Matrix Method for Computing Approximations in Variable Consistency Dominance-Based Rough Set Approach

Lei Wang¹,², Ziqi Wang³, Jun Ye¹,²

¹College of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, Jiangxi, China
²Jiangxi Provincial Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing, Nanchang 330099, Jiangxi, China
³College of Science, Beijing Forestry University, Beijing 100083, China

*Corresponding author(E-mail: ezhoulei@163.com)

Abstract
Lower and upper approximations of upward (downward) union of decision classes in Variable Consistency Dominance-Based Rough Set Approach (VC-DRSA) are the basis for extraction of decision rules, so computations of approximations is a necessary step in knowledge representing and reduction based on VC-DRSA. In this paper, a matrix-based method for computing lower approximations of upward and downward union of decision classes at consistency level is proposed. Firstly, concepts such as dominance relation matrix, consistency level, Boolean column vector of subset and the comparison operator between matrices are introduced, then a new method for computing approximations in VC-DRSA given from matrix perspective and its correctness are proved theoretically. Furthermore, the corresponding algorithms are put forward. Finally, the matrix-based algorithms are applied to the calculation of approximations on UCI datasets and its results are compared with a previous non-matrix algorithm. The experimental results demonstrate the feasibility, conciseness and validity of the proposed matrix-based method.

Key words: Rough Sets, Variable Consistency Dominance-based Rough Set Approach, Rough Approximation, Matrix method.

1. INTRODUCTION
Rough sets theory (RST, for short), which put forward by Pawlak in 1982 (Pawlak, 1982), is proved to be a very useful mathematic tool as well as an information processing tool, owing to its wide range of applications in scientific and technological fields such as data mining, machine learning, pattern recognition, clinical diagnosis and so on (Chen and Ziarko, 2011; Qian, Liang, and Pedrycz, 2011). The key idea of RST is that any subset of universe can be characterized by two precise sets. Many scholars are engaged in the researches on RST and its applications in various fields. RST has some drawbacks, however, one is the preference-ordered on attributes domains (The attribute involves preference-ordered on domains is called criterion) is not considered, so RST is not able to handle the problems with preferential information. In this problem, there exists a preference relation on the sets of objects with respect to a particular attribute. For this reason, Greco et al. have proposed an extension method of RST called dominance-based rough set approach (DRSA, for short) to deal with the above preferential information (Greco, Matarazzo, and Slowinski, 1999). The main change of DRSA lies in the substitution of the indiscernibility relation by a dominance relation, and then the DRSA was applied to multi-criteria decision analysis (MCDA, for short) and knowledge acquisition (Greco, Matarazzo, and Slowinski, 1999; Xu, Zhang, and Zhang, 2009; Hu, Gao, and Yu, 2010; Greco, Matarazzo, and Slowinski, 2001; Shao and Zhang, 2015) such as sorting, choice and ranking problems. Analysis of many real-world information systems shows, however, that the definitions of rough approximations in DRSA are excessively restrictive. In consequence, the restriction undoubtedly influences its applications in real-world information processing problems. An extension approach of DRSA (Greco, Matarazzo, Slowinski and et al, 2001) which called variable consistency dominance-based rough set approach (VC-DRSA, for short) are born at the right moment to cope with such kind of problems (Blaszczynski, Greco, and Slowinski, 2007).

Lower and upper approximations are a pair of key concepts in all rough set approach, they are the basis of rule acquisition and attribute reduction, and so how to compute the approximations efficiently and fast has always been the focus topic in research fields of rough set theory, there is no exception to the VC-DRSA. To our best knowledge, computing approximations in VC-DRSA by matrix method has not yet discussed so far. In this paper, we focus on the calculations of approximations in VC-DRSA from matrix perspective. Matrix-based algorithms are proposed by considering the variable inconsistency in VC-DRSA. Experimental results indicate that the proposed algorithms have better performance than the traditional non-matrix algorithms.
The paper is organized as follows. In section 2, the related notions concerned to the DRSA and VC-DRSA are briefly introduced. The matrix-based method for computing approximations is focus on Section 3. Section 4 contains description and analysis of the corresponding matrix-based algorithms. Experimental evaluation on UCI datasets as well as the results analysis and comparison are presented in section 5. The last section concludes the research work and outlooks the researches in the future.

2. PRELIMINARIES

In this section, some basic concepts as well as the notations concerned with DRSA and VC-DRSA are introduced briefly.

2.1. Dominance Relation and Dominance-Based Rough Set Approach

**Definition 1.** (Liu, 2001) An Information system is a quadruple $SI = (U, A, V, f)$, where $U$ is a non-empty finite set of objects, which called universe. $A = C \cup \{d\}$ and $C \cap \{d\} = \emptyset$, $C$ is a non-empty finite set of condition attributes, $d$ is the decision attribute. $V = \bigcup_{a \in A} V_a$. $V_a$ is the domain of attribute $a$. $f : U \times A \rightarrow V$ is an information function such that: $\forall a \in A, x \in U$, $f(x, a) \in V_a$.

Considering the information system in real-world applications is with preference-ordered attribute domains, all the attributes may be further divided into two categories, one category is called criterion which values have ordinal properties, while another category is called regular attribute. Assuming that $a \in A$ is a criterion, $\forall x, y \in U, x \succeq y$ means “$x$ is at least as good as $y$ with respect to criterion $a$.” For a nonempty finite criteria set $P$, $P \subseteq C$. $x \succeq y$ means $\forall a \in P, x \succeq y$, showing that $x$ dominates $y$ with respect to criteria $P$. So $D_x$ is a dominance relation on universe $U$ with respect to the criteria set $P$.

**Definition 2.** (Greco, Matarazzo, and Slowinski, 2001; Greco, Matarazzo, and Slowinski, 2002) P-dominated set is a set of objects dominating $x$, denoted as:

$$D_x^+(y) = \{ y \in U | y D_x y \}$$

(1)

P-dominated set is a set of objects dominated by $x$, denoted as:

$$D_x^-(y) = \{ y \in U | y D_x y \}$$

(2)

Assume $Cl = U / Ind(d) = \{ Cl_n | n \in I \}$ be a set of decision classes partitioned by attribute $d$. $I = \{1, 2, \cdots, m\}$. $\forall r, s \in I$ such that $r > s$, the objects from $Cl_r$ dominate the objects from $Cl_s$.

The extension of the classical Pawlak rough approach under dominance relation is called dominance-based rough set approach. As for DRSA, the concepts to be approximated are not the subset of the universe, but rather the union of decision classes such as the followings (Greco, Matarazzo, and Slowinski, 2002).

**Definition 3.** The upward union of decision classes is defined as: $Cl^+_r = \bigcup_{a \in A} Cl_a$. The downward union of decision classes is defined as: $Cl^-_r = \bigcup_{a \in A} Cl_a$. The corresponding Boolean column vectors are denoted as: $Cl^+_a = \sum_{a \in A} Cl_a$.

**Example 1.** In Table 1, $C$ is conditional attribute set and $d$ is a decision attribute. $P \subseteq C = \{C, C\}, U/Ind(d) = \{Cl, Cl, Cl\}$, the decision class induced by attribute $d$ and the upward (downward) union can be figured out:

$Cl_1 = \{x, x, x, x\}$, $Cl_1^+ = U$, $Cl_1^- = Cl_1$; $Cl_2 = \{x, x, x, x, x, x\}$, $Cl_2^+ = \{x, x, x, x, x, x\}$; $Cl_3 = \{x, x, x, x, x, x, x, x\}$; $Cl_3^- = \{x, x, x, x, x, x, x, x\}$.

$Cl_1 = \{x, x, x, x\}$, $Cl_1^+ = Cl_1$, $Cl_1^- = U$.

So: $Cl_1^+ = \sum_{a \in A} Cl_a = [1,0,0,0,0,0,0,0]^T$, $Cl_2^+ = \sum_{a \in A} Cl_a = [1,0,0,0,0,0,0,0]^T$, $Cl_3^+ = Cl_3 = [0,0,0,0,0,0,0,0]^T$.

$Cl_1^- = \sum_{a \in A} Cl_a = [1,0,0,0,0,0,0,0]^T$, $Cl_2^- = \sum_{a \in A} Cl_a = [1,0,0,0,0,0,0,0]^T$, $Cl_3^- = \sum_{a \in A} Cl_a = [1,0,0,0,0,0,0,0]^T$. Given an information system $SI = (U, C \cup \{d\}, V, f)$, Let $P \subseteq C$, $Cl_1^+ = \bigcup_{a \in A} Cl_a$ and $Cl_1^- = \bigcup_{a \in A} Cl_a$. The lower and upper approximation of $Cl_1^+$ are defined respectively as followings (Greco, Matarazzo, and Slowinski, 2001; Greco, Matarazzo and Slowinski, 2002).
Table 1. An information system with preference-ordered domains

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>x₁</td>
<td>60</td>
<td>75</td>
<td>2</td>
</tr>
<tr>
<td>x₂</td>
<td>65</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>x₃</td>
<td>70</td>
<td>75</td>
<td>1</td>
</tr>
<tr>
<td>x₄</td>
<td>60</td>
<td>65</td>
<td>1</td>
</tr>
<tr>
<td>x₅</td>
<td>80</td>
<td>90</td>
<td>2</td>
</tr>
<tr>
<td>x₆</td>
<td>90</td>
<td>80</td>
<td>3</td>
</tr>
<tr>
<td>x₇</td>
<td>80</td>
<td>85</td>
<td>3</td>
</tr>
<tr>
<td>x₈</td>
<td>90</td>
<td>90</td>
<td>3</td>
</tr>
</tbody>
</table>

Definition 4.

\[ P(C₁^u) = \{ x ∈ U : D'_x(x) ⊆ C₁^u \} \]  \hspace{1cm} (5)
\[ \overline{P}(C₁^u) = \{ x ∈ U : D'_x(x) \cap C₁^u ≠ \emptyset \} \]  \hspace{1cm} (6)

Lower and the upper approximations of \( C₁^u \) are defined respectively as followings:

Definition 5.

\[ P(C₁^u) = \{ x ∈ U : D'_x(x) ⊆ C₁^u \} \]  \hspace{1cm} (7)
\[ \overline{P}(C₁^u) = \{ x ∈ U : D'_x(x) \cap C₁^u ≠ \emptyset \} \]  \hspace{1cm} (8)

Obviously, these two approximations of \( C₁^u \) can divided \( U \) into three disjoint regions: the positive region \( POS_{C₁^u}(C₁^u) \), the boundary region \( BND_{C₁^u}(C₁^u) \) (i.e., the difference of upper approximation set and lower approximation set) and the negative region \( NEG_{C₁^u}(C₁^u) \) (i.e., the complementary set of upper approximation).

\[
\begin{align*}
POS_{C₁^u}(C₁^u) &= P_{C₁^u}(C₁^u) \\
BND_{C₁^u}(C₁^u) &= P_{C₁^u}(C₁^u) - P_{C₁^u}(C₁^u) \\
NEG_{C₁^u}(C₁^u) &= U - P_{C₁^u}(C₁^u)
\end{align*}
\]  \hspace{1cm} (9)

It is similarity to two approximations of \( C₁^u \) in the partition of universe \( U \).

\[
\begin{align*}
POS_{C₁^u}(C₁^u) &= P_{C₁^u}(C₁^u) \\
BND_{C₁^u}(C₁^u) &= P_{C₁^u}(C₁^u) - P_{C₁^u}(C₁^u) \\
NEG_{C₁^u}(C₁^u) &= U - P_{C₁^u}(C₁^u)
\end{align*}
\]  \hspace{1cm} (10)

2.2. An Extension of Dominance-Based Rough Set Approach

The dominance-based rough set approach in definition 4 is also called original dominance-based rough set approach because it can be obtained through extending the classical Pawlak rough set approach by substitution of the indiscernibility relation in RST by a dominance relation. Inspire of the idea of the probability rough sets model (Yao, 2003; Yao, 2011), a measure index called rough inclusion is introduced and the dominance relation rough set approach is extended further in order to enhance the ability to tolerate the inconsistency data in real applications. This extended approach is called variable consistency dominance-based rough set approach (VC-DRSA). The definition of rough inclusion is as follows.

Definition 5. Given two non-empty subsets \( X \) and \( Y \) of universe \( U \), the inclusion degree of the set \( X \) in set \( Y \) is defined as:

\[ I(X,Y) = \frac{|X \cap Y|}{|X|} \]  \hspace{1cm} (11)
Where $|\cdot|$ denotes the cardinal number of a set. The inclusion degree is also called consistency level. Obviously, $0 \leq I(X,Y) \leq 1$.

**Definition 6.** Given $IS = (U, C \cup \{d\}, V, f)$, $P \subseteq C$, we say $x \in U$ belongs to $CI^+_l$ at consistency level $l$ ($0 < l \leq 1$) if and only if $x \in CI^+_l$ and $I(D^+_l(x), CI^+_l) \geq l$.

we say $x \in U$ belongs to $CI^+_l$ at consistency level $l$ ($0 < l \leq 1$) if and only if $x \in CI^+_l$ and $I(D^+_l(x), CI^+_l) \geq l$.

The definition of P-lower approximations of $CI^+_l$ and $CI^+_l$ at certain consistency level $l$ in VC-DRSA can be derived naturally from the above considerations respectively (Greco, Matarazzo and Slowinski, 2001).

**Definition 7.**
\[
P^+ (CI^+_l) = \{x \in CI^+_l | I(D^+_l(x), CI^+_l) \geq l\} \tag{12}
\]
\[
P^- (CI^+_l) = \{x \in CI^+_l | I(D^-_l(x), CI^+_l) \geq l\} \tag{13}
\]

Where $0.5 < l \leq 1$.

The P-upper approximations of $CI^+_l$ and $CI^+_l$ at certain consistency level $l$ can be defined as complementary sets of the P-lower approximations of $CI^+_l$ and $CI^+_l$ with respect to universe $U$.

\[
\overline{P}^+ (CI^+_l) = U - P^+ (CI^+_l) \tag{14}
\]
\[
\overline{P}^- (CI^+_l) = U - P^- (CI^+_l) \tag{15}
\]

The setting of parameter $l$ is aim to degrade the threshold value by which whether a certain object $x(x \in CI^+_l)$ is the element of the lower approximation of $CI^+_l$ can be judged. In the consequence, the criterion, by which judging whether a given object $x$ belongs to the lower approximation of $CI^+_l$, is that the inclusion degree of $D^+_l(x)$ in set $CI^+_l$ should equal to or more than $l$ instead of $D^+_l(x)$ should fully included in set $CI^+_l$. Compared with DRSA, the number of elements of P-lower approximation under the VC-DRSA takes on the rising trend, meanwhile the number of elements of P-upper approximation set shows a descending trend, so the boundary set becomes small. The introduction of parameter $l$ can not only make DRSA be applied to knowledge acquisition better, but also be tolerant with partial data which violate the dominance principle on attribute domains.

Specially, if $l$ equals to 1, then the VC-DRSA will be degenerated to the DRSA (Ziarko, 1993).

### 3. MATRIX-BASED APPROACH FOR COMPUTING APPROXIMATIONS IN VC-DRSA

#### 3.1. Computational Method

On the basis of Boolean column matrix representation for subset of universe, the dominance relation matrix is introduced combining with the definition of dominance relation, and the matrix-based computational methods for approximations under VC-DRSA are proposed.

**Definition 8.** (Liu, 2006) Let $U = \{u_1, u_2, \ldots, u_n\}$ be a non-empty set, called universe, $X$ be subset of $U$ (X can be regarded as a concept). $V(X)$ denotes Boolean column vector of $X$, $V(X)$ is defined as

\[
V(X) = (x_1, x_2, \ldots, x_n)^T \tag{16}
\]

Where $T$ denotes the transpose operation, $x_i = \begin{cases} 1, & u_i \in X \\ 0, & u_i \notin X \end{cases}$ $V(X)$ assigns 1 to the element that belongs to $X$ and 0 to the element that does not belong to $X$.

**Example 3.** Let $U = \{u_1, u_2, u_3, u_4\}$ and $X = \{u_1, u_2, u_4\}$, then $V(X) = [1, 0, 1, 1]^T$.

**Definition 8.** $IS = (U, C \cup \{d\}, V, f)$, $|U| = m$, $P \subseteq C$, $D_r$ is the dominance relation in universe $U$, then dominance relation matrix is defined as:

\[
D_r = (m_{i,j})_{m \times n}, \quad m_{i,j} = \begin{cases} 1, & (u_i, u_j) \in D_r \\ 0, & (u_i, u_j) \notin D_r \end{cases} \tag{17}
\]

Where $m$ denotes/stands for the cardinal number of universe $U$. Obviously, $D_r$ is an $m$-th order matrix.

The following properties of matrix of dominance relation can be derived from its definition:

**Property 1** $D_r(u_i, u_i) = D_r(i, :)$, $D_r(i, :)$ denotes the matrix which is composed of the $i$-th row’s elements of $D_r$. 

272
Property 2 \( D^r_j(u_i) = D^r_j(., j) \). \( D^r_j(i, .) \) denotes the matrix which is composed of the \( j \)-th column’s elements of \( D^r_p \).

\( \text{sum}(D^r, 2) \) treats the rows of \( D^r \) as vectors, returning a column vector of the sums of each row. So it is a column matrix composed by the sum of the \( i \)-th row of \( D^r \).

Property 3 \( \text{sum}_p(i) = [D^r_j(u_i)] \). \( \text{sum}_p(i) \) denotes the \( i \)-th row’s element of matrix \( \text{sum}(D^r, 2) \).

Property 4 \( \text{sum}_p(i) = [D^r_j(u_i)] \). \( \text{sum}_p(i) \) denotes the \( i \)-th row’s element of matrix \( \text{sum}(D^r^T, 2) \).

According to the definition 7, the key for computing approximations of upward(downward)union of decision classes is the judgements on intersection relation and inclusion relation between the union and \( P \)-dominating set(\( P \)-dominated set). Considering the inner product between the corresponding column vector(matrix) can judge intersection relation and inclusion relation between two sets, the matrix-based approach for computing the lower approximation of \( C^r_l \) and \( C^r_u \) at certain consistency level \( l \) in VC-DRSA is presented in the form of theorems as follows. In formula (18) and (19), \( \cdot \) \( \cdot \) denotes element-by-element multiplication between two column matrices which have the same dimensions; \( \cdot \) \( \cdot \) denotes matrix multiply and \( \geq \) is the comparison operator between the same dimension matrices, its comparison result is the choice between 0 or 1.

Theorem 1

\[
P'(C^r_l) = \left( \begin{array}{c}
D^r_j(u_i)\\
D^r_j(u_i) \\
\vdots \\
D^r_j(u_i)
\end{array} \right) \bullet C^r_l \geq \left[ \begin{array}{c}
l \cdot |D^r_j(u_i)| \\
l \cdot |D^r_j(u_i)| \\
\vdots \\
l \cdot |D^r_j(u_i)|
\end{array} \right] \bullet C^r_l = (D^r_j \cdot C^r_l \geq l \cdot \text{sum}(D^r, 2)) \cdot C^r_l (0.5 < l \leq 1)
\]

Proof. Let \( p_i \) be element of the \( i \)-th row of \( P'(C^r_l) \).

1) if \( u_i \in P'(C^r_l) \) and \( u_i \in C^r_l \), then \( p_i = 1 \).

According to the definition 7 and the property 4 of dominance relation matrix, we have followings:

\[
I(D^r_j(u_i), C^r_l) \geq l \Leftrightarrow D^r_j(u_i) \bullet C^r_l \geq l \cdot |D^r_j(u_i)| \Leftrightarrow D^r_j(u_i) \bullet C^r_l = l \cdot \text{sum}_p(i) .
\]

And because \( u_i \in C^r_l \), \( \cdot (D^r_j(u_i) \bullet C^r_l = l \cdot \text{sum}_p(i)) \cdot C^r_l = 1 \Leftrightarrow p_i = 1 \). Where \( C^r_l(i) \) is the \( i \)-th row’s element of \( C^r_l \).

2) if \( u_i \notin P'(C^r_l) \), then \( p_i = 0 \).

According to the definition 7 and the property 4 of dominance relation matrix, we have followings:

\[
I(D^r_j(x), C^r_l) < l \Leftrightarrow D^r_j(u_i) \bullet C^r_l < l \cdot |D^r_j(u_i)| \Leftrightarrow D^r_j(u_i) \bullet C^r_l < l \cdot \text{sum}_p(i) .
\]

So theorem 1 holds. □

Theorem 2

\[
P(C^r_l) = \left( \begin{array}{c}
D^r_j(u_i) \\
D^r_j(u_i) \\
\vdots \\
D^r_j(u_i)
\end{array} \right) \cdot C^r_l \geq \left[ \begin{array}{c}
l \cdot |D^r_j(u_i)| \\
l \cdot |D^r_j(u_i)| \\
\vdots \\
l \cdot |D^r_j(u_i)|
\end{array} \right] \cdot C^r_l = (D^r_j \cdot C^r_l \geq l \cdot \text{sum}(D^r, 2)) \cdot C^r_l (0.5 < l \leq 1)
\]

Proof. The proof is similar to that of theorem 1.□

The upper approximations of \( C^r_l \) and \( C^r_u \) at certain consistency level \( l \) in VC-DRSA can be gained easily through formula (14) and formula (15).

The positive region, boundary region and negative region of \( C^r_l \) and \( C^r_u \) can be easily obtained from upper and lower approximations of \( C^r_l \) and \( C^r_u \) at certain consistency level \( l \) in VC-DRSA.

Example 2. Continuation of Example 1.

The dominance-based matrix can be derived from table 2:
8.0

\[
D_c = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix},
\]

\[
D'_c = \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

So we have: \( \text{sum}(D_c,2) = [2,1,4,1,6,5,5,7]^T \), \( \text{sum}(D'_c,2) = [6,6,5,7,2,2,3,1]^T \).

Let \( l=0.8 \). Take calculation of \( P(C'_1) \), \( \overrightarrow{P}(C'_1) \), and \( \overrightarrow{P}(C'_1) \) as examples:

\[ P^8(C'_1) = (D_c \bullet C'_1 \geq 0.8 \cdot \text{sum}(D_c,2)). \]

\[ \text{sum}(D_c,2) = (2,1,4,1,6,5,5,7)^T \geq [1.6,0.8,3.2,0.8,4.8,4.4,5.6]^T. \]

\[ \text{sum}(D'_c,2) = [6,6,5,7,2,2,3,1]^T \geq [1.6,0.8,3.2,0.8,4.8,4.4,5.6]^T. \]

4. DESCRIPTIONS AND ANALYSIS OF MATRIX-BASED ALGORITHMS

The matrix-based algorithm for computing approximations of upward union of decision class level \( l \) can be gained easily according to above theorems and definitions.

Algorithm 1: Matrix-based algorithm for computing approximations of upward union of decision class level \( l \).

Input: \( I = (U, C \cup \{d, V, f\}) \), \( l \in \{C \cup \{d, V, f\} \} \) and consistency level \( l \).

Output: \( P(C'_1) \), \( \overrightarrow{P}(C'_1) \).

Step 1: Construct the dominance relation matrix \( D_c \) and derive the column vectors \( C'_1 \), \( C'_1 \) from upward (downward) union of decision classes \( C'_1 \) and \( C'_1 \).

Step 2: Compute \( D'_c \), \( C'_1 \), and \( D'_c \), \( C'_1 \), denoted as \( \text{Product 1} \) and \( \text{Product 2} \) respectively.

Step 3: Compute column vector \( \text{sum}(D'_c,2) \) of \( D'_c \).

Step 4: Compute \( P = (\text{Product 1} \geq l \cdot \text{sum}(D'_c,2)) \) \( \text{according to theorem 1} \).

Step 5: Compute \( P^8(C'_1) = (D_c \bullet C'_1 \geq 0.8 \cdot \text{sum}(D_c,2)). \)

Step 6: Compute \( P(C'_1) = P \cdot (C'_1)^T \) and \( \overrightarrow{P}(C'_1) = P \cdot (C'_1)^T \).

Step 7: Derive \( P(C'_1) \) and \( \overrightarrow{P}(C'_1) \) from \( P(C'_1) \) and \( \overrightarrow{P}(C'_1) \) according to definition 8.

Step 8: Obtain \( \overrightarrow{P}(C'_1) \) by computing the complementation set of \( P^8(C'_1) \) universe \( U \).

The matrix-based algorithm for computing the approximations of downward union of decision class level \( l \) can be gained easily according to above theorems and definitions.

Algorithm 2: Matrix-based algorithm for computing approximations of downward union of decision class level \( l \).
Input: $I = (U, C \cup d, V, f)$, $U \setminus d$, $C^l_1, C^l_2$, and consistency level $l$.

Output: $\overline{P}(C^l_1), \overline{P}(C^l_2)$.

Step 1: Construct the dominance relation matrix $D_x$ and derive the column vectors $C^l_1, C^l_2$ from downward(upward)union of $C^l_1$ and $C^l_2$;

Step 2: Compute $D_x \bullet C^l_1$ and $D_x \bullet C^l_2$, denoted as Product 3 and Product 4 respectively;

Step 3: Compute column vector $\text{sum}(D_x, 2)$ of $D_x$;

Step 4: Compute $P = (\text{Product} 3 \geq l \cdot \text{sum}(D_x, 2))$; according to theorem 2

Step 5: Compute $P' = (\text{Product} 4 \geq l \cdot \text{sum}(D_x, 2))$; according to theorem 1

Step 6: Compute $P(C^l_1) = P \bullet \overline{C^l_1}$ and $P(C^l_2) = P \bullet \overline{C^l_2}$;

Step 7: Derive $P(C^l_1)$ and $P(C^l_2)$ from $P(C^l_1)$ and $P(C^l_2)$ according to definition 8;

Step 8: Obtain $\overline{P}(C^l_1)$ by computing the complementation set of $\overline{P}(C^l_1)$ universe $U$.

In order to illustrate the validity of the proposed matrix approach (Algorithm 1, Algorithm 2), a traditional non-matrix algorithm computing approximations (Algorithm 3) is adopted and compared with Algorithm 1 and Algorithm 2 with respect to computational complexity. Algorithm 3 is originated from the idea of computing the equivalent class and support subset proposed by Zhang Wen-xiu in literature (Zhang, Wu, and Liang, 2006, 57-62) and was obtained by generalizing the idea to the calculations of $P$-dominated sets ($P$-dominating sets) and approximations of upward(downward)union of decision classes at certain level $l$ in VC-DRSA.

Algorithm 1, Algorithm 2 and Algorithm 3 have the same computational complexity with respect to compute the $P$-dominated sets or $P$-dominating sets (where $m$ denotes the number of attribute). However, in the process of computing approximations, Algorithm 3 need to compare each element of each dominated set or dominating set (there are $n$ different $P$-dominated sets and $n$ different $P$-dominating sets, the average number of elements in each dominated sets or dominating sets is equal to $n/2$) with each element $x$ of upward(downward)union of decision classes at certain consistency level $l$ (the average number of elements in upward or downward union of decision classes is equal to $n/2$) in order to judge whether the element $x$ of upward(downward)union in the $P$-dominated sets($P$-dominating sets) of the element $x$. So the computational complexity of Algorithm 3 is equal to $O(n^2 \log n)$ under the prerequisite of sorted elements of $U$. Algorithm 1 and Algorithm 2 take the inherent advantage of MATLAB in matrix operation, the upper approximation and the lower approximation of upward(downward)union of decision classes at certain level $l$ can be derived directly from the product operation and comparison operation among the Boolean column matrix of the upward(downward)union, the dominance relation matrix and the column matrix composed by the sum of each row, it needs $n^2$ Boolean multiplication, $n^2(n-1)$ addition operations and $n$ comparison operations, so the complexity is approximately equal to $O(n^3)$. It shows Algorithm 1 and Algorithm 2 outperforms Algorithm 3 slightly.

5. ANALYSIS ON EXPERIMENTAL RESULTS

A large dataset with preference-domain attribute—car in UCI datasets (Frank and Asuncion, 2010) is selected and the performance of Algorithm 1 and Algorithm 2 are tested on car so as to further demonstrate the validity of Algorithm 1 and Algorithm 2. Algorithm 1, Algorithm 2 and Algorithm 3 are all applied to the calculation of approximations of upward(downward)union of decision classes in car dataset respectively and compared the time consumption of two algorithms.

Experimental Platform: CPU Intel Core i7-4510U(2.00GHz), 8.0G Memory, Windows 8 operation system, Matlab7.0 development tool.

Experimental method: Above all, the ‘car’ dataset is divided into five subsets; each sub-dataset is called car1, car2, car3, car4 and car5 respectively. Then test the performance(time consuming) of two kinds of algorithm on each of these five sub-datasets. The corresponding programs of two kinds of algorithm are developed on Matlab7.0 platform for computing the two approximations.

The main principles for selecting the sub-dataset are as follows. The size of universe car1, car2, car3 and car4 is one-fifth, two-fifths, three-fifths, four-fifths of that of dataset car respectively. And dataset car5 is the copy of soybean. There exists the following inclusion relation among these sub-datasets. The car5 contains car4, car4 contains car3, car3 contains car2 and car2 contains car1. Besides, the number of objects with a certain decision class in each sub-dataset is proportional to the size of their universe. The size of universe and the number of elements of four decision classes in each sub-dataset is showed in the Table 2.
Table 2. Basic information about each sub-dataset of car dataset

<table>
<thead>
<tr>
<th>Name of Sub-dataset</th>
<th>Size of Universe</th>
<th>Number of each Decision Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>unacc</td>
</tr>
<tr>
<td>car 1</td>
<td>346</td>
<td>243</td>
</tr>
<tr>
<td>car 2</td>
<td>692</td>
<td>485</td>
</tr>
<tr>
<td>car 3</td>
<td>1038</td>
<td>727</td>
</tr>
<tr>
<td>car 4</td>
<td>1384</td>
<td>970</td>
</tr>
<tr>
<td>car 5</td>
<td>1728</td>
<td>1210</td>
</tr>
</tbody>
</table>

The experimental results are shown on figure 1 and figure 2. Y-coordinate represents the time consuming on computing for approximations of $Cl^L_i$ (figure 1) and $Cl^G_i$ (figure 2), measured in second. 1, 2, 3, 4 and 5 denotes sub-dataset car1, car2, car3, car4 and car5 respectively in X-coordinate.

**Figure 1.** The comparison on the performance of algorithm 1 and algorithm 3

**Figure 2.** The comparison on the performance of algorithm 2 and algorithm

The following conclusions can be drawn from the experimental results shown as figure 1 and figure 2:

1) The computational time of two algorithms is increase with the increase in size of universe of sub-dataset.
2) The computational time of algorithm 1 is less than that of algorithm 2 for the same sub-dataset.
3) The difference of time consuming between algorithm 1 and algorithm 2 becomes large with the increase in size of universe of sub-dataset.

It can be seen from the above three conclusions that the matrix approach for calculation of approximations is feasible and outperforms the non-matrix approach.

6. CONCLUSIONS

Aiming at an extension of DRSA—VC-DRSA which has good ability to tolerate the data with non-consistency to the principle of preference attribute domain and considering that the computation of approximations is a crucial step in the process of extracting rules, a new matrix-based method for computing the lower approximation of upward(downward) union of decision classes has been proposed and tested on UCI datasets. The test results have demonstrated the proposed method is feasible, concise and valid in comparison with the precious non-matrix method. The matrix-based method is not the best approach for computing the approximations, after all the researches provide a novel and valid method from the viewpoint of matrix. To apply the matrix method to the dynamic knowledge maintenance under Dominance-based Rough Set Approach will be the main research contents in the future.

Acknowledgements

This work was supported by the Natural Science Foundation of China under grant No. 61562061 and No. 61363047, the Natural Science Foundation of Jiangxi Province of China under grant No. 20151BAB207032 and
REFERENCES


