) = K(Cl_1^{\geqslant}) = \emptyset$.

$\underline{P \cup Q}(Cl_4^{\geqslant}) = \underline{P}(Cl_4^{\geqslant}) \cup T(Cl_4^{\geqslant}) = \{x_1\}$, $\underline{P \cup Q}(Cl_3^{\geqslant}) = \{x_1, x_2, x_{10}\}$, $\underline{P \cup Q}(Cl_2^{\geqslant}) = \{x_1, x_2, x_3, x_4, x_7, x_8, x_{10}\}$, $\underline{P \cup Q}(Cl_1^{\geqslant}) = U$.

$\overline{P \cup Q}(Cl_4^{\geqslant}) = \overline{P}(Cl_4^{\geqslant}) - K(Cl_4^{\geqslant}) = \{x_1\}$,
$\overline{P \cup Q}(Cl_3^{\geqslant}) = \{x_1, x_2, x_3, x_6, x_8, x_9, x_{10}\}$. $\overline{P}(Cl_2^{\geqslant}) = \overline{P}(Cl_1^{\geqslant}) = U$.

3.2. Incremental updating approximations when attributes deleted

Lemma 2 *Let IODS* $= (U, A, V, f)$, $P \subset C$ *and* $Q \subset P \subseteq C$. *For any* $x \in U$, *we have*

(1)$D_{P-Q}^{CDR+}(x) = D_P^{CDR+}(x) \cup \mathrm{A}_Q^{\geqslant}(x)$,

where $\mathrm{A}_Q^{\geqslant}(x) = \{y \in U - D_P^{CDR+}(x) | (U - D_P^{CDR+}(x)) \wedge D_{P-Q}^{CDR+}(x)\}$.

(2) $D_{P-Q}^{CDR-}(x) = D_P^{CDR-}(x) \cup \mathrm{A}_Q^{\leqslant}(x)$,

where $\mathrm{A}_Q^{\leqslant}(x) = \{y \in U - D_P^{CDR-}(x) | (U - D_P^{CDR-}(x)) \wedge D_{P-Q}^{CDR-}(x)\}$.

Proof. (1) Proof by contradiction. Assume $y \in D_{P-Q}^{CDR+}(x)$ satisfies $y \in U - D_P^{CDR+}(x)$ and $y \notin \mathrm{A}_Q^{\geqslant}(x)$, namely, $y \notin \{U - D_P^{CDR+}(x) \wedge D_{P-Q}^{CDR+}(x)\}$, then $y \in D_P^{CDR+}(x)$ or $y \in U - D_{P-Q}^{CDR+}(x)$. (a) $y \in D_P^{CDR+}(x)$ is in contradiction with the assumption $y \in U - D_P^{CDR+}(x)$; (b) $y \in U - D_{P-Q}^{CDR+}(x)$, we have $(x, y) \subseteq U - D_{P-Q}^{CDR+}(x)$ which means $y \notin D_{P-Q}^{CDR+}(x)$. Contradiction. $D_{P-Q}^{CDR+}(x) = D_P^{CDR+}(x) \cup \mathrm{A}_Q^{\geqslant}(x)$ holds.

(2) It is similar to Lemma 2(1). □

Theorem 3 *Let IODS* $= (U, A, V, f)$, $P \subset C$ *and* $Q \subset P \subseteq C$. *For any* $x \in U$, *we have*

(1) $\underline{P - Q}(Cl_t^{\geqslant}) = \underline{P}(Cl_t^{\geqslant}) - W(Cl_t^{\geqslant})$,

where $W(Cl_t^{\geqslant}) = \{x | l_{P-Q}(\mathrm{A}_Q^{\geqslant}(x)) < Cl_t, x \in \underline{P}(Cl_t^{\geqslant})\}$

(2) $\overline{P - Q}(Cl_t^{\geqslant}) = \overline{P}(Cl_t^{\geqslant}) \cup R(Cl_t^{\geqslant})$,

where $R(Cl_t^{\geqslant}) = \{\mathrm{A}_Q^{\geqslant}(x) : x \in \overline{P}(Cl_t^{\geqslant})\}$

Proof. (1) $\underline{P - Q}(Cl_t^{\geqslant}) = \{x \in U | D_{P-Q}^{CDR+}(x) \subseteq Cl_t^{\geqslant}\} = \underline{P}(Cl_t^{\geqslant}) - \{x \in \underline{P}(Cl_t^{\geqslant}) | D_{P-Q}^{CDR+}(x) \nsubseteq Cl_t^{\geqslant}\}$. Let $W(Cl_t^{\geqslant}) = \{x \in \underline{P}(Cl_t^{\geqslant}) | D_{P-Q}^{CDR+}(x) \nsubseteq Cl_t^{\geqslant}\} = \{x | l_{P-Q}(\mathrm{A}_Q^{\geqslant}(x)) < Cl_t, x \in \underline{P}(Cl_t^{\geqslant})\}$.
$\underline{P - Q}(Cl_t^{\geqslant}) = \underline{P}(Cl_t^{\geqslant}) - W(Cl_t^{\geqslant})$ holds.

(2) $\overline{P - Q}(Cl_t^{\geqslant}) = \bigcup_{x \in Cl_t^{\geqslant}} D_{P-Q}^{CDR+}(x) = \overline{P}(Cl_t^{\geqslant}) \cup \{\mathrm{A}_Q^{\geqslant}(x) : x \in \overline{P}(Cl_t^{\geqslant})\}$. Let $R(Cl_t^{\geqslant}) = \{\mathrm{A}_Q^{\geqslant}(x) : x \in \overline{P}(Cl_t^{\geqslant})\}$.
$\overline{P - Q}(Cl_t^{\geqslant}) = \overline{P}(Cl_t^{\geqslant}) \cup R(Cl_t^{\geqslant})$ holds. □

Algorithm 1 (IUA) Incremental **U**pdating approximations when new attributes **A**dded

```
Require: $P$ confidential dominating/dominated sets; $\underline{P}(Cl_t^{\geqslant})$, $\overline{P}(Cl_t^{\geqslant})$
Ensure: $\overline{P\cup Q}(Cl_t^{\geqslant})$, $\underline{P\cup Q}(Cl_t^{\geqslant})$
Compute $Q$ confidential dominating/dominated sets
for i=1:$|U|$ do
    $D_{P\cup Q}^{CDR+}(x)=D_P^{CDR+}(x)\cap D_Q^{CDR+}(x)$
    $D_{P\cup Q}^{CDR-}(x)=D_P^{CDR-}(x)\cap D_Q^{CDR-}(x)$
end for
for i=1:$|Cl_t^{\geqslant}|$ do
    Compute $T(Cl_t^{\geqslant})$ and $K(Cl_t^{\geqslant})$
    Compute $Y(Cl_t^{\geqslant})$ and $Z(Cl_t^{\geqslant})$
    $\underline{P\cup Q}(Cl_t^{\geqslant})=\underline{P}(Cl_t^{\geqslant})\cup T(Cl_t^{\geqslant})$ and $\underline{P\cup Q}(Cl_t^{\leqslant})=\underline{P}(Cl_t^{\leqslant})\cup Y(Cl_t^{\geqslant})$
    $\overline{P\cup Q}(Cl_t^{\geqslant})=\overline{P}(Cl_t^{\geqslant})-K(Cl_t^{\geqslant})$ and $\overline{P\cup Q}(Cl_t^{\leqslant})=\overline{P}(Cl_t^{\leqslant})-Z(Cl_t^{\leqslant})$
end for
```

Theorem 4 *Let* $IODS=(U,A,V,f)$, $P\subset C$ *and* $Q\subset P\subseteq C$. *For any* $x\in U$, *we have*

(1) $\underline{P-Q}(Cl_t^{\leqslant})=\underline{P}(Cl_t^{\leqslant})-S(Cl_t^{\leqslant})$,
where $S(Cl_t^{\leqslant})=\{x|u_{P-Q}(\mathrm{A}_Q^{\leqslant}(x))>Cl_t,x\in\underline{P}(Cl_t^{\leqslant})\}$

(2) $\overline{P-Q}(Cl_t^{\leqslant})=\overline{P}(Cl_t^{\leqslant})\cup V(Cl_t^{\leqslant})$,
where $V(Cl_t^{\leqslant})=\{\mathrm{A}_Q^{\leqslant}(x):x\in\overline{P}(Cl_t^{\leqslant})\}$

Proof. It is similar to proof of Theorem 3. □

Based on Theorems 3-4, an algorithm is developed for updating approximations of CDRSA when some attributes are deleted from the IODS (see Algorithm 2 **IUD**).

Example 3 (Continued from Example 1) Let $x\in U$, $P=C-Q$.

According to Lemma 2, for any x, $\mathrm{A}_Q^{\geqslant}(x)$ is easily calculated as follows.

$\mathrm{A}_Q^{\geqslant}(x_1)=\{x_{10}\}$, $\mathrm{A}_Q^{\geqslant}(x_2)=\mathrm{A}_Q^{\geqslant}(x_3)=\mathrm{A}_Q^{\geqslant}(x_9)=\mathrm{A}_Q^{\geqslant}(x_{10})=\emptyset$,
$\mathrm{A}_Q^{\geqslant}(x_4)=\{x_3,x_5,x_6,x_8,x_9,x_{10}\}$, $\mathrm{A}_Q^{\geqslant}(x_5)=\mathrm{A}_Q^{\geqslant}(x_6)=\{x_3,x_9\}$,
$\mathrm{A}_Q^{\geqslant}(x_7)=\{x_3,x_5,x_6,x_8,x_9\}$, $\mathrm{A}_Q^{\geqslant}(x_8)=\{x_3\}$

Approximation of $Cl_t^{\geqslant}$ with respect to $P=C-Q$ confidential dominating sets are computing as follows.

$W(Cl_4^{\geqslant})=\{x|l_{C-Q}(\mathrm{A}_Q^{\geqslant}(x))<Cl_4,x\in\underline{P}(Cl_4^{\geqslant})\}=\{x_1\}$, $W(Cl_3^{\geqslant})=\emptyset$, $W(Cl_2^{\geqslant})=\{x_4,x_7\}$, $W(Cl_1^{\geqslant})=\emptyset$.

$R(Cl_4^{\geqslant})=\{\mathrm{A}_Q^{\geqslant}(x):x\in\overline{P}(Cl_4^{\geqslant})\}=\{x_{10}\}$, $R(Cl_3^{\geqslant})=\{x_3,x_9,x_{10}\}$, $R(Cl_2^{\geqslant})=\{x_3,x_5,x_6,x_8,x_9,x_{10}\}$, $R(Cl_1^{\geqslant})=\{x_3,x_5,x_6,x_8,x_9,x_{10}\}$.

$\underline{P}=\underline{C-Q}(Cl_4^{\geqslant})=\underline{C}-W(Cl_4^{\geqslant})=\emptyset$, $\underline{C-Q}(Cl_3^{\geqslant})=\{x_1,x_2,x_{10}\}$, $\underline{C-Q}(Cl_2^{\geqslant})=\{x_1,x_2,x_3,x_8,x_{10}\}$, $\underline{C-Q}(Cl_1^{\geqslant})=U$

$\overline{P}(Cl_4^{\geqslant})=\underline{C-Q}(Cl_4^{\geqslant})=\underline{C}(Cl_4^{\geqslant})\cup R(Cl_4^{\geqslant})=\{x_1,x_{10}\}$,
$\overline{P}(Cl_3^{\geqslant})=\{x_1,x_2,x_3,x_6,x_8,x_9,x_{10}\}$, $\overline{P}(Cl_2^{\geqslant})=U$, $\overline{P}(Cl_1^{\geqslant})=U$

4. Incremental updating approximations when subsets of objects are merged

In this section, incremental updating approximations of CDRSA is discussed when subsets of objects are merged into IODS.

Assume the universe of $IODS=(U,A,V,f)$ is composed by m subsets of objects, where $U=\bigcup_{i=1}^{m}U_i$, $i=1,2,\cdots,m$.

Definition 7 *Let* $U_k\subset U$, $k=1,2,\cdots,m$. $Cl_t^{\geqslant k}$ *and* $Cl_t^{\leqslant k}$ *in subset* U_k *are defined as follows.*

$$Cl_t^{\geqslant k}=\{x\in U_k|f(x,d)\geqslant Cl_t\}$$

$$Cl_t^{\leqslant k}=\{x\in U_k|f(x,d)\leqslant Cl_t\}$$

Algorithm 2 (IUD) Incremental **U**pdating approximations when some attributes **D**eleted

```
Require: $P$ confidential dominating/dominated sets; $\underline{P}(Cl_t^{\geqslant})$, $\overline{P}(Cl_t^{\geqslant})$
Ensure: $\overline{P-Q}(Cl_t^{\geqslant})$, $\underline{P-Q}(Cl_t^{\geqslant})$
for i=1:$|U|$ do
    Compute $A_Q^{\geqslant}(x)$ and $A_Q^{\leqslant}(x)$
    $D_{P-Q}^{CDR+}(x) = D_P^{CDR+}(x) \cup A_Q^{\geqslant}(x)$
    $D_{P-Q}^{CDR-}(x) = D_P^{CDR-}(x) \cup A_Q^{\leqslant}(x)$
end for
for i=1:$|Cl_t^{\geqslant}|$ do
    Compute $W(Cl_t^{\geqslant})$ and $R(Cl_t^{\geqslant})$
    $\underline{P-Q}(Cl_t^{\geqslant}) = \underline{P}(Cl_t^{\geqslant}) - W(Cl_t^{\geqslant})$ and $\underline{P-Q}(Cl_t^{\leqslant}) = \underline{P}(Cl_t^{\leqslant}) - S(Cl_t^{\leqslant})$
    $\overline{P-Q}(Cl_t^{\geqslant}) = \overline{P}(Cl_t^{\geqslant}) \cup R(Cl_t^{\geqslant})$ and $\overline{P-Q}(Cl_t^{\leqslant}) = \overline{P}(Cl_t^{\leqslant}) \cup V(Cl_t^{\leqslant})$
end for
```

Definition 8 *Let $U_i \subset U$, $i = 1, 2, \cdots, m$. The P confidential dominating/dominated sets of the object x related to U_i are defined as follows.*

$$D_P^{CDR+}(x)^i = \{y \in U_i | y D_P^{CDR} x\}$$

$$D_P^{CDR-}(x)^i = \{y \in U_i | x D_P^{CDR} y\}$$

Definition 9 *Let $U_k \subset U$, $i, k = 1, 2, \cdots, m$. Lower and upper approximations of $Cl_t^{\geqslant k}$ related to U_i are defined as follows.*

$$\underline{P}(Cl_t^{\geqslant k})^i = \{x \in U_i | D_P^{CDR+}(x)^i \subseteq Cl_t^{\geqslant k}\}$$

$$\overline{P}(Cl_t^{\geqslant k})^i = \bigcup_{x \in Cl_t^{\geqslant k}} D_P^{CDR+}(x)^i = \{x \in U_i | D_P^{CDR-}(x)^i \cap Cl_t^{\geqslant k} \neq \emptyset\}$$

Similarly, lower and upper approximations of $Cl_t^{\leqslant k}$ related to U_i are defined as follows.

$$\underline{P}(Cl_t^{\leqslant k})^i = \{x \in U_i | D_P^{CDR-}(x)^i \subseteq Cl_t^{\leqslant k}\}$$

$$\overline{P}(Cl_t^{\leqslant k})^i = \bigcup_{x \in Cl_t^{\leqslant k}} D_P^{CDR-}(x)^i = \{x \in U_i | D_P^{CDR+}(x)^i \cap Cl_t^{\leqslant k} \neq \emptyset\}$$

Theorem 5 *Let $i = 1, 2, \cdots, m$, $t = 1, 2, \cdots, n$, we have*

(1) $Cl_t^{\geqslant} = \bigcup_{i=1}^m Cl_t^{\geqslant i}$

(2) $Cl_t^{\leqslant} = \bigcup_{i=1}^m Cl_t^{\leqslant i}$

Proof. (1) Since $Cl_t^{\geqslant} = \{x \in U | f(x, D) \geqslant Cl_t\} = \{x \in \bigcup_{i=1}^m U_i | f(x, D) \geqslant Cl_t\} = \bigcup_{i=1}^m \{x \in U_i | f(x, D) \geqslant Cl_t\} = \bigcup_{i=1}^m Cl_t^{\geqslant i}$, $Cl_i^{\geqslant} = \bigcup_{k=1}^m Cl_t^{\geqslant i}$ holds.

(2) It is similar to proof Theorem 5(1). □

Theorem 6 *Let $x \in U, i = 1, 2, \cdots, m$, we have*

(1) $D_P^{CDR+}(x) = \bigcup_{i=1}^m D_P^{CDR+}(x)^i$

(2) $D_p^{CDR-}(x) = \bigcup_{i=1}^m D_P^{CDR-}(x)^i$

Proof. (1) Since $D_P^{CDR+}(x) = \{x \in U | y D_P^{CDR} x\} = \{x \in \bigcup_{i=1}^m U_i | y D_P^{CDR} x\} = \bigcup_{i=1}^m \{x \in U_i | y D_P^{CDR} x\} = \bigcup_{i=1}^m D_P^{CDR+}(x)^i$, $D_P^{CDR+}(x) = \bigcup_{i=1}^m D_P^{CDR+}(x)^i$ holds.

(2) It is similar to proof Theorem 6(1). □

Theorem 7 *Given $t = 1, 2, \cdots, n, i = 1, 2, \cdots, m$, the following items hold.*

(1) $\underline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m (\bigcap_{k=1}^m \underline{P}(Cl_t^{\geqslant k})^i)$

(2) $\overline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m \bigcup_{k=1}^m \overline{P}(Cl_t^{\geqslant k})^i$

Proof. (1) Since $\underline{P}(Cl_t^{\geqslant}) = \{x \in U | D_p^{CDR+}(x) \subseteq Cl_t^{\geqslant}\} =$
$\{x \in \bigcup_{i=1}^m U_i | \bigcup_{k=1}^m D_p^{CDR+}(x)^k \subseteq \bigcup_{k=1}^m Cl_t^{\geqslant k}\} =$
$\bigcup_{i=1}^m \{x \in U_i | D_p^{CDR+}(x)^1 \subseteq Cl_t^{\geqslant 1} \wedge \cdots \wedge D_p^{CDR+}(x)^m \subseteq Cl_t^{\geqslant m}\} =$
$\bigcup_{i=1}^m (\bigcap_{k=1}^m \{x \in U_i | D_p^{CDR+}(x)^k \subseteq Cl_t^{\geqslant k}\}) = \bigcup_{i=1}^m (\bigcap_{k=1}^m \underline{P}(Cl_t^{\geqslant k})^i)$,
$\underline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m (\bigcap_{k=1}^m \underline{P}(Cl_t^{\geqslant k})^i)$ holds.

(2) Since $\overline{P}(Cl_t^{\geqslant}) = \{x \in U | D_p^{CDR-}(x) \bigcap Cl_t^{\geqslant} \neq \emptyset\} =$
$\{x \in \bigcup_{i=1}^m U_i | \bigcup_{k=1}^m D_p^{CDR-}(x)^k \bigcap \bigcup_{k=1}^m Cl_t^{\geqslant k} \neq \emptyset\} =$
$\bigcup_{i=1}^m \{x \in U_i | D_p^{CDR-}(x)^1 \bigcap Cl_t^{\geqslant 1} \neq \emptyset \wedge \cdots \wedge D_p^{CDR-}(x)^m \bigcap Cl_t^{\geqslant m} \neq \emptyset\} = \bigcup_{i=1}^m (\bigcup_{k=1}^m \{x \in U_i | D_p^{CDR-}(x)^k \bigcap Cl_t^{\geqslant k} \neq \emptyset\}) = \bigcup_{i=1}^m \bigcup_{k=1}^m \overline{P}(Cl_t^{\geqslant k})^i$,
$\overline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m \bigcup_{k=1}^m \overline{P}(Cl_t^{\geqslant k})^i$ holds. □

Theorem 8 *Given* $t = 1,2,\cdots,n, i = 1,2,\cdots,m$, *the following items hold.*

(1) $\underline{P}(Cl_t^{\leqslant}) = \cup_{i=1}^{m}(\cap_{k=1}^{m}\underline{P}(Cl_t^{\leqslant k})^i)$

(2) $\overline{P}(Cl_t^{\leqslant}) = \cup_{i=1}^{m}\cup_{k=1}^{m}\overline{P}(Cl_t^{\leqslant k})^i$

Proof. It is similar to proof of Theorem 7. □

Based on Theorems 5-8, an algorithm is designed for updating approximations of CDRSA when subsets of objects are merged IODS (see Algorithm 3 **IUM**).

Example 4 (Continued from Example 1) Let $U = U_1 \cup U_2$, where $U_1 = \{x_1,x_2,\cdots,x_7\}$ and $U_2 = \{x_8,x_9,x_{10}\}$.

According to Definition, $k = 1,2$. $Cl_t^{\geqslant k}$ in subset U_k are calculated as follows.

$Cl_4^{\geqslant 1} = \{x_1\}$, $Cl_3^{\geqslant 1} = \{x_1,x_2,x_3\}$, $Cl_2^{\geqslant 1} = \{x_1,x_2,x_3,x_4,x_5,x_7\}$, $Cl_1^{\geqslant 1} = U_1$.

$Cl_3^{\geqslant 2} = \{x_9,x_{10}\}$, $Cl_2^{\geqslant 2} = \{x_8,x_9,x_{10}\}$, $Cl_1^{\geqslant 2} = U_2$.

The C confidential dominating sets of the object x related to $U_i (i = 1,2)$ are calculated as follows.

$D_C^{CDR+}(x_1)^1 = \{x_1\}$, $D_C^{CDR+}(x_2)^1 = \{x_1,x_2\}$, $D_C^{CDR+}(x_3)^1 = \{x_1,x_3\}$, $D_C^{CDR+}(x_4)^1 = \{x_1,x_2,x_4\}$, $D_C^{CDR+}(x_5)^1 = \{x_1,x_2,x_4,x_5,x_6\}$, $D_C^{CDR+}(x_6)^1 = \{x_1,x_2,x_6\}$, $D_C^{CDR+}(x_7)^1 = \{x_1,x_2,x_4,x_7\}$.

$D_C^{CDR+}(x_1)^2 = \emptyset$, $D_C^{CDR+}(x_2)^2 = \{x_{10}\}$, $D_C^{CDR+}(x_3)^2 = \{x_8,x_{10}\}$, $D_C^{CDR+}(x_4)^2 = \{x_{10}\}$, $D_C^{CDR+}(x_5)^2 = \{x_8,x_{10}\}$, $D_C^{CDR+}(x_6)^2 = \{x_8,x_{10}\}$, $D_C^{CDR+}(x_7)^2 = \{x_{10}\}$.

$D_C^{CDR+}(x_8)^1 = \{x_1\}, D_C^{CDR+}(x_9)^1 = \{x_1,x_2,x_3,x_6\}$, $D_C^{CDR+}(x_{10})^1 = \{x_1\}$.

$D_C^{CDR+}(x_8)^2 = \{x_8,x_{10}\}$, $D_C^{CDR+}(x_9)^2 = \{x_8,x_9,x_{10}\}$, $D_C^{CDR+}(x_{10})^2 = \{x_{10}\}$.

Lower approximations of $Cl_t^{\geqslant k}$ related to U_i, where $i,k = 1,2$, are computed as follows.

$\underline{C}(Cl_4^{\geqslant 1})^1 = \{x_1\}, \underline{C}(Cl_4^{\geqslant 2})^1 = \{x_1\}$, $\underline{C}(Cl_3^{\geqslant 1})^1 = \{x_1,x_2,x_3\}, \underline{C}(Cl_3^{\geqslant 2})^1 = \{x_1,x_2,x_4,x_7\}$, $\underline{C}(Cl_2^{\geqslant 1})^1 = \{x_1,x_2,x_3,x_4,x_7\}, \underline{C}(Cl_2^{\geqslant 2})^1 = U_1$.

$\underline{C}(Cl_4^{\geqslant 1})^2 = \emptyset, \underline{C}(Cl_4^{\geqslant 2})^2 = \{x_8,x_{10}\}$, $\underline{C}(Cl_3^{\geqslant 1})^2 = \{x_8,x_{10}\}$, $\underline{C}(Cl_3^{\geqslant 2})^2 = \{x_{10}\}$, $\underline{C}(Cl_2^{\geqslant 1})^2 = \{x_8,x_{10}\}$, $\underline{C}(Cl_2^{\geqslant 2})^2 = U_2$.

Upper approximations of $Cl_t^{\geqslant k}$ related to U_i, where $i,k = 1,2$, are computed as follows.

$\overline{C}(Cl_4^{\geqslant 1})^1 = \{x_1\}$, $\overline{C}(Cl_3^{\geqslant 1})^1 = \{x_1,x_2,x_3\}$, $\overline{C}(Cl_2^{\geqslant 1})^1 = \overline{C}(Cl_1^{\geqslant 1})^1 = U_1$

$\overline{C}(Cl_4^{\geqslant 1})^2 = \emptyset$, $\overline{C}(Cl_3^{\geqslant 1})^2 = \overline{C}(Cl_2^{\geqslant 1})^2 = \overline{C}(Cl_1^{\geqslant 1})^2 = \{x_1,x_2,x_3,x_6\}$.

$\overline{C}(Cl_3^{\geqslant 2})^1 = \overline{C}(Cl_2^{\geqslant 2})^1 = \{x_8,x_9,x_{10}\}$.

$\overline{C}(Cl_3^{\geqslant 2})^2 = \overline{C}(Cl_2^{\geqslant 2})^2 = \{x_8,x_9,x_{10}\}$.

Approximations of $Cl_t^{\geqslant}$ when U_1 and U_2 merged are updated as follows.

$\underline{C}(Cl_t^{\geqslant}) = \bigcup_{i=1}^{2}(\bigcap_{k=1}^{2}\underline{C}(Cl_t^{\geqslant k})^i) = \{x_1\}$, $\underline{C}(Cl_3^{\geqslant}) = \{x_1,x_2,x_{10}\}$, $\underline{C}(Cl_2^{\geqslant}) = \{x_1,x_2,x_3,x_4,x_7,x_8,x_{10}\}$, $\underline{C}(Cl_1^{\geqslant}) = U_1 \cup U_2$.

Approximations of $Cl_t^{\geqslant}$ when U_1 and U_2 merged are updated as follows.

$\overline{C}(Cl_4^{\geqslant}) = \bigcup_{i=1}^{2}\bigcup_{k=1}^{2}\overline{C}(Cl_4^{\geqslant k})^i = \{x_1\}, \overline{C}(Cl_3^{\geqslant}) = \{x_1,x_2,x_3,x_6,x_8,x_9,x_{10}\}, \overline{C}(Cl_2^{\geqslant}) = U, \overline{C}(Cl_1^{\geqslant}) = U$.

5. Expriment

Some experiments are conducted to evaluate the performance of the incremental approaches. Six UCI data sets are selected from UCI Machine Learning Repository[50], described in Table 2, as benchmark for performance test. The incremental algorithms and non-incremental algorithm are developed by Matlab 2014Ra on Macbook Pro 2015 with OS X EI Capitan on Intel Core i5 2.7GHz and 16G memory.

There are two classes of the experiments: experiments on updating approximations of CDRSA under the variation of attributes and experiments on updating approximations when two subsets are combined.

Table 2. Data sets description

ID	Data set	$\lvert U\rvert$	$\lvert C\rvert$	*class*
1	heart-disease(cleveland)	303	13	5
2	lonosphere	351	34	2
3	arrhythmia	452	279	12
4	winequality(white)	4898	11	7
5	statlog (Landsat Satellite)	6435	36	6
6	mushroom	7910	22	7

Algorithm 3 (IUM) Incremental **U**pdating approximations when subsets of objects are **M**erged

```
Require: approximations of $Cl_t^{\geqslant i}$ related to every $U_i$;
Ensure: $\overline{P}(Cl_t^{\geqslant}), \underline{P}(Cl_t^{\geqslant})$
for each $x \notin U_i$ do
    Compute $D_p^{CDR+}(x)^i$
end for
for i=1:$m$ do
    for each $k \neq i$ in $Cl_t^{\geqslant k}, Cl_t^{\leqslant k}$ do
        compute $\underline{P}(Cl_t^{\geqslant k})^i$ and $\overline{P}(Cl_t^{\geqslant k})^i$
        compute $\underline{P}(Cl_t^{\leqslant k})^i$ and $\overline{P}(Cl_t^{\leqslant k})^i$
    end for
end for
$\underline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m (\bigcap_{k=1}^m \underline{P}(Cl_t^{\geqslant k})^i)$ and $\underline{P}(Cl_t^{\leqslant}) = \bigcup_{i=1}^m (\bigcap_{k=1}^m \underline{P}(Cl_t^{\leqslant k})^i)$
$\overline{P}(Cl_t^{\geqslant}) = \bigcup_{i=1}^m \bigcup_{k=1}^m \overline{P}(Cl_t^{\geqslant k})^i$ and $\overline{P}(Cl_t^{\leqslant}) = \bigcup_{i=1}^m \bigcup_{k=1}^m \overline{P}(Cl_t^{\leqslant k})^i$
```

5.1. The variation of attributes

For this case, a comparison of incremental updating approximations and non-incremental computing approximations is investigated to evaluate the performance of updating approximations of CDRSA under the variation of the attribute set. We make two groups of experiments on each data sets when some attributes are deleted and added correspondingly.

In the first group of experiments, some attributes are randomly selected from the condition attribute C to constitute a deleted attribute set Q. The size of Q is from 10% to 50% in percentage in C, where the step length is 10%. The remaining attributes are $P = C - Q$. Time consuming is recorded to compare the incremental approach of updating approximations **IUD** and non-incremental method of computing approximations with respect to P when attribute set Q is deleted from C.

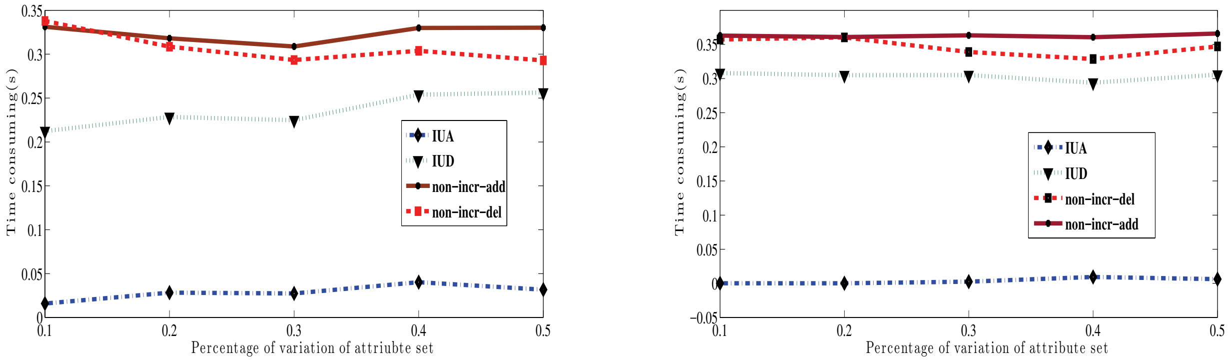
In the second group of experiments, the deleted attribute set Q are added back into P. Thus, we have $C = P \cup Q$. Similarly, computational time of incremental approach **IUA** is compared with non-incremental method when attribute set Q is added into P.

The experimental results of the six data sets are shown in Figure 1. In each of sub-figures, the x-coordinate pertains to percentage of the size of Q in C from 10% to 50%, and y-coordinate pertains the time consuming. The 'non-incr-add' denotes non-incremental method of computing approximations with respect to C, while the 'non-incr-del' denotes non-incremental computing approximations with respect to P.

From the Figure 1, the incremental approaches **IUA** and **IUD** are always faster than non-incremental approach 'non-incr-add' and 'non-incr-del' respectively. Furthermore, **IUA** is always effective than **IUD** because calculation of the confidential dominating/dominated sets in **IUA** is faster.

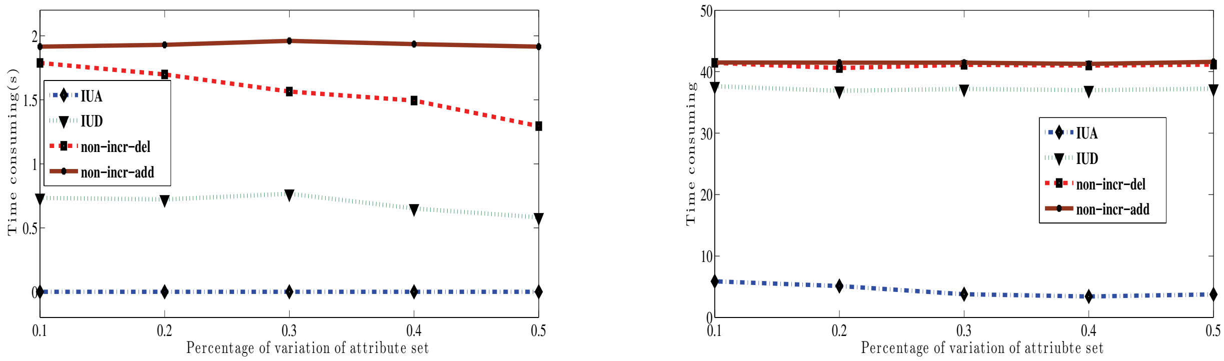
The experimental results of data sets arrhythmia and statlog (Landsat Satellite) with large condition attributes, as in Figure 1 (c) and Figure 1(e), show that the trend of time consuming of both **IUD** and 'non-incr-del' goes down when the cardinality of deleted attribute set increases.

In addition, from the results of winequality(white), statlog (Landsat Satellite) and mushroom with more objects, the trend of time consuming of **IUA** goes down when the cardinality of added attribute set increases. Intuitively, the trend of time consuming should go up when the cardinality of added attribute set increases. However, the boundary of remain attributes may decrease after the cardinality of deleted attributes increases. According to **IUA**, it follows that time consuming may decrease.



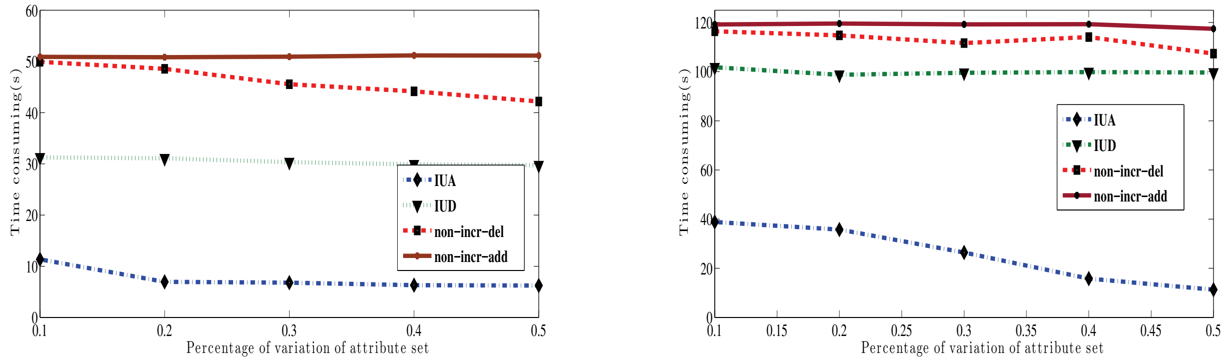
(a)heart-disease(cleveland).

(b)Ionosphere



(c)arrhythmia

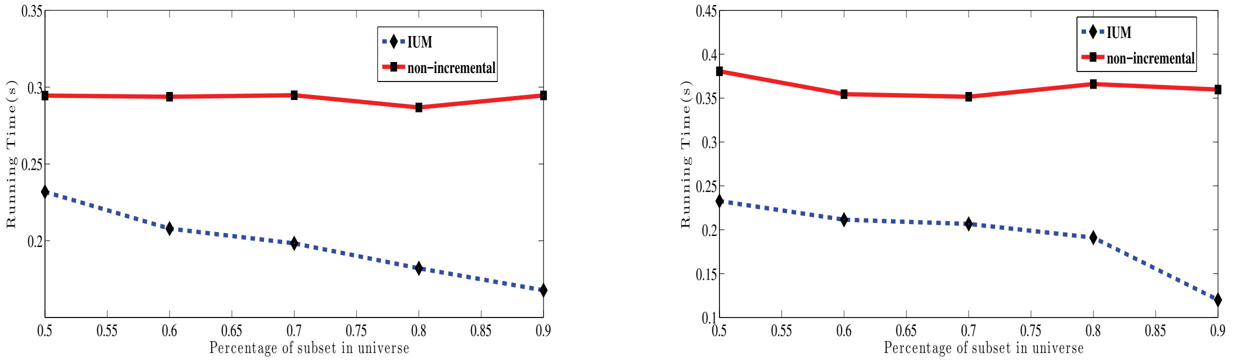
(d)winequality(white)



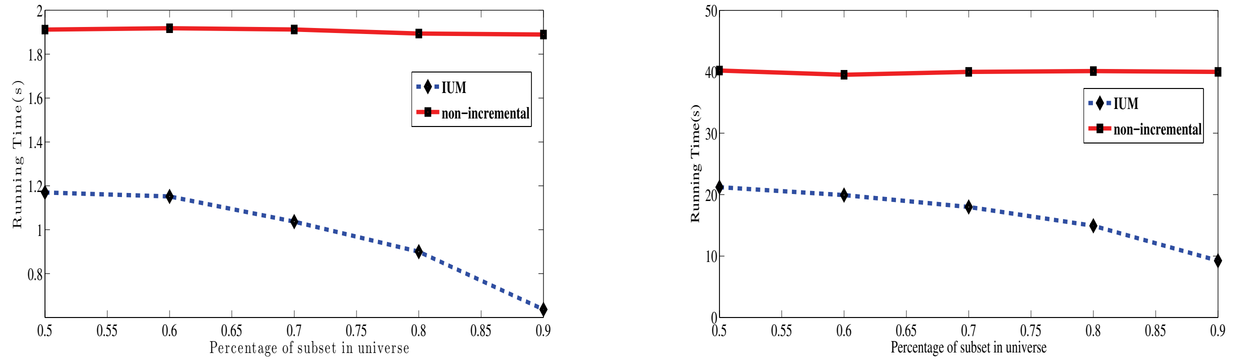
(e)statlog (Landsat Satellite)

(f)mushroom

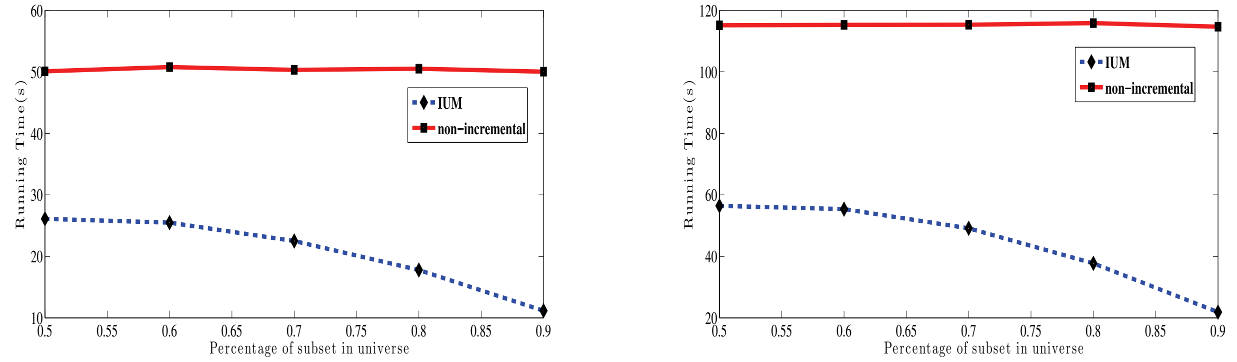
Fig. 1. A comparison of incremental and non-incremental of computing approximations under the variation of attribute set



(a)heart-disease(cleveland). (b)Ionosphere



(c)arrhythmia (d)winequality(white)



(e)statlog (Landsat Satellite) (f)mushroom

Fig. 2. A comparison of incremental and non-incremental of updating approximations when two subsets of objects are merged

5.2. Two subsets are merged

For this section, we focus on update approximations when two subsets are merged. Analogously, time consuming of computation is recorded to compare the incremental approach **IUM** with non-incremental method.

The strategy of the experiment is the universe of each data sets are divided into two subsets. A group experiments are conducted, one subset is partitioned in percentage of universe from 50% to 90% and the other is remaining.

The experimental results of the six data sets are shown in Figure 2. In each of sub-figures, the x-coordinate pertains to percentage of the one subset in the universe, and y-coordinate pertains the time consuming of updating approximations when the other subset is combined into the one.

Clearly, from Figure 2(a)-(f), **IUM** is more effective than non-incremental approach when two subsets are combined into the universe. And the trend of time consuming of **IUM** goes down when percentage of the one subset in the universe increases.

6. Conclusion and future work

An information system evolves commonly in the dynamic data environment. It is important to accelerate the calculation of approximations in CDRSA in such dynamic environment. In this paper, we focus on the incremental approaches while the variation of attributes or the objects in IODS. The properties of confidential dominating/dominated sets under the variation of attributes or objects are discussed. Furthermore, incremental approaches of updating approximations are proposed when attributes or objects are dynamic changed. With the result of experiment, incremental approaches of updating approximations of CDRSA are feasible and effectively reduce the time consuming.

Our future work will focus on speeding up these approaches on the distributed computing environment for massive data sets.

Ackownledgements

This work is supported by the National Natural Science Foundation of China (Nos. 61272060 and 61572091), National key research and development program of China (No. 2016YFB1000905), Scientific and Technological Research Program of Chongqing Municipal Education Commission(No. KJ1600933), and Scientific Research Foundation of Sichuan Provincial Education Department (No. 16ZA0329)

References

1. Z. Pawlak, Rough sets. *International Journal of Computer and Information Sciences*, 11(**5**), (1982) 341-356.
2. J. Li, Y. Ren, C. Mei, et al., A comparative study of multigranulation rough sets and concept lattices via rule acquisition, *Knowledge-Based Systems*, **85**, (2015) 1-23.
3. C. H. Weng and C. K. Huang, Knowledge discovery of customer purchasing intentions by plausible-frequent itemsets from uncertain data, *Applied Intelligence*, 43(**3**), (2015) 598-613.
4. T. R. Li, D. Ruan, Y. J. Shen, et al., A new weighting approach based on rough set theory and granular computing for road safety indicator analysis, *Computational Intelligence*, 32(**4**), (2016) 517-534.
5. R. Bello and J. L. Verdegay, Knowledge Engineering for Rough Sets Based Decision-Making Models, *International Journal of Intelligent Systems*, 29(**9**), (2014) 823-835.
6. J. T. Yao and N. Azam, Web-Based Medical Decision Support Systems for Three-Way Medical Decision Making With Game-Theoretic Rough Sets, *IEEE Transactions on Fuzzy Systems*, 23(**1**), (2015) 3-15.
7. T. R. Li, C. Luo, H. M. Chen, et al. (eds.), *The Principles and Methodologies of Big Data Mining—-From the Perspectives of Granular Computing and Rough Sets*, (Science Press, Beijing, 2016) (In Chinese).
8. J. B. Zhang, Y. Zhu, Y. Pan, et al., Efficient Parallel Boolean Matrix Based Algorithms for Computing Composite Rough Set Approximations, *Information Sciences*, **329**, (2016) 287-302.
9. L. Bellodi, M. Battaglia, M. Gasperini, et al., Rough approximation by dominance relations, *International Journal of Intelligent Systems*, 17(**2**), (2002) 153-171.
10. S. Greco, Matarazzo B and R. Słowinski, Rough sets methodology for sorting problems in presence of multiple attributes and criteria, *European Journal of Operational Research*, 138(**2**), (2002) 247-259.

11. J. Błaszczynski, S. Greco and Słowinski, Multicriteria classification : A new scheme for application of dominance-based decision rules, *European Journal of Operational Research*, 181(**3**), (2007) 1030-1044.
12. T. Lee, S. Yeung and C. Tsang, Rough sets and ordinal reducts, *Soft Computing*, 10(**1**), (2006) 27-33.
13. H. M. Chen, T. R. Li, Y. Cai, et al., Parallel attribute reduction in dominance-based neighborhood rough set, *Information Sciences*, **373**, (2016) 351-368.
14. J. J. H. Liou and G. H. Tzeng, A Dominance-based Rough Set Approach to customer behavior in the airline market, *Information Sciences*, 180(**11**), (2010) 2230-2238.
15. S. Greco, B. Matarazzo and R. Słowinski, A New Rough Set Approach to Evaluation of Bankruptcy Risk, *Operational Tools in the Management of Financial Risks*, (Springer US, 1998), pp.121-136.
16. J. Karami, A. Alimohammadi and T. Seifouri, Water quality analysis using a variable consistency dominance-based rough set approach, *Computers Environment and Urban Systems*, 43(**1**), (2014) 25-33.
17. S. Greco, B. Matarazzo and R. Słowinski, Handling Missing Values in Rough Set Analysis of Multi-attribute and Multi-criteria Decision Problems. In *Zhong, N., Skowron, A., Ohsuga, S. (eds.) RSFDGrC 1999. LNCS (LNAI), vol. 1711, Springer, Heidelberg* pp.146-157.
18. S. Greco, B. Matarazzo and R. Słowinski, Dealing with Missing Data in Rough Set Analysis of Multi-Attribute and Multi-Criteria Decision Problems, *Decision Making: Recent Developments and Worldwide Applications*, (Springer US, 2000).
19. J. Błaszczynski, R. Słowinski, (eds.) and M. Szelag, Induction of Ordinal Classification Rules from Incomplete Data, *Rough Sets and Current Trends in Computing*, (Springer Berlin Heidelberg, 2012), pp.56-65.
20. Y. Q. He and S. S. Hu, Rough analysis method of multi-attribute decision making with incomplete information, *Journal of Systems Engineering*, 19(**2**), (2004) 44-57. (in Chinese)
21. M. W. Shao and W. X. Zhang, Dominance relation and rules in an incomplete ordered information system, *International Journal of Intelligent Systems*, 20(**1**), (2005) 13-27.
22. M. L. Hu and S. F. Liu, An Analysis Method of Rough Multi-attribute Decision Making Based on Limited Extended Dominance Relation, *Systems Engineering*, 24(**4**), (2006) 106-110. (in Chinese)
23. M. L. Hu and S. F. Liu, Rough analysis method of multi-attribute decision making based on generalized extended dominance relation, *Control and Decision*, 22(**12**), (2007) 1347-1351. (in Chinese)
24. H. M. Chen, T. R. Li and D. Ruan, Maintenance of approximations in incomplete ordered decision systems while attribute values coarsening or refining, *Knowledge-Based Systems*, 31(**7**), (2012) 140-161.
25. W. S. Du and B. Q. Hu, Dominance-based rough set approach to incomplete ordered information systems, *Information Sciences*, **346-347**, (2016) 106-129.
26. X. B. Yang, J.Y.Yang, C. Wu et al., Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, *Information Sciences*, **178** (2008) 1219-1234.
27. G. Z. Luo, X. J. Yang and D. Q. Zhou, Rough Analysis Model of Multi-attribute Decision Making Based on Limited Dominance Relation, *Chinese Journal of Management Science*, 18(**4**), (2009) 391-396. (in Chinese)
28. X. B. Yang and H. L. Dou, Valued Dominance-Based Rough Set Approach to Incomplete Information System, *Data and Knowledge Engineering*, 68(**11**), (2011) 1331-1347.
29. G. L. Gou, G. Y. Wang, J. Li et al., Confidential dominance relation based rough approximation model, *Control and Decision*, 29(**7**), (2014) 1325-1329. (in Chinese)
30. C. C. Chan. A rough set approach to attribute generalization in data mining, *Information Sciences*, 107(**1-4**), (1998) 169-176.
31. T. R. Li , D. Ruan, W. Geert, et al. A rough sets based characteristic relation approach for dynamic attribute generalization in data mining, *Knowledge-Based Systems*, 2007, 20(**5**):485-494.
32. J. B. Zhang, T. R. Li, D. Ruan, et al., Rough sets based matrix approaches with dynamic attribute variation in set-valued information systems, *International Journal of Approximate Reasoning*, 53(**4**), (2012) 620-635.
33. S. Y. Li , T. R. Li and D. Liu, Incremental updating approximations in dominance-based rough sets approach under the variation of the attribute set, *Knowledge-Based Systems*, 40(**1**), (2013) 17-26.
34. Y. Y. Zhang, T. R. Li, C. Luo, et al., Incremental Updating of Rough Approximations in Interval-valued Information Systems under Attribute Generalization, *Information Sciences*. **373**, (2016) 461-475.
35. D. Liu, T. R. Li and J. B. Zhang. Incremental updating approximations in probabilistic rough sets under the variation of attributes, *Knowledge-Based Systems*, **73**, (2015) 81-96.
36. X. B. Yang, Y. Qi, H. L. Yu, et al., Updating multigranulation rough approximations with increasing of granular structures, *Knowledge-Based Systems*, **64**, (2014) 59-69.
37. N. Shan and W. Ziarko, Data-based acquisition and incremental modification of classification rules, *Computational Intelligence*, 11(**2**), (1995) 357-370.
38. Z. Zheng and G. Y. Wang, RRIA: A rough set and rule tree-based incremental knowledge acquisition algorithm, *Fundamenta Informaticae*, 59(**3**), (2004) 299-313.

39. J. B. Zhang, T. R. Li, D. Ruan, et al., A parallel method for computing rough set approximations, *Information Sciences*, 194(**5**), (2012) 209-223.
40. J. B. Zhang, T. R. Li, D. Ruan, et al., Neighborhood rough sets for dynamic data mining, *International Journal of Intelligent Systems*, 27(**4**), (2012) 317-342.
41. C. Luo, T. R. Li, Z. Yi, et al., Matrix approach to decision-theoretic rough sets for evolving data, *Knowledge-Based Systems*, **99**, (2016) 123-134.
42. C. Luo, T. R. Li, H. M. Chen, et al., Efficient updating of probabilistic approximations with incremental objects, *Knowledge-Based Systems*, **109**, (2016) 71-83.
43. S. Y. Li, T. R. Li and J. Hu, Update of approximations in composite information systems, *Knowledge-Based Systems*, **83**, (2015) 138-148.
44. S. Y. Li, T. R. Li and D. Liu, Dynamic Maintenance of Approximations in Dominance-Based Rough Set Approach under the Variation of the Object Set, *International Journal of Intelligent Systems*, 28(**8**), (2013) 729-751.
45. S. Y. Li, T. R. Li, Z Zhang, et al., Parallel computing of approximations in dominance-based rough sets approach, *Knowledge-Based Systems*, **87**, (2015) 102-111.
46. J. Błaszczynski and R.Słowinski, Incremental induction of decision rules from dominance-based rough approximations, *Electronic Notes in Theoretical Computer Science*, 82(**4**), (2003) 40-51.
47. S. Wang, T. R. Li, C. Luo, et al., Efficient updating rough approximations with multi-dimensional variation of ordered data, *Information Sciences*, **372**, (2016) 690-708.
48. A. P. Zeng, T. R. Li, J. Hu, et al., Dynamical updating fuzzy rough approximations for hybrid data under the variation of attribute values, *Information Sciences*, **378**, (2017) 363-388.
49. H. M. Chen, T. R. Li, D. Ruan, et al., A rough-set-based incremental approach for updating approximations under dynamic maintenance environments, *IEEE Transactions on Knowledge and Data Engineering*, 25(**2**), (2013) 274-284.
50. M. Lichman,, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science, (2013).