

Received July 22, 2020, accepted August 6, 2020, date of publication August 11, 2020, date of current version August 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3015813

Accelerating Update of Approximations Under a Dominance Relation

ZHIYONG HONG^{1,2}, SHAOYONG LI^{1,2}, AND LIN YU¹

¹Facility of Intelligence Manufacture, Wuyi University, Jiangmen 529020, China

²Yue-Gang-Ao Industrial Big Data Collaborative Innovation Center, Wuyi University, Jiangmen 529020, China

Corresponding author: Shaoyong Li (meterer@163.com)

This work was supported in part by the Foundation for Key Platform and Scientific Research in Higher Education of Guangdong under Grant 2018KTSCX235 and Grant 2019KTSCX189; in part by the Natural Science Foundation of Guangdong Province under Grant 2016A030310003 and Grant 2020A1515011468; and in part by the Open Project of Shanghai Key Laboratory of Information Security Integrated Management Technology, China, under Grant AGK2018006.

ABSTRACT In real-time applications, the information system may often evolves over time. Dynamic information processing is one of main challenges in the research fields related to information processing. Rough sets theory is an excellent information processing tool which has been applied in dynamic information processing researches in recent years. Dominance-based rough sets approach is an outstanding generation of rough sets theory, which can process the ordered information related to the problems of multi-criteria decision analysis and multi-criteria sort. In this article, we investigate some new strategies to unleash the performance of updating approximations in dominance-based rough sets approach further. An original concept of coarse boundary of approximations is given, which can be applied to prune more unnecessary computation in updating approximations further. Then a new incremental approach for updating approximations under a dominance relation is proposed and the corresponding algorithm is designed. A numeric illustration shows the feasibility of the approach. By extensive experimental estimation, the performance of the approach outperforms that of its counterpart on the computational time.


INDEX TERMS Rough sets, dominance relation, coarse boundary, approximation, dynamic data.

I. INTRODUCTION

In 1982, Zdzislaw Pawlak, a Polish scientist proposed the original rough sets theory (RST) [1]. The theory has two essential characteristics. The one is the granularity of universe with a discernibility relation, which produces a family of information granule. The another is to approximate the rough concepts (rough sets) by their lower and upper approximations in order to obtain certain and uncertain knowledge. In recent years, it has been proved that RST is one of most excellent information processing tools, which can be applied widely in different application fields [2] such as faults diagnose, image processing, knowledge discovery, feature selection, intelligent control systems and so on [3]–[6].

However, the original RST introduced by Pawlak can not process information with preference-ordered attributes effectively. Greco *et al.* introduced a dominance-based rough sets approach (DRSA) to solve multi-criteria decision analysis and multi-criteria sort problems [7], in which a dominance

relation was used to granulate universe instead of the discernibility relation in RST forming two families of basic information granule, i.e., dominating set and dominated set. In DRSA, the concepts approximated are upward unions and downward unions of decision classes, which are characterized by the corresponding lower and upper approximations, respectively [7]. In order to deal with errors, missing values in the domains of condition attributes and the strong inconsistency of given data, Inuiguchi *et al.* proposed a variation-precision dominance-based rough sets approach (VP-DRSA) [8] following the idea of variable precision rough sets approach (VP-RSA) introduced by Ziarko [9]. Considering the combination of the “do not care” and unknown attribute values in the incomplete information system and the preference-ordered domains of the attributes in the ordered information system, Yang *et al.* proposed a similarity dominance relation by combining a similarity relation and a dominance relation, then developed a generation of DRSA for incomplete ordered information processing [10]. Qian *et al.* considered that some of the attribute values in many real-time information systems may be set-valued to

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Hua Chen .

characterize uncertain information and missing information and summarized two types of set-valued information systems. Then they introduced two dominance relations to the two types of set-valued information systems [11]. Du *et al.* proposed a characteristic-based dominance relation under the mainframe of DRSA to process information with the losing or absenting attribute values [12]. DRSA is an useful generation of RST, which can be used to analyzing the ordered data. Hence it has been applied widely in many fields related to multi-criteria decision analysis problems solving, *e.g.*, periodic prediction, group decision, multi-criteria web mining, business indicator analysis, public security, territorial sustainable policies making, text processing and so on [13]–[18].

In fact, the information from the real-world evolves over time, which often reflects to object set, attribute set, attribute values in an information system. The evolution of these three aspects in the information system had been paid many attentions by many scholars in rough sets society [19]–[35]. Many research works had been done by based on DRSA, *e.g.*, Jerzy *et al.* realized that concept descriptions from facts incrementally refine when some new facts become available. They introduced an approach for incrementally inducing decision rules and selecting the most interesting representatives from the final set of rules under a dominance-based relation [19]. For multiple criteria decision analysis problems in the dynamic data environment, Greco *et al.* proposed an incremental strategy for updating decision rules [20]. Jia *et al.* proposed an incremental algorithm INRIDDM based on DRSA to update decision rule sets when a new object being added into the universe [21]. Computing lower and upper approximations of concepts is a key step for applying RST in data mining and knowledge discovery. The computational time taken by computing lower and upper approximations relates directly to whether RST can be used to solve some real-time problems with high timeliness. In recent years, many approach for dynamically updating lower and upper approximations of upward and downward unions of decision classes in DRSA in dynamic data environment were introduced *e.g.*, Li *et al.* analyzed the variations happened in the computation of approximations in DRSA when the information system varying and then explored the dynamic mechanism of updating approximations of DRSA by incremental learning. They proposed an incremental approaches for updating approximations of DRSA under the object set varying [23]. Luo *et al.* proposed an incremental approach for updating approximation of set-valued dominance-based rough sets approach when the set of objects varying [25]. Chen *et al.* proposed an incremental approach for update approximations of DRSA while attribute values refining or coarsening in incomplete information system [22]. Wang *et al.* propose an incremental algorithm which can efficiently update approximations of DRSA when objects and attributes increase simultaneously [28]. Huang *et al.* introduced a composite dominance-based rough sets model and redefined the lower and upper approximations of upward and downward unions of decision classes.

They proposed a matrix-based incremental approaches for the update of approximations in composite ordered decision systems while the attribute set varies over time [33]. Guo *et al.* analyzed the variations happened in equivalence classes, decision classes, conditional probability, internal grade and external grade while the set of objects evolves over time and explored the updating mechanisms for the concept approximations of two types of double-quantitative decision-theoretic rough sets models with the incremental learning technique [34].

Although there are many excellent approaches of incrementally updating approximations of rough sets, they can hardly be directly applied to update approximations of rough sets under a dominance relation when the object set evolves over time except for the one and its generalization presented in [23], [28]. The incremental approach in literature [23] can not directly deal with the cases that the insertion or deletion of multiple objects at the same time. If the insertion or deletion of multiple objects are processed in the way of single object accumulation, redundant calculation will be caused. The aim of the paper is to improve the performance of updating approximations under a dominance relation when some objects adding. An original concept of coarse boundary is introduced with respect to lower and upper approximations of downward and upward unions of decision classes in DRSA. Combined with the coarse boundaries of approximations, a new available object can be directly assigned into some approximations according to its available information description if no variations happen in the corresponding coarse boundaries of approximations. Then some unnecessary computation in approximations updating may be avoided under this case. In another case that some variations happen in coarse boundaries of approximations when an object becoming available, the unnecessary computation in updating the dominance relation may be pruned by lessening the scan range. Thus there is an alternative incremental approach that can be used to update approximations under a dominance relation. A numeric example illustrates the feasibility of the approach. A corresponding algorithm is developed to estimate the performance of the approach by several different data sets downloaded from UCI [36].

The rest of this article is organized as follows. Section II reviews the definitions of information system, dominance relation, information granule and approximations in DRSA. In Section III, the definitions and some properties of coarse boundaries of lower and upper approximations of downward and upward unions of decision classes are introduced respectively. In Section IV, we investigate that the variations may be happened on coarse boundaries in dynamic data environment. In Section V, an alternative incremental approach is introduced. An algorithm of updating approximations of DRSA is developed and analyzed in Section VI. Section VII employs a numerical example to validate the feasibility of the approach. The experimental estimations are introduced in Section VIII. The paper ends with conclusions and further research work in Section IX.

II. DOMINANCE-BASED ROUGH SETS APPROXIMATION

In this section, we review briefly some basic notations, concepts and terminologies of DRSA on the basis of references [7], [37], [38].

Definition 1: An information system can be presented formally as a four-tuple $S = (U, A, V, f)$, in which

- U is called as universe, a non-empty finite set of objects;
- A is a set of attributes, which can be divided into a condition attribute set C and a decision attribute set $\{d\}$, $A = C \cup \{d\}$, in where d is a decision attribute;
- V is regarded as the domain of all attributes;
- f is an information function that can be present as $f : U \times A \rightarrow V$ such that $f(x, a) \in V_a, \forall a \in A$ and $x \in U$, where V_a is a domain of attribute a .

$\forall a \in C$, let \succeq_a present a preference relation on universe U with respect to the criterion a . $\forall x, y \in U, x \succeq_a y$ means that x is at least as good as y with respect to the criterion a . Let $P \subseteq C$, if there exists $x \succeq_a y$ for all $a \in P$, then x dominates y with respect to P , denoted by $xDPy$. D_P is a dominance relation on universe U with respect to the set of criteria P which can be presented formally as

$$D_P = \{(x, y) \in U \times U \mid f(x, a) \geq f(y, a), \forall a \in P\}.$$

In DRSA, the basic information granules are called as the knowledge granules, which can be obtained by granulating the universe with a dominance relation. Then the universe U can be granulated by the dominance relation D_P into two families of knowledge granules under the mainframe of DRSA, which can be presented as follows.

- $D_P^+(x) = \{y \in U \mid yDPx\}$ is the set of the objects that are dominating the object x with respect to P , which is called as P -dominating set of the object x ;
- $D_P^-(x) = \{y \in U \mid xDPy\}$ is the set of the objects that are dominated by the object x with respect to P , which is called as P -dominated set of the object x .

The concepts (rough sets) approximated in DRSA are the downward and upward unions of decision classes. The universe U can be partitioned by the decision attribute d into a family of decision classes C_1, C_2, \dots, C_m . For any decision class C_n , its upward union and downward union present respectively as follows.

$$Cl_n^{\geq} = \bigcup_{n' \geq n} Cl_{n'} \tag{1a}$$

$$Cl_n^{\leq} = \bigcup_{n' \leq n} Cl_{n'} \tag{1b}$$

In equations (1a) and (1a), let $n, n' \in \{1, 2, \dots, m\}$. Cl_n^{\geq} is the upward union of decision class Cl_n which means that if $x \in Cl_n^{\geq}$ then x at least belongs to class Cl_n ; Cl_n^{\leq} is the downward union of decision class Cl_n which means that if $x \in Cl_n^{\leq}$ then x at most belongs to class Cl_n .

$\forall P \subseteq C$, the lower and upper approximations of Cl_n^{\geq} with respect to P are defined respectively as:

$$\underline{P}(Cl_n^{\geq}) = \{x \in U \mid D_P^+(x) \subseteq Cl_n^{\geq}\} \tag{2a}$$

$$\overline{P}(Cl_n^{\geq}) = \{x \in U \mid D_P^-(x) \cap Cl_n^{\geq} \neq \emptyset\} \tag{2b}$$

Analogously, the lower and upper approximations of Cl_n^{\leq} with respect to P are defined respectively as:

$$\underline{P}(Cl_n^{\leq}) = \{x \in U \mid D_P^-(x) \subseteq Cl_n^{\leq}\} \tag{3a}$$

$$\overline{P}(Cl_n^{\leq}) = \{x \in U \mid D_P^+(x) \cap Cl_n^{\leq} \neq \emptyset\} \tag{3b}$$

In RST, the lower and upper approximations of any rough sets can partition the universe into three regions as positive region, boundary region and negative region. The lower and upper approximations of the concept Cl_n^{\geq} can partition the universe U into three regions as follows.

- Positive region of Cl_n^{\geq} : $Pos_P(Cl_n^{\geq}) = \underline{P}(Cl_n^{\geq})$;
- Negative region of Cl_n^{\geq} : $Neg_P(Cl_n^{\geq}) = U - \overline{P}(Cl_n^{\geq})$;
- Boundary region of Cl_n^{\geq} : $Bn_P(Cl_n^{\geq}) = \overline{P}(Cl_n^{\geq}) - \underline{P}(Cl_n^{\geq})$.

Analogously, the universe U can be divided by the lower and upper approximations of the concept Cl_n^{\leq} into three regions as follows.

- Positive region of Cl_n^{\leq} : $Pos_P(Cl_n^{\leq}) = \underline{P}(Cl_n^{\leq})$;
- Negative region of Cl_n^{\leq} : $Neg_P(Cl_n^{\leq}) = U - \overline{P}(Cl_n^{\leq})$;
- Boundary region of Cl_n^{\leq} : $Bn_P(Cl_n^{\leq}) = \overline{P}(Cl_n^{\leq}) - \underline{P}(Cl_n^{\leq})$.

In addition, if $n > 1$ then there are the following properties.

- $Pos_P(Cl_n^{\geq}) = Neg_P(Cl_{n-1}^{\leq})$;
- $Neg_P(Cl_n^{\geq}) = Pos_P(Cl_{n-1}^{\leq})$;
- $Bn_P(Cl_n^{\geq}) = Bn_P(Cl_{n-1}^{\leq})$.

III. COARSE BOUNDARIES

Following our previous work [23], we attempt to explore some new strategies to reduce more computation than those we had. Here we will introduce some concepts of coarse boundaries of lower and upper approximations of upward unions and downward unions of decision classes, respectively.

Definition 2: $\forall a_i \in P$ and $\forall x \in \underline{P}(Cl_n^{\geq})$, let \underline{a}_i^n present the minimal value of the information function $f(x, a_i)$ in its value domain, denoted by

$$\underline{a}_i^n = \min\{f(x, a_i) \mid x \in \underline{P}(Cl_n^{\geq})\}.$$

The coarse lower boundary of $\underline{P}(Cl_n^{\geq})$ is the vector formed by \underline{a}_i^n for all a_i in P , denoted by

$$\underline{B}_P(\underline{P}(Cl_n^{\geq})) = \langle \underline{a}_1^n, \underline{a}_2^n, \dots, \underline{a}_{|P|}^n \rangle. \tag{4}$$

Similarly, $\forall x \in \overline{P}(Cl_n^{\geq})$, let \overline{a}_i^n present the minimal value of the information function $f(x, a_i)$ in its value domain, denoted by

$$\overline{a}_i^n = \min\{f(x, a_i) \mid x \in \overline{P}(Cl_n^{\geq})\}.$$

The coarse lower boundary of $\overline{P}(Cl_n^{\geq})$ is the vector formed by \overline{a}_i^n for all a_i in P , denoted by

$$\underline{B}_P(\overline{P}(Cl_n^{\geq})) = \langle \overline{a}_1^n, \overline{a}_2^n, \dots, \overline{a}_{|P|}^n \rangle. \tag{5}$$

In equations (4) and (5), $|P|$ presents the cardinality of the criterion set P . The coarse lower boundary of lower or upper approximation of an upward union of decision classes

may be seen as the description of available information of a virtual or real object which is dominated by all objects in the corresponding approximation. Then, we may regard $\underline{B}_P(\underline{P}(Cl_n^{\geq}))$ as the description of available information of an object x' . No matter whether x' is a virtual or real object, all objects assigned into $\underline{P}(Cl_n^{\geq})$ are dominating x' . If x' is a real object, then the set $\underline{P}(Cl_n^{\geq})$ is same to the set $D_P^+(x')$, i.e., $\underline{P}(Cl_n^{\geq}) = D_P^+(x')$. Otherwise, for all $x \in \underline{P}(Cl_n^{\geq})$, it is obvious that $D_P^+(x) \subset D_P^+(x')$. So does $\underline{B}_P(\underline{P}(Cl_n^{\geq}))$. Based on three regions divided from the universe U by lower and upper approximations of upward union of decision classes, the following equations can be obtained.

$$\begin{aligned} \underline{B}_P(\underline{P}(Cl_n^{\geq})) &= \underline{B}_P(Pos_P(Cl_n^{\geq})) \\ \underline{B}_P(\overline{P}(Cl_n^{\geq})) &= \underline{B}_P(Bn_P(Cl_n^{\geq})) \end{aligned}$$

For the coarse upper boundaries of lower and upper approximations of upward union of decision classes, we have

$$\overline{B}_P(\underline{P}(Cl_n^{\geq})) = \overline{B}_P(\overline{P}(Cl_n^{\geq})). \quad (6)$$

Assume that $x \in \underline{P}(Cl_n^{\geq})$, $y \in \overline{P}(Cl_n^{\geq})$ and $y \notin \underline{P}(Cl_n^{\geq})$, then $x \succeq y$ and the equation (6) holds.

Next, the definition of the coarse upper boundaries of lower and upper approximations of downward union of decision classes is introduced in the following.

Definition 3: $\forall a_i \in P$ and $\forall x \in \underline{P}(Cl_n^{\leq})$, let \underline{a}_i^n present the maximal value of the information function $f(x, a_i)$ in its value domain, denoted by

$$\underline{a}_i^n = \max\{f(x, a_i) | x \in \underline{P}(Cl_n^{\leq})\}.$$

The coarse upper boundary of $\underline{P}(Cl_n^{\leq})$ is the vector formed by \underline{a}_i^n for all a_i in P , denoted by

$$\overline{B}_P(\underline{P}(Cl_n^{\leq})) = \langle \underline{a}_1^n, \underline{a}_2^n, \dots, \underline{a}_{|P|}^n \rangle. \quad (7)$$

Similarly, $\forall x \in \overline{P}(Cl_n^{\leq})$, let \overline{a}_i^n present the maximal value of the information function $f(x, a_i)$ in its value domain, denoted by

$$\overline{a}_i^n = \max\{f(x, a_i) | x \in \overline{P}(Cl_n^{\leq})\}.$$

The coarse upper boundary of $\overline{P}(Cl_n^{\leq})$ is the vector formed by \overline{a}_i^n for all a_i in P , denoted by

$$\overline{B}_P(\overline{P}(Cl_n^{\leq})) = \langle \overline{a}_1^n, \overline{a}_2^n, \dots, \overline{a}_{|P|}^n \rangle. \quad (8)$$

The coarse upper boundary of lower or upper approximations of an downward union of decision classes may also be seen as the description of available information of a virtual or real object which is dominating all objects in the corresponding approximation. Such as, $\overline{B}_P(\underline{P}(Cl_n^{\leq}))$ may be regarded as the description of available information of a virtual or real object x' which is dominating all objects assigned into $\underline{P}(Cl_n^{\leq})$. Based on three regions divided by lower and upper approximations of downward union of decision classes, we have two equations as follows.

$$\begin{aligned} \overline{B}_P(\underline{P}(Cl_n^{\leq})) &= \overline{B}_P(Pos_P(Cl_n^{\leq})) \\ \overline{B}_P(\overline{P}(Cl_n^{\leq})) &= \overline{B}_P(Bn_P(Cl_n^{\leq})) \end{aligned}$$

For lower or upper approximations of downward unions of decision classes, we do not discuss their coarse lower boundaries any more.

Property 1: $\forall n, k \in \{1, \dots, m\}$, if $n < k$ then the following items hold.

- 1) $\underline{B}_P(\underline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\underline{P}(Cl_k^{\geq}))$;
- 2) $\underline{B}_P(\overline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\overline{P}(Cl_k^{\geq}))$;
- 3) $\overline{B}_P(\underline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\underline{P}(Cl_k^{\leq}))$;
- 4) $\overline{B}_P(\overline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\overline{P}(Cl_k^{\leq}))$.

In where \leq presents the meaning of ‘‘dominated’’.

Proof: 1) If $n < k$, then $Cl_k^{\geq} \subseteq Cl_n^{\geq}$ and $\underline{P}(Cl_k^{\geq}) \subseteq \underline{P}(Cl_n^{\geq})$. Hence, $\underline{B}_P(\underline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\underline{P}(Cl_k^{\geq}))$ holds.

Similarly, the items 2), 3) and 4) also hold. ■

Property 2: Let $i \in \{1, \dots, |P|\}$ and $\forall n, k \in \{1, \dots, m\}$, the following items hold.

- 1) If $\underline{a}_i^n \leq \underline{a}_i^k$, then $\underline{B}_P(\underline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\underline{P}(Cl_k^{\geq}))$;
- 2) If $\overline{a}_i^n \leq \overline{a}_i^k$, then $\underline{B}_P(\overline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\overline{P}(Cl_k^{\geq}))$;
- 3) If $\underline{a}_i^n \leq \underline{a}_i^k$, then $\overline{B}_P(\underline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\underline{P}(Cl_k^{\leq}))$;
- 4) If $\overline{a}_i^n \leq \overline{a}_i^k$, then $\overline{B}_P(\overline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\overline{P}(Cl_k^{\leq}))$.

Proof: 1) $\because \underline{a}_i^n \leq \underline{a}_i^k$ for all $i \in \{1, \dots, |P|\}$, then there is $\langle \underline{a}_1^n, \underline{a}_2^n, \dots, \underline{a}_{|P|}^n \rangle \leq \langle \underline{a}_1^k, \underline{a}_2^k, \dots, \underline{a}_{|P|}^k \rangle \Leftrightarrow \underline{B}_P(\underline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\underline{P}(Cl_k^{\geq}))$. \therefore The item 1) holds.

Similarly, the items 2), 3) and 4) also hold. ■

Property 3: $\forall n, k \in \{1, \dots, m\}$, the following items hold.

- 1) If $\underline{B}_P(\underline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\underline{P}(Cl_k^{\geq}))$, then $\underline{P}(Cl_k^{\geq}) \subset \underline{P}(Cl_n^{\geq})$;
- 2) If $\underline{B}_P(\overline{P}(Cl_n^{\geq})) \leq \underline{B}_P(\overline{P}(Cl_k^{\geq}))$, then $\overline{P}(Cl_k^{\geq}) \subset \overline{P}(Cl_n^{\geq})$;
- 3) If $\overline{B}_P(\underline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\underline{P}(Cl_k^{\leq}))$, then $\underline{P}(Cl_n^{\leq}) \subset \underline{P}(Cl_k^{\leq})$;
- 4) If $\overline{B}_P(\overline{P}(Cl_n^{\leq})) \leq \overline{B}_P(\overline{P}(Cl_k^{\leq}))$, then $\overline{P}(Cl_n^{\leq}) \subset \overline{P}(Cl_k^{\leq})$.

Proof: The proof of Property 3 can be proved easily by based on Properties 1 and 2 and is omitted. ■

IV. COARSE BOUNDARIES EVOLVES OVER TIME

Some variations will be happened in the steps of approximations computing when some new object being added into the universe [23]. Here we focus on variations happened on the coarse boundaries of approximations under the case when a new object becoming available. The case of some new objects becoming available simultaneously may be regarded as the accumulation of single object becoming available.

Let U be the original universe in the initial time of a dynamic course. Within the dynamic course, the new object x' becomes available. This means that the original universe U is outdated. Then we need to discovery new knowledge on the universe $U' = U \cup \{x'\}$ instead of that obtained from the universe U to ensure the timeliness of useful knowledge. The coarse boundary of an approximation can be seen as the description of the available information of an object no matter which is a virtual or real. To discuss whether variations will be happened on coarse boundaries when a new object adding into the universe, the available information description of the object is introduced in the following.

$\forall a_i \in P$, the available information of x' with respect to P can be described as

$$I(x') = \langle f(x', a_1), f(x', a_2), \dots, f(x', a_{|P|}) \rangle. \quad (9)$$

As for the coarse boundaries when a new object becoming available, we discuss unchanged and changed coarse boundaries respectively. In the rest of the paper, let $Cl'_n = Cl_n \cup \{x'\}$.

A. UNCHANGED COARSE BOUNDARIES

If $f(x', d) \geq d_n$ then x' must be assigned into Cl'_n . The coarse lower boundaries remain unchanged if satisfying the following proposition.

Proposition 1: There exists $\underline{B}(P(Cl'_n)) = \underline{B}(P(Cl_n))$ if satisfying the following items.

- 1) $f(x', d) \geq d_n$;
- 2) $\underline{B}(P(Cl_n)) \leq I(x')$;
- 3) $\overline{B}(\overline{P}(Cl_{n-1})) \leq I(x')$.

Proof: $\because f(x', d) \geq d_n \Rightarrow x' \in Cl'_n$ and $\because \overline{B}(\overline{P}(Cl_{n-1})) \leq I(x') \Rightarrow \overline{P}(Cl'_{n-1}) = \overline{P}(Cl_{n-1})$. Then, there is $\underline{P}(Cl'_n) = \underline{P}(Cl_n) \cup \{x'\}$. $\because \underline{B}(P(Cl_n)) \leq I(x')$, $\therefore \underline{B}(P(Cl'_n)) = \underline{B}(P(Cl_n))$. ■

Proposition 2: If $\underline{B}(P(Cl'_n)) \leq I(x')$, then $\underline{B}(\overline{P}(Cl'_n)) = \underline{B}(\overline{P}(Cl_n))$.

Proof: $\because \underline{B}(\overline{P}(Cl'_n)) \leq I(x') \Rightarrow \underline{B}_p(\overline{P}(Cl'_n)) = (\overline{a}_1^n, \overline{a}_2^n, \dots, \overline{a}_{|P|}^n) \Rightarrow \underline{B}(\overline{P}(Cl'_n)) = \underline{B}(\overline{P}(Cl_n))$. $\therefore \underline{B}(\overline{P}(Cl'_n)) = \underline{B}(\overline{P}(Cl_n))$. ■

As for the coarse upper boundaries remain unchanged, we obtain the following proposition.

Proposition 3: There exists $\overline{B}(P(Cl'_n)) = \overline{B}(P(Cl_n))$ if satisfying the following items.

- 1) $f(x', d) \leq d_n$;
- 2) $\overline{B}(P(Cl_n)) \geq I(x')$;
- 3) $\underline{B}(\overline{P}(Cl_{n+1})) \geq I(x')$.

In where \geq presents the meaning of ‘‘dominating’’.

Proof: The proof is similar to the proof of Proposition 1. ■

Proposition 4: If $\overline{B}(\overline{P}(Cl'_n)) \geq I(x')$, then $\overline{B}(\overline{P}(Cl'_n)) = \overline{B}(\overline{P}(Cl_n))$.

Proof: The proof is similar to the proof of Proposition 2. ■

B. CHANGED COARSE BOUNDARIES

The case that coarse boundaries will be changed when some objects becoming available is inevitable in dynamic data environment. Here we discuss the changing trends of coarse boundaries after an object adding into the universe.

Proposition 5: There exists $\underline{B}(P(Cl'_n)) \leq \underline{B}(P(Cl_n))$ if satisfying the following items.

- 1) $f(x', d) \geq d_n$;
- 2) $\underline{B}(P(Cl_n)) \geq I(x')$;
- 3) $\overline{B}(\overline{P}(Cl_{n-1})) \leq I(x')$.

Proof: $\because f(x', d) \geq d_n \Rightarrow x' \in Cl'_n$ and $\because \overline{B}(\overline{P}(Cl_{n-1})) \leq I(x') \Rightarrow \overline{P}(Cl'_{n-1}) = \overline{P}(Cl_{n-1})$. Then, there is $\underline{P}(Cl'_n) = \underline{P}(Cl_n) \cup \{x'\}$. $\because \underline{B}(P(Cl_n)) \geq I(x')$, $\therefore \underline{B}(P(Cl'_n)) \leq \underline{B}(P(Cl_n))$. ■

Proposition 6: If $f(x', d) \geq d_n$ and $\underline{B}(\overline{P}(Cl'_n)) \geq I(x')$, then $\underline{B}(\overline{P}(Cl'_n)) \leq \underline{B}(\overline{P}(Cl_n))$.

Proof: $\because f(x', d) \geq d_n \Rightarrow x' \in Cl'_n$. And $\because \underline{B}(\overline{P}(Cl'_n)) \geq I(x')$, then there is at least an information

function $f(x', a_i) \leq \overline{a}_i^n$ for $i = 1, 2, \dots, |P|$. $\therefore \underline{B}(\overline{P}(Cl'_n)) \leq \underline{B}(\overline{P}(Cl_n))$ holds. ■

Proposition 7: There exists $\overline{B}(\underline{P}(Cl'_n)) \leq \overline{B}(\underline{P}(Cl_n))$ if satisfying the following items.

- 1) $f(x', d) \leq d_n$;
- 2) $\overline{B}(\underline{P}(Cl_n)) \leq I(x')$;
- 3) $\underline{B}(\overline{P}(Cl_{n+1})) \geq I(x')$.

Proof: The proof is similar to the proof of Proposition 5. ■

Proposition 8: If $f(x', d) \leq d_n$ and $\overline{B}(\overline{P}(Cl'_n)) \leq I(x')$, then $\overline{B}(\overline{P}(Cl'_n)) \geq \overline{B}(\overline{P}(Cl_n))$.

Proof: The proof is similar to the proof of Proposition 6. ■

V. APPROXIMATIONS UPDATING

After some new objects becoming available, the knowledge discovered from the original information system may be outdated. In order to obtain the timeliness and effectiveness of knowledge, we need to recompute the approximations of RST which is applied in our problem solving. In many cases, the new available information just cause a small part of approximations to be outdated. In [23], an incremental approach for updating approximations had been introduced. However, that approach requires the beginning from computing the dominating and dominated sets of the new object over the whole universe. In addition, it also needs to update all approximations that are related to the dominating and dominated sets of the new object. Obviously, we still can prune unnecessary computation from that approach. Then, we investigate a streamlined strategy of updating approximations of DRSA in this section.

When a new object becomes available, approximations may be updated by adding it according to the following theorems.

Theorem 1: $\underline{P}(Cl'_n) = \underline{P}(Cl_n) \cup \{x'\}$ holds if satisfying the following items.

- 1) $f(x', d) \geq n$;
- 2) $I(x') \geq \overline{B}(\overline{P}(Cl_{n-1}))$;

Proof: $\because f(x', d) \geq d_n \Rightarrow x' \in Cl'_n$ and $x' \notin Cl'_{n-1}$. Then we have $\underline{P}(Cl'_n) = \{x \in U \cup \{x'\} | D_p^+(x) \subseteq Cl'_n\} = \underline{P}(Cl_n) \cup \{x \in \{x'\} | D_p^+(x) \subseteq Cl'_n\}$. And $\because I(x') \geq \overline{B}(\overline{P}(Cl_{n-1})) \Rightarrow x' \notin \overline{P}(Cl_{n-1}) \Rightarrow D_p^+(x') \cap \overline{P}(Cl_{n-1}) = \emptyset \Rightarrow D_p^+(x') \subseteq Cl'_n$. Then, there is $x' \in \underline{P}(Cl'_n)$. $\therefore \underline{P}(Cl'_n) = \underline{P}(Cl_n) \cup \{x'\}$. ■

Theorem 2: $\overline{P}(Cl'_n) = \overline{P}(Cl_n) \cup \{x'\}$ holds if satisfying the following items.

- 1) $f(x', d) \geq n$;
- 2) $I(x') \geq \overline{B}(\underline{P}(Cl_{n-1}))$;

Proof: $\because f(x', d) \geq d_n \Rightarrow x' \in Cl'_n$ and $x' \notin Cl'_{n-1}$. Then we have $\overline{P}(Cl'_n) = \{x \in U \cup \{x'\} | D_p^-(x) \cap Cl'_n \neq \emptyset\} = \overline{P}(Cl_n) \cup \{x \in \{x'\} | D_p^-(x) \cap Cl'_n \neq \emptyset\}$. And $\because I(x') \geq \overline{B}(\underline{P}(Cl_{n-1})) \Rightarrow x' \notin \underline{P}(Cl_{n-1}) \Rightarrow x' \in \overline{P}(Cl'_n)$. $\therefore \overline{P}(Cl'_n) = \overline{P}(Cl_n) \cup \{x'\}$. ■

Theorem 3: $\underline{P}(Cl'_n) = \underline{P}(Cl_n) \cup \{x'\}$ holds if satisfying the following items.

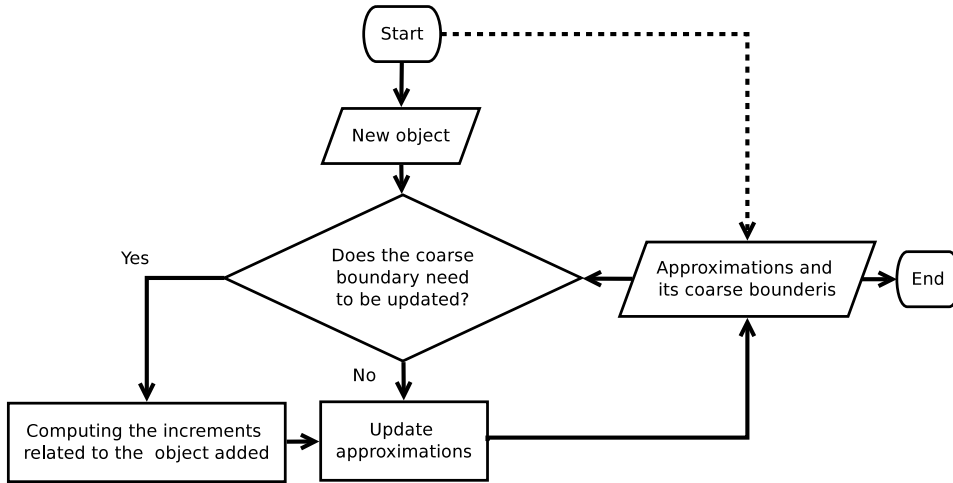


FIGURE 1. A flowchart of updating approximations in DRSA by using the coarse boundaries of approximations when new objects being inserted.

- 1) $f(x', d) \leq n$;
- 2) $I(x') \leq \underline{B}(\overline{P}(CI_{n+1}^{\geq}))$;

Proof: The proof of Theorem 3 is similar to the proof of Theorem 1 and is omitted. ■

Theorem 4: $\overline{P}(CI_n^{\leq}) = \overline{P}(CI_n^{\leq}) \cup \{x'\}$ holds if satisfying the following items.

- 1) $f(x', d) \leq n$;
- 2) $I(x') \leq \underline{B}(\underline{P}(CI_{n+1}^{\leq}))$;

Proof: The proof of Theorem 4 is similar to the proof of Theorem 2 and is omitted. ■

In the cases satisfying Theorems 1, 2, 3 and 4, approximations may be updated by comparing the description of available information of the object added to their corresponding coarse boundaries. In other cases, if these Theorems do not work at all, then we may improve the incremental approach introduced in [23] to update approximations.

According the concept of coarse boundary, we can obtain the following Remark easily.

- Remark 1:*
- 1) If $I(x') \leq \underline{B}(\underline{P}(CI_n^{\geq}))$ then $\underline{P}(CI_n^{\geq}) \subset D_P^+(x')$;
 - 2) If $I(x') \leq \underline{B}(\overline{P}(CI_n^{\geq}))$ then $\overline{P}(CI_n^{\geq}) \subset D_P^+(x')$;
 - 3) If $I(x') \geq \overline{B}(\underline{P}(CI_n^{\leq}))$ then $\underline{P}(CI_n^{\leq}) \subset D_P^-(x')$;
 - 4) If $I(x') \geq \overline{B}(\overline{P}(CI_n^{\leq}))$ then $\overline{P}(CI_n^{\leq}) \subset D_P^-(x')$.

Then Remark 2 may be used to reduce the scan range in computing the dominating or dominated sets of the object x' .

Remark 2: Let $n > 1$,

- 1) If $I(x') \leq \underline{B}(\underline{P}(CI_n^{\geq}))$ then $D_P^+(x') = \{x \in \overline{P}(CI_{n-1}^{\leq}) | x D_P x'\} \cup \underline{P}(CI_n^{\geq})$;
- 2) If $I(x') \leq \underline{B}(\overline{P}(CI_n^{\geq}))$ then $D_P^+(x') = \{x \in \underline{P}(CI_{n-1}^{\leq}) | x D_P x'\} \cup \overline{P}(CI_n^{\geq})$;

and let $n \geq 1$,

- 1) If $I(x') \geq \overline{B}(\underline{P}(CI_n^{\leq}))$ then $D_P^-(x') = \{x \in \overline{P}(CI_{n+1}^{\geq}) | x D_P x'\} \cup \underline{P}(CI_n^{\leq})$;
- 2) If $I(x') \geq \overline{B}(\overline{P}(CI_n^{\leq}))$ then $D_P^-(x') = \{x \in \underline{P}(CI_{n+1}^{\geq}) | x D_P x'\} \cup \overline{P}(CI_n^{\leq})$.

The procedure of approximations updating based on their corresponding coarse boundaries is shown in Figure 1. When a new object becomes available, at first we decide whether the coarse boundaries of approximations need to be updated according to its basic available information description. Next, for some coarse boundaries that need not to be updated, the corresponding approximations may be updated by Theorems 1, 2, 3 and 4. Otherwise, the dominating and dominated set of the new object can be computed according to Remark 2. Then we use the incremental approach [23] to update the approximations related to the new object. Finally, we obtain the updated approximations and the procedure ends.

VI. ALGORITHM

According to the strategies of updating approximations with their corresponding coarse boundaries, we develop a new incremental algorithm. The main idea of the algorithm is to reduce more computation in approximations updating than the counterpart introduced in [23]. By using the coarse boundaries, some approximations may be updated by adding the new object directly. For the cases that Theorems 1, 2, 3 and 4 do not work, the scan range in the computation of the dominating and dominated sets may be reduced by using the corresponding coarse boundaries. Combination of these strategies, Algorithm 1 may be designed for updating approximations in DRAS when the new object adding.

In Algorithm 1, the steps 3-11 and 13-15 are to update lower and upper approximations of an upward union of decision classes under the cases satisfying Theorems 1 and 2, respectively. The time complexity of these steps is $O(|P|)$. The steps 12 and 16 are to update lower and upper approximations by calling function *fun-IncAlg1* under the cases when Theorems 1 and 2 do not work, respectively. The time complexity of function *fun-IncAlg1* is $O(|U'|)$. Analogously, the steps 19-26 and 28-30 are to update lower and upper approximations of a downward union of decision classes under the cases satisfying Theorems 3 and 4, respectively.

Algorithm 1: A New Algorithm for Updating Approximations of DRSA When the Object Being Added

Input:
 1. The new object will be added x' ;
 2. Approximations and its coarse boundaries;
Output:
 The updated approximations and the updated coarse boundaries of approximations.

```

1 begin
2   forall  $n \in V_d$  do
3     if  $f(x', d) > n$  then
4       if  $I(x') \geq \underline{B}(P(CI_n^{\leq}))$  then
5         if  $I(x') \geq \underline{B}(P(CI_{n-1}^{\leq}))$  then
6            $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq}) \cup \{x'\}$ ;
7         end
8         if  $I(x') \leq \underline{B}(P(CI_{n-1}^{\leq}))$  then
9            $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq})$ ;
10        end
11       end
12       else Call fun-IncAlg1( $x'$ , 1, 1,  $\underline{P}(CI_n^{\leq})$ );
13       if  $I(x') \geq \underline{B}(P(CI_{n-1}^{\leq}))$  then
14          $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq}) \cup \{x'\}$ 
15       end
16       else Call fun-IncAlg1( $x'$ , 1, 0,  $\underline{P}(CI_n^{\leq})$ );
17     end
18     else if  $f(x', d) < n$  then
19       if  $I(x') \leq \underline{B}(P(CI_n^{\leq}))$  then
20         if  $I(x') \leq \underline{B}(P(CI_{n+1}^{\leq}))$  then
21            $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq}) \cup \{x'\}$ ;
22         end
23         if  $I(x') \geq \underline{B}(P(CI_{n+1}^{\leq}))$  then
24            $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq})$ ;
25         end
26       end
27       else Call fun-IncAlg1( $x'$ , 0, 1,  $\underline{P}(CI_n^{\leq})$ );
28       if  $I(x') \leq \underline{B}(P(CI_{n+1}^{\leq}))$  then
29          $\underline{P}(CI_n^{\leq}) \leftarrow \underline{P}(CI_n^{\leq}) \cup \{x'\}$ 
30       end
31       else Call fun-IncAlg1( $x'$ , 0, 0,  $\underline{P}(CI_n^{\leq})$ );
32     end
33   end
34   Do the steps 4-16 and 19-31;
35 end
36 end
    
```

Their time complexity also is $O(|P|)$. The steps 27 and 31 are to update the corresponding approximations by calling function *fun-IncAlg1* under the cases when Theorems 3 and 4 do not work, respectively. In addition, the steps 33-35 are to update approximations when $f(x', d) = d_n$.

The function *fun-IncAlg1* is designed based on the improvement of the incremental approach introduced in [23]. Due to reduce the scan range in computing the dominating

Function fun-IncAlg1(x, i, j, X)

```

1 Compute  $D_P^+(x)$  and  $D_P^-(x)$  /* According to Remark 2. */;
2 Compute  $l_P(x)$  and  $u_P(x)$  /*  $l_P(x)$  and  $u_P(x)$  were defined in Definition 3 in [23]. */;
3 if  $i == 1$  then
4   if  $j == 1$  then
5     Compute  $M$  /*  $M$  is  $M_n$  that was defined in Proposition 4 in [23]. */;
6      $X \leftarrow X - M$ ;
7   end
8   else
9      $X \leftarrow X \cup D_P^+(x)$ ;
10  end
11 end
12 else
13   if  $j == 1$  then
14     Compute  $X$  /*  $X$  is  $X_n$  that was defined in Proposition 2 in [23]. */;
15      $X \leftarrow X - X$ ;
16   end
17   else
18      $X \leftarrow X \cup D_P^-(x)$ ;
19   end
20 end
21 return  $X$ ;
    
```

and dominated sets, the function *fun-IncAlg1* reduce unnecessary computation further. Then, even though the time complexity of function *fun-IncAlg1* is same to the corresponding steps in the incremental algorithm designed in [23], it is better to certain extent.

From the analysis above, the time complexity of Algorithm 1 is $O(|V_d|(|P| + |U'|))$. Algorithm 1 may reduce more unnecessary computation in approximations updating than its counterpart, the Algorithm 2 developed in [23].

VII. A NUMERIC ILLUSTRATION

Example 1: Table 1 shows a information system, in which $U = \{u_1, u_2, \dots, u_{16}\}$, $A = \{a_1, a_2, d\}$, $C = \{a_1, a_2\}$, $V_{a_1} = V_{a_2} = \{1, 2, \dots, 7\}$ and $V_d = \{1, 2, 3\}$.

TABLE 1. An information table.

U	a_1	a_2	d	U	a_1	a_2	d
u_1	1	2	1	u_9	6	4	3
u_2	2	3	1	u_{10}	3	5	2
u_3	3	1	2	u_{11}	5	5	3
u_4	3	3	1	u_{12}	7	5	3
u_5	4	2	1	u_{13}	2	6	2
u_6	5	2	2	u_{14}	5	6	2
u_7	6	3	2	u_{15}	6	6	3
u_8	4	4	2	u_{16}	6	7	3

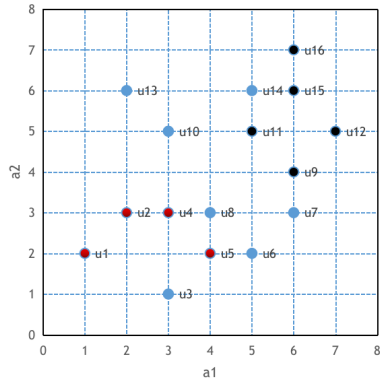


FIGURE 2. The scatter chart of the information system with respect to attributes a_1 and a_2 .

According to condition attributes a_1 and a_2 , we can draw a scatter chart as Figure 2.

In Figure 2, the red, blue and black points present the objects whose decision attribute values are 1, 2 and 3, respectively.

Let $P = \{a_1, a_2\}$, the lower and upper approximations of upward union and downward union of decision classes can be calculated, respectively. The results are listed as follows.

$$\begin{aligned} \underline{P}(Cl_3^{\geq}) &= \{u_9, u_{12}, u_{15}, u_{16}\}; \\ \overline{P}(Cl_3^{\geq}) &= \{u_9, u_{11}, u_{12}, u_{14}, u_{15}, u_{16}\}; \\ \underline{P}(Cl_2^{\geq}) &= \{u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, u_{16}\}; \\ \overline{P}(Cl_2^{\geq}) &= \{u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, \\ &\quad u_{15}, u_{16}\}; \\ \underline{P}(Cl_1^{\leq}) &= \{u_1, u_2\}; \\ \overline{P}(Cl_1^{\leq}) &= \{u_1, u_2, u_3, u_4, u_5\}; \\ \underline{P}(Cl_2^{\leq}) &= \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_{10}, u_{13}\}; \\ \overline{P}(Cl_2^{\leq}) &= \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_{10}, u_{11}, u_{12}, u_{13}, \\ &\quad u_{14}\}. \end{aligned}$$

The coarse boundaries of approximations above are listed as follows.

$$\begin{aligned} \underline{B}(P(Cl_3^{\geq})) &= \langle 6, 4 \rangle; & \underline{B}(\overline{P}(Cl_3^{\geq})) &= \langle 5, 4 \rangle; \\ \underline{B}(P(Cl_2^{\geq})) &= \langle 2, 2 \rangle; & \underline{B}(\overline{P}(Cl_2^{\geq})) &= \langle 2, 1 \rangle; \\ \underline{B}(P(Cl_1^{\leq})) &= \langle 2, 3 \rangle; & \underline{B}(\overline{P}(Cl_1^{\leq})) &= \langle 4, 3 \rangle; \\ \underline{B}(P(Cl_2^{\leq})) &= \langle 6, 6 \rangle; & \underline{B}(\overline{P}(Cl_2^{\leq})) &= \langle 6, 6 \rangle. \end{aligned}$$

Assumed that a new object x' is added into the information system, which $I(x') = \langle 4, 4 \rangle$ and $f(x', d) = 3$. The scatter chart is updated as Fig.3.

It is not difficult to find that $\underline{B}(\overline{P}(Cl_3^{\geq}))$ need to be updated, and the rest of coarse boundaries do not change after the object x' being added. Then the updated coarse boundaries are listed as follows.

$$\begin{aligned} \underline{B}(P(Cl_3^{\geq})) &= \langle 6, 4 \rangle; & \underline{B}(\overline{P}(Cl_3^{\geq})) &= \langle 4, 4 \rangle; \\ \underline{B}(P(Cl_2^{\geq})) &= \langle 2, 2 \rangle; & \underline{B}(\overline{P}(Cl_2^{\geq})) &= \langle 2, 1 \rangle; \\ \underline{B}(P(Cl_1^{\leq})) &= \langle 2, 3 \rangle; & \underline{B}(\overline{P}(Cl_1^{\leq})) &= \langle 4, 3 \rangle; \end{aligned}$$

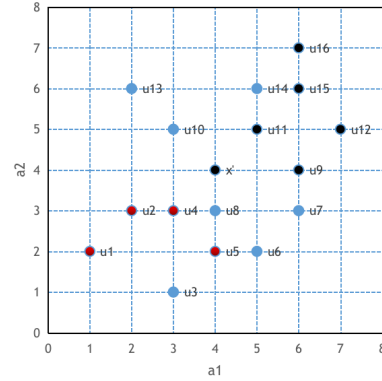


FIGURE 3. The scatter chart of the information system with the object x' .

$$\underline{B}(P(Cl_2^{\leq})) = \langle 6, 6 \rangle; \quad \underline{B}(\overline{P}(Cl_2^{\leq})) = \langle 6, 6 \rangle.$$

The lower and upper approximations of upward union and downward union of decision classes can be updated, respectively. The results are listed as follows.

$$\begin{aligned} \underline{P}(Cl_3^{\geq}) &= \{u_9, u_{12}, u_{15}, u_{16}\}; \\ \overline{P}(Cl_3^{\geq}) &= \{x', u_9, u_{11}, u_{12}, u_{14}, u_{15}, u_{16}\}; \\ \underline{P}(Cl_2^{\geq}) &= \{x', u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, u_{14}, u_{15}, \\ &\quad u_{16}\}; \\ \overline{P}(Cl_2^{\geq}) &= \{x', u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}, u_{11}, u_{12}, u_{13}, \\ &\quad u_{14}, u_{15}, u_{16}\}; \\ \underline{P}(Cl_1^{\leq}) &= \{u_1, u_2\}; \\ \overline{P}(Cl_1^{\leq}) &= \{u_1, u_2, u_3, u_4, u_5\}; \\ \underline{P}(Cl_2^{\leq}) &= \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_{10}, u_{13}\}; \\ \overline{P}(Cl_2^{\leq}) &= \{x', u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_{10}, u_{11}, u_{12}, \\ &\quad u_{13}, u_{14}\}. \end{aligned}$$

VIII. EXPERIMENTAL EVALUATIONS

In order to estimate the performance of Algorithm 1, in this section some experiments are designed and executed. These experiments are to compare the computational time of updating approximations of DRSA taken by it and its counterpart, the incremental algorithm designed in [23]. Here we call them as *BAlgorithm* and *IAlgorithm*, respectively. The platform of experiments is a personal computer with GNU/Linux Debian 9 (Stretch) and Intel(R) Core(TM) i5 CPU and 4 GB memory. *BAlgorithm* and *IAlgorithm* are coded in GNU Octave, version 4.0.3. The Datasets of experiments are downloaded from machine learning data repository, University of California at Irvine [36]. The Datasets include Abalone, Car Evaluation, Ionosphere and Sonar. The basic information of them are listed in Table 2.

For satisfying the experimental requirements, we select some attributes from these Datasets to act as the condition criteria and the decision criterion, respectively. Such as Dataset Ablone, we select Length, Diameter, Height, Whole weight, Shucked weight, Shucked weight and Shell weight as the condition criteria, meanwhile, we select Rings as the decision

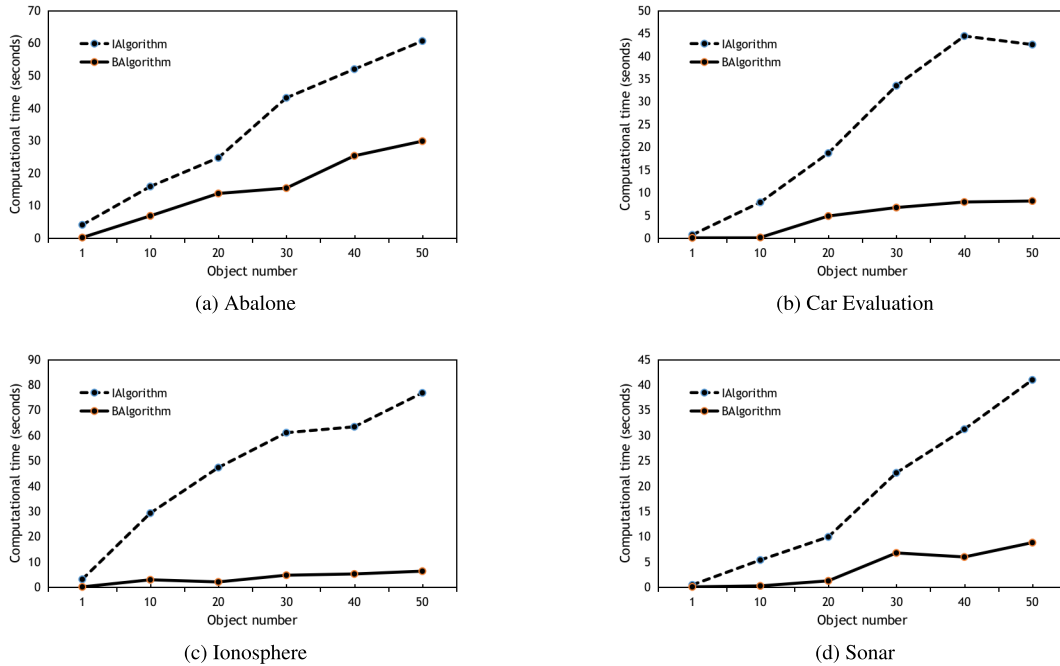


FIGURE 4. Comparison of the computational time taken by *IAlgorithm* and *BAlgorithm* versus the objects added on four different datasets.

TABLE 2. The basic information of four datasets in the experiments.

Data Set	Objects	Attributes	Classes
Abalone	4177	8	29
Car Evaluation	1728	6	4
Ionosphere	351	34	2
Sonar	208	60	2

criterion. The similar operations has been done on other Datasets.

In order to obtain the more convincing experimental results, we scaled Car Evaluation two times, Ionosphere and Sonar five times to better results in the experiments, respectively. We divided four Datasets of experiment into five subsets of same size and called them as Dataset 1, . . . , Dataset 5, respectively. For each of these four experiment Datasets, we united Dataset 1, Dataset 2, Dataset 3 and Dataset 4 as the training set to compute approximations and the corresponding coarse boundaries. We select randomly 1, 10, 20, 30, 40 and 50 objects from Dataset 5 as the new objects being added, respectively. *IAlgorithm* and *BAlgorithm* were employed to update approximations of DRSA when the new objects adding. The trends of computational time taken by two algorithms are shown respectively in Figure 4.

Figure 4 shows the trends of the computation time taken by *IAlgorithm* and *BAlgorithm* on four experimental Datasets, respectively. In each of four sub-figures, the horizontal axis is related to the number of objects added while the vertical axis concerns the computational time. From Figure 4, one can see that the computational time taken by *IAlgorithm* and *BAlgorithm* increase monotonically with the increasing number of objects added. As shown in each sub-figure of Figure 4,

BAlgorithm is always faster than *IAlgorithm*. Moreover, the difference of computational time taken by *IAlgorithm* and *BAlgorithm* enlarges with the increasing of the number of objects added.

The results from experiments show that the performance of *BAlgorithm* outperforms that of *IAlgorithm* in updating approximations in dynamic data environment. Hence, *BAlgorithm* is more effective than *IAlgorithm* for updating approximations of DRSA when some objects becoming available.

IX. CONCLUSION AND FUTURE WORK

How to obtain knowledge effectively and timely from the dynamic data is a hot issue in the research field related to data mining, which has been paid highly attention in recent years. The ordered data is often used to present people’s preference in daily lives. The analysis of ordered data is one of main tasks for multi-criteria decision making and multi-criteria sort. DRSA is an effective tool to the ordered data mining. Approximations computing is one of main steps for applying RST in data mining and knowledge discovery, which is the starting point to attribute reduction and rule extraction. In this article, an alternative incremental approach for updating approximations of DRSA when some new objects becoming available. The main innovation is to give the concept of coarse boundaries of approximations of DRSA. According to the coarse boundaries, we can reduce more unnecessary computation in approximations updating further. By the experimental estimation and the numeric illustration, we can draw conclusions as follows: (1) The approach has the feasibility in updating approximations of DRSA; (2)The performance of the approach outperforms that of its counterpart;

(3) The approach can be applied in updating approximations of DRSA when multiple objects adding. Due to the incremental algorithm can not directly be used to massive big processing, then our future work is to design the corresponding parallel algorithm for updating approximations of DRSA on cluster and multi-core platforms.

REFERENCES

- [1] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, Oct. 1982.
- [2] C. Zhang, D. Li, X. Kang, D. Song, A. K. Sangaiah, and S. Broumi, "Neutrosophic fusion of rough set theory: An overview," *Comput. Ind.*, vol. 115, pp. 103–117, 2020.
- [3] Q. Zhang, Q. Xie, and G. Wang, "A survey on rough set theory and its applications," *CAAI Trans. Intell. Technol.*, vol. 1, no. 4, pp. 323–333, Oct. 2016.
- [4] C. Wang, Y. Wang, M. Shao, Y. Qian, and D. Chen, "Fuzzy rough attribute reduction for categorical data," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 5, pp. 818–830, May 2020, doi: [10.1109/TFUZZ.2019.2949765](https://doi.org/10.1109/TFUZZ.2019.2949765).
- [5] C. Wang, Y. Shi, X. Fan, and M. Shao, "Attribute reduction based on k-nearest neighborhood rough sets," *Int. J. Approx. Reasoning*, vol. 106, pp. 18–31, Mar. 2019, doi: [10.1016/j.ijar.2018.12.013](https://doi.org/10.1016/j.ijar.2018.12.013).
- [6] C. Wang, Y. Huang, M. Shao, Q. Hu, and D. Chen, "Feature selection based on neighborhood self-information," *IEEE Trans. Cybern.*, early access, Jul. 9, 2019, doi: [10.1109/TCYB.2019.2923430](https://doi.org/10.1109/TCYB.2019.2923430).
- [7] S. Greco, B. Matarazzo, and R. Slowinski, "Rough sets theory for multicriteria decision analysis," *Eur. J. Oper. Res.*, vol. 129, no. 1, pp. 1–47, Feb. 2001.
- [8] M. Inuiguchi, Y. Yoshioka, and Y. Kusunoki, "Variable-precision dominance-based rough set approach and attribute reduction," *Int. J. Approx. Reasoning*, vol. 50, no. 8, pp. 1199–1214, Sep. 2009.
- [9] W. Ziarko, "Variable precision rough set model," *J. Comput. Syst. Sci.*, vol. 46, no. 1, pp. 39–59, Feb. 1993.
- [10] X. Yang, J. Yang, C. Wu, and D. Yu, "Dominance-based rough set approach and knowledge reductions in incomplete ordered information system," *Inf. Sci.*, vol. 178, no. 4, pp. 1219–1234, Feb. 2008.
- [11] Y. Qian, C. Dang, J. Liang, and D. Tang, "Set-valued ordered information systems," *Inf. Sci.*, vol. 179, no. 16, pp. 2809–2832, Jul. 2009.
- [12] W. S. Du and B. Q. Hu, "Dominance-based rough set approach to incomplete ordered information systems," *Inf. Sci.*, vols. 346–347, pp. 106–129, Jun. 2016.
- [13] S. Chakhar, A. Ishizaka, A. Labib, and I. Saad, "Dominance-based rough set approach for group decisions," *Eur. J. Oper. Res.*, vol. 251, no. 1, pp. 206–224, May 2016.
- [14] A. B. G. D. Couto and L. F. A. M. Gomes, "Multi-criteria Web mining with DRSA," *Procedia Comput. Sci.*, vol. 91, pp. 131–140, Jan. 2016.
- [15] Couto and A. B. G. Do, "Using a dominance-based rough set approach for analysing business indicators," *Procedia Comput. Sci.*, vol. 55, pp. 350–359, Dec. 2015.
- [16] C. M. D. M. Mota, C. J. J. D. Figueiredo, and D. V. E. S. Pereira, "Identifying areas vulnerable to homicide using multiple criteria analysis and spatial analysis," *Omega*, Feb. 2020, Art. no. 102211. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S030504831930547X>
- [17] P. Ferretti, M. B. Zolin, and G. Ferraro, "Relationships among sustainability dimensions: Evidence from an alpine area case study using dominance-based rough set approach," *Land Use Policy*, vol. 92, Mar. 2020, Art. no. 104457. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S026483771931720X>
- [18] R. Cekik and A. K. Uysal, "A novel filter feature selection method using rough set for short text data," *Expert Syst. Appl.*, vol. 160, Dec. 2020, Art. no. 113691. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417420305157>
- [19] B. Jerzy and R. Slowinski, "Incremental induction of decision rules from dominance-base rough approximations," *Electron. Notes Theor. Comput. Sci.*, vol. 82, pp. 40–45, 2003.
- [20] S. Greco, R. Slowinski, J. Stefanowski, and M. Zurawski, "Incremental versus non-incremental rule induction for multicriteria classification," in *Transactions on Rough Sets II*, 2004, pp. 33–53.
- [21] X. Jia, L. Shang, J. Chen, and X. Dai, "Incremental versus non-incremental: Data and algorithms based on ordering relations," *Int. J. Comput. Intell. Syst.*, vol. 4, no. 1, pp. 112–122, Feb. 2011.
- [22] H. Chen, T. Li, and D. Ruan, "Maintenance of approximations in incomplete ordered decision systems while attribute values coarsening or refining," *Knowl.-Based Syst.*, vol. 31, pp. 140–161, Jul. 2012.
- [23] S. Li, T. Li, and D. Liu, "Dynamic maintenance of approximations in dominance-based rough set approach under the variation of the object set," *Int. J. Intell. Syst.*, vol. 28, no. 8, pp. 729–751, Aug. 2013.
- [24] S. Li and T. Li, "Incremental update of approximations in dominance-based rough sets approach under the variation of attribute values," *Inf. Sci.*, vol. 294, pp. 348–361, Feb. 2015.
- [25] C. Luo, T. Li, H. Chen, and D. Liu, "Incremental approaches for updating approximations in set-valued ordered information systems," *Knowl.-Based Syst.*, vol. 50, pp. 218–233, Sep. 2013.
- [26] C. Luo, T. Li, and H. Chen, "Dynamic maintenance of approximations in set-valued ordered decision systems under the attribute generalization," *Inf. Sci.*, vol. 257, pp. 210–228, Feb. 2014.
- [27] C. Luo, T. Li, H. Chen, and L. Lu, "Fast algorithms for computing rough approximations in set-valued decision systems while updating criteria values," *Inf. Sci.*, vol. 299, pp. 221–242, Apr. 2015.
- [28] S. Wang, T. Li, C. Luo, and H. Fujita, "Efficient updating rough approximations with multi-dimensional variation of ordered data," *Inf. Sci.*, vol. 372, pp. 690–708, Dec. 2016.
- [29] X. Yang, T. Li, H. Fujita, D. Liu, and Y. Yao, "A unified model of sequential three-way decisions and multilevel incremental processing," *Knowl.-Based Syst.*, vol. 134, pp. 172–188, Oct. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950705117303532>
- [30] Y. Jing, T. Li, H. Fujita, Z. Yu, and B. Wang, "An incremental attribute reduction approach based on knowledge granularity with a multi-granulation view," *Inf. Sci.*, vol. 411, pp. 23–38, Oct. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0020025516308775>
- [31] Y. Huang, T. Li, C. Luo, H. Fujita, and S.-J. Horng, "Dynamic fusion of multisource interval-valued data by fuzzy granulation," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3403–3417, Dec. 2018.
- [32] W. Shu, W. Qian, and Y. Xie, "Incremental feature selection for dynamic hybrid data using neighborhood rough set," *Knowl.-Based Syst.*, vol. 194, Apr. 2020, Art. no. 105516. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950705120300289>
- [33] Q. Huang, T. Li, Y. Huang, X. Yang, and H. Fujita, "Dynamic dominance rough set approach for processing composite ordered data," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104829. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950705119303041>
- [34] Y. Guo, E. C. C. Tsang, M. Hu, X. Lin, D. Chen, W. Xu, and B. Sang, "Incremental updating approximations for double-quantitative decision-theoretic rough sets with the variation of objects," *Knowl.-Based Syst.*, vol. 189, Feb. 2020, Art. no. 105082. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0950705119304642>
- [35] P. Ni, S. Zhao, X. Wang, H. Chen, C. Li, and E. C. C. Tsang, "Incremental feature selection based on fuzzy rough sets," *Inf. Sci.*, vol. 536, pp. 185–204, Oct. 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0020025520303467>
- [36] D. Dua and C. Graff. (2017). *UCI Machine Learning Repository*. [Online]. Available: <http://archive.ics.uci.edu/ml>
- [37] S. Greco, B. Matarazzo, and R. Slowinski, "Rough approximation by dominance relations," *Int. J. Intell. Syst.*, vol. 17, no. 2, pp. 153–171, Feb. 2002.
- [38] W. Kotłowski, K. Dembczyński, S. Greco, and R. Słowiński, "Stochastic dominance-based rough set model for ordinal classification," *Inf. Sci.*, vol. 178, no. 21, pp. 4019–4037, Nov. 2008.



ZHIYONG HONG received the B.S. and M.S. degrees in computer science and technology from the Shenyang Institute of Technology, Shenyang, China, in 2001 and 2004, respectively, and the Ph.D. degree in computer science and technology from Southwest Jiaotong University, Chengdu, China, in 2014. From 2006 to 2014, he was a Lecturer with the School of Computer, Wuyi University, Jiangmen, China. Since 2015, he has been an Associate Professor. His research interests include intelligent information processing, block chain, and rough set theory.



SHAORYONG LI received the Ph.D. degree in computer science and technology from Southwest Jiaotong University, Chengdu, China, in 2014. He is currently an Associate Professor with the Facility of Intelligence Manufacture, Wuyi University, Jiangmen, China. His research interests include data mining, parallel computing, and rough set theory.



LIN YU received the B.S. degree in computer science and technology from the Huali College Guangdong University of Technology, Guangzhou, China, in 2019. He is currently pursuing the M.S. degree in computer science and technology with the Facility of Intelligence Manufacture, Wuyi University, Jiangmen, China. His research interests include data mining and rough set theory.

...