HK.NCCP
International Journal of Intelligent Information and Management Science
ISSN: 2307-0692 Volume 4, Issue 5, October 2015

A Rough Set Model Based on Extended Dominance Relations

Yan LI, Shuai SUN, Yongfei JIN, Yan ZHAN
Key Lab. In Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding, CHINA

Abstract: Dominance relation rough set approach (DRSA) is a useful mathematical tool to deal with preference-ordered data. The main idea is using dominance relations to replace equivalent relations in classical rough set theory. However, the definition of conventional dominance relation is very strict which may limit its application to information systems with relative large number of attributes. In this paper, we relax the conditions in the definition of dominance relation and introduce the concept of extended dominance relation. The proprieties of this new concept are also discussed and it is found that all the properties of classical dominance relation are still satisfied.

Keywords: Rough Set; Information System; Dominance Relation; DRSA, Extended Dominance Relation

1. Introduction

Rough set theory is proposed by Pawlark [1] in 1982 which is a powerful mathematical tool to deal with inconsistency and ambiguity information. Classical rough set theory is developed based on equivalent relation and the basic knowledge granules are the equivalent classes. Based on the knowledge granules, every concept in the universe can be described by its two approximations (i.e., lower and upper approximations) which are the union of some basic knowledge granules. Further, knowledge reduction methods can be developed using the two approximations and decision rules are finally generated. The quality and quantity of the generated rules depend on the objects in the lower and upper approximations. However, the classical rough set theory cannot deal with preference-ordered data such as “high”, “medium” and “low”. This kind of ordered information often occurs in many real-life applications especially multi-criteria decision-making problems. In order to address this issue, Greco et al. extended the classical rough set theory by introducing the concept of dominance relation and proposed dominance-based rough sets approach (DRSA) [2-6]. In these references, the conditional attributes in an information system with preference order are called criteria. The knowledge granules generated based on equivalent relations in classical rough sets are replaced by dominance classes. The concepts of lower/upper approximations, knowledge reductions and decision rules are consequently defined based on dominance relations. Currently, DRSA has been widely applied in multi-criteria decision problems[7,8].

According to the definition of dominance relation, an object x is said to dominate another object y only when x dominates y on all the conditional attributes (criteria). This requirement can be hardly satisfied especially when there are many conditional attributes. In most situations, an object x dominates another object y on part of the attributes but is dominated by y on some other attributes. In this case, the dominance classes based on conventional dominance relations would be very small which results in small approximation sets. This will finally leads to relative small set of decision rules, which may decrease the coverage of rule set and problem-solving ability. In other words, some objects cannot be covered by any generated rule. In this paper, we extend the concept of dominance relation to relax its strict requirement in the definition. As a result, larger approximations of target concepts would be obtained and more decision rules would be generated from these approximations.

The remainder of this paper is organized as follows: We present basic notions of DRSA in Section 2; and point out the problem of the definition of dominance relation with examples in Section 3. In section 4, we define the new extended dominance relation and provide its properties, and conclusions are given finally.

2. Basic Concepts and Problem Statement

2.1. Basic Concepts in DRSA

As a prior knowledge, this section describes the involved concepts based on dominance relations of rough set theory.

Definition 1(Information System) A quadruple S = (U, A, V, f) is an information system, where U is a nonempty finite set of objects, called the universe. A is a nonempty finite set of attributes, $A = C \cup D, C \cap D = \varnothing$, where C and D denote the sets of condition attributes and decision attributes, respectively. $V = \bigcup_{a \in A} V_a$, V_a is the domain of attribute a. f: $U \times A \to V$ is an information function, which gives values to every object on each attribute, namely, $\forall a \in A, x \in U, f(x,a) \in V_a$.

Definition 2 (Dominance Relation) Let S = (U, A, V, f) is an information system, for $B \subseteq A$, we denote

$$R_B^{\leq} = \{(x_i, x_j) \in U \times U : f_l(x_i) \leq f_l(x_j), \forall a_l \in B\}$$

$R_B^{\leq}$ is the dominance relations of information system. Based on definition 2,

$D_P^+(x_i) = \{x_j \in U : (x_i, x_j) \in R_P^{\leq}\} = \{x_j \in U : f_l(x_i) \leq f_l(x_j), \forall a_l \in P, P \subseteq A\}$ is the dominance class of x_i .

Definition 3 (Lower/upper approximations) Let $X \subseteq U$ is a target concept in the universe, and the lower and upper approximations are respectively defined as follows:

$$\underline{R}^{\leq}(X) = \{x \in U : D_P^+(x) \subseteq X\},$$

$$\overline{R}^{\leq}(X) = \{x \in U : D_P^+(x) \cap X \neq \varnothing\}.$$

2.2. Problem Statement

From the definition of dominance relation in Section 2.1, an object x is said to dominate another object y on the attribute set A only when x dominates y on all the attributes in A. This is a strict requirement especially when the attribute set is large. As a result, most dominance classes are relative small and so the approximation sets. Since the decision rules will generated from these approximations, the size of the approximations determines the quality and quantity of the rules.

In the following, we use an example to explain our statement.

Example 1. Given information system as in Table 1, U is the universe and there are 15 objects x1-x20 and 7 condition attributes a1-a7.

Table 1. An information system

U	a_1	a_2	a_3	a_4	a_5	a_6	a_7	d
x_1	6.7	2.8	6.7	2	5.8	2.7	4.1	2
x_2	5.1	2.5	3	1.1	5.7	4.4	1.5	2
x_3	5.4	3.4	1.5	0.4	6.9	3.1	5.4	1
x_4	5.1	3.4	1.5	0.2	6.1	2.6	5.6	3
x_5	5.5	3.5	1.6	0.5	7.1	3.2	5.5	3
x_6	6.3	2.5	3.2	1.3	5.8	4.6	1.7	2
x_7	6.7	2.9	6.8	1.9	6.2	2.8	4.2	1
x_8	6.3	3.4	3.1	1.2	6.9	4.5	1.8	2
x_9	6.7	3	6.8	2.1	5.9	2.8	4.3	1
x_{10}	6.6	3.1	6.9	2.2	5.9	2.6	4.4	2
x_{11}	6.6	3	6.9	2.1	6.1	2.7	4.4	3
x_{12}	5.5	3.6	1.7	0.6	7.1	3.3	5.6	3
x_{13}	6.3	3	3.2	1.4	6.1	2.8	1.8	1
x_{14}	6.5	3.1	5.9	2.1	6.2	2.4	3.8	2
x_{15}	6.3	2.8	6.5	2.2	5.9	2.9	4.5	3
x_{16}	5.2	2.6	3.1	1.3	5.8	4.5	1.6	2
x_{17}	6.7	2.4	6.8	2.1	6.2	2.8	4.3	1
x_{18}	6.5	2.6	6.3	1.8	6.1	2.7	4.3	2
x_{19}	6.6	2.5	6.2	1.9	6.1	2.5	3.8	1
x_{20}	6.8	3.1	6.9	2.2	6.3	2.9	4.4	3

According to the definitions in Section 2.1, all the dominance classes can be computed as:

$D_P^+(x_1) = \{x_1, x_9, x_{20}\}$, $D_P^+(x_2) = \{x_2, x_6, x_8, x_{16}\}$,
$D_P^+(x_3) = \{x_3, x_5, x_{12}\}$, $D_P^+(x_4) = \{x_4, x_{12}\}$,
$D_P^+(x_5) = \{x_5, x_{12}\}$, $D_P^+(x_6) = \{x_6\}$, $D_P^+(x_7) = \{x_7, x_{20}\}$,
$D_P^+(x_8) = \{x_8\}$, $D_P^+(x_9) = \{x_9, x_{20}\}$, $D_P^+(x_{10}) = \{x_{10}, x_{20}\}$,
$D_P^+(x_{11}) = \{x_{11}, x_{20}\}$, $D_P^+(x_{12}) = \{x_{12}\}$,
$D_P^+(x_{13}) = \{x_{13}, x_{20}\}$,
$D_P^+(x_{14}) = \{x_{14}, x_{20}\}$, $D_P^+(x_{15}) = \{x_{15}\}$, $D_P^+(x_{16}) = \{x_{16}\}$,
$D_P^+(x_{17}) = \{x_{17}, x_{20}\}$, $D_P^+(x_{18}) = \{x_{11}, x_{18}, x_{20}\}$,
$D_P^+(x_{19}) = \{x_7, x_{11}, x_{19}, x_{20}\}$, $D_P^+(x_{20}) = \{x_{20}\}$.

Let $X = \{x_3, x_7, x_9, x_{13}, x_{17}, x_{19}\}$, then the lower and upper approximations of X are:

$\underline{R}^{\leq}(X) = \{x \in U : D_P^+(x) \subseteq X\} = \Phi$;

$\overline{R}^{\leq}(X) = \{x_3, x_7, x_9, x_{13}, x_{17}, x_{19}\}$ =X.

Based on the two approximations, only 5 decision rules can be generated from $\overline{R}^{\leq}(X)$. If we consider the 20 objects as training examples, then there are only 10 objects are covered by these rules. Here, an object is said to be covered by a rule when this object meets the precondition of the rule, i.e., when this object dominates at least one object in the upper approximation $\overline{R}^{\leq}(X)$.

Therefore, we can compute the set of objects which are covered by the generated rules as follows:

$D_P^+(x_3) = \{x_3, x_5, x_{12}\} \cup D_P^+(x_7) = \{x_7, x_{20}\} \cup D_P^+(x_9) = \{x_9, x_{20}\} \cup D_P^+(x_{13}) = \{x_{13}, x_{20}\} \cup D_P^+(x_{17}) = \{x_{17}, x_{20}\} \cup D_P^+(x_{19}) = \{x_7, x_{11}, x_{19}, x_{20}\} = \{x_3, x_5, x_7, x_9, x_{11}, x_{12}, x_{13}, x_{17}, x_{19}, x_{20}\}$.

That is to say, 50% of the objects in the universe cannot be classified by the rules.

3. Extended Dominance Relation and Numerical Examples

Definition 4.(Extended Dominance Relation) Let $S = (U, A, V, f)$ be a target information system, $P \subseteq A$, $M = [\alpha \cdot m]$, where m is the number of the attributes in P, [*] denotes the round number of *, we define

$$R_{P_\alpha}^{\leq} = \left\{ \begin{array}{l} (x_i, x_j) \in U \times U : f(x_i, a_k) \leq f(x_j, a_k), \\ \exists a_k \in P \text{ and } \sum_{k=1}^{|P|} |a_k| \geq M \end{array} \right\}$$

$R_{P_\alpha}^{\leq}$ is called the extended dominance relation of the information system S.

Definition 5.(Extended Dominance/dominated classes)Let $\alpha > 0.5$ and $M = [\alpha \cdot m]$, where m is the number of attributes, $[\cdot]$ is the rounding operator, $P \subseteq C$, then the dominance/dominated classes can be respectively defined as:

$$D_{P_\alpha}^{+}(x_i)=\{x_j\in U:(x_i,x_j)\in R_{P_\alpha}^{\leq}\}$$

$$D_{P_\alpha}^{-}(x_i)=\{x_j\in U:(x_j,x_i)\in R_{P_\alpha}^{\leq}\}$$

Note: In definitions 4-5, we relax the conditions in definitions 2-3. Different from the original definition of dominance relation, on a attribute set P, we define an object x_i dominating another object x_j when x_i dominates x_j on the majority of the attributes in P which is determined by the parameter α. Based on these definitions, the lower and upper approximations under extended dominance relations can be also defined as follows.

Definition 6. (Lower/Upper approximations based on extended dominance relation)

$\underline{R}_\alpha^{\leq}(X)=\{x\in U: D_{P_\alpha}^{+}(x)\subseteq X\}$,

$\overline{R}_\alpha^{\leq}(X)=\{x\in U: D_{P_\alpha}^{+}(x)\cap X\neq\varnothing\}$.

Example 2: Given an information system as in Table 1. Let α=0.8, and then according to the definition of the extended dominance relation (definition 4), the dominance classes are computed as

$D_{P_\alpha}^{+}(x_1)=\{x_1,x_7,x_9,x_{10},x_{11},x_{15},x_{17},x_{20}\}$,

$D_{P_\alpha}^{+}(x_2)=\left\{\begin{matrix}x_1,x_2,x_6,x_7,x_8,x_9,x_{10},x_{11},x_{13},\\ x_{14},x_{15},x_{16},x_{17},x_{18},x_{19},x_{20}\end{matrix}\right\}$,

$D_{P_\alpha}^{+}(x_3)=\{x_3,x_5,x_8,x_{12}\}$,

$D_{P_\alpha}^{+}(x_4)=\{x_3,x_4,x_5,x_7,x_8,x_{11},x_{12},x_{13},x_{17},x_{18},x_{20}\}$,

$D_{P_\alpha}^{+}(x_5)=\{x_5,x_{12}\}$,

$D_{P_\alpha}^{+}(x_6)=\{x_1,x_6,x_7,x_9,x_{10},x_{11},x_{13},x_{14},x_{15},x_{17},x_{18},x_{19},x_{20}\}$,

$D_{P_\alpha}^{+}(x_7)=\{x_7,x_9,x_{17},x_{20}\}$, $D_{P_\alpha}^{+}(x_8)=\{x_8\}$,

$D_{P_\alpha}^{+}(x_9)=\{x_9,x_{10},x_{11},x_{17},x_{20}\}$, $D_{P_\alpha}^{+}(x_{10})=\{x_{10},x_{11},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{11})=\{x_{10},x_{11},x_{20}\}$, $D_{P_\alpha}^{+}(x_{12})=\{x_{12}\}$,

$D_{P_\alpha}^{+}(x_{13})=\{x_7,x_8,x_9,x_{10},x_{11},x_{13},x_{14},x_{15},x_{17},x_{18},x_{19},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{14})=\{x_7,x_9,x_{10},x_{11},x_{14},x_{17},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{15})=\{x_{10},x_{15},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{16})=\{x_1,x_6,x_7,x_8,x_9,x_{10},x_{11},x_{13},x_{14},x_{15},x_{16},x_{17},x_{18}\}$, $D_{P_\alpha}^{+}(x_{17})=\{x_7,x_9,x_{17},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{18})=\{x_1,x_7,x_9,x_{10},x_{11},x_{15},x_{17},x_{18},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{19})=\{x_1,x_7,x_9,x_{10},x_{11},x_{15},x_{17},x_{18},x_{19},x_{20}\}$,

$D_{P_\alpha}^{+}(x_{20})=\{x_{20}\}$.

Let $X=\{x_3,x_7,x_9,x_{13},x_{17},x_{19}\}$ which is the same as that in Example 1, and then the lower and upper approximations of X are:

$\underline{R}_\alpha^{\leq}(X)=\{x\in U: D_{P_\alpha}^{+}(x)\subseteq X\}=\Phi$;

$\overline{R}_\alpha^{\leq}(X)=\{x\in U: D_{P_\alpha}^{+}(x)\cap X\neq\varnothing\}=$
$\{x_1,x_2,x_3,x_4,x_6,x_7,x_9,x_{13},$
$x_{14},x_{16},x_{17},x_{18},x_{19}\}\supset\overline{R}^{\leq}(X)$.

It is obvious that, after extending the concept of dominance relation, the upper approximation set becomes larger than that in Example 1. Consequently, the number of rules extracted from the approximations also becomes larger which means more objects can be covered by these rules. Since $\bigcup_{x\in\overline{R}_\alpha^{\leq}(X)} D_{P_\alpha}^{+}(x)=U$, all the objects in U can be covered by the extracted rules. In Example 1, however, only 50% objects can be covered due to the strict definition of classical dominance relation.

4. Properties of the Extended Dominance Relation

All the properties of classical dominance relations and approximations still hold in the extended counterparts as follows:

(1) $\underline{R}_\alpha^{\leq}(U)=\underline{R}_\alpha^{\geq}(U)=U,\overline{R}_\alpha^{\leq}(\varnothing)=\overline{R}_\alpha^{\geq}(\varnothing)=\varnothing$;

(2) $\underline{R}_\alpha^{\leq}(X)\subseteq X\subseteq\overline{R}_\alpha^{\leq}(X),\underline{R}_\alpha^{\geq}(X)\subseteq X\subseteq\overline{R}_\alpha^{\geq}(X)$;

(3) $\underline{R}_\alpha^{\leq}(X\cap Y)=\underline{R}_\alpha^{\leq}(X)\cap\underline{R}_\alpha^{\leq}(Y),\underline{R}_\alpha^{\geq}(X\cap Y)=\underline{R}_\alpha^{\geq}(X)\cap\underline{R}_\alpha^{\geq}(Y)$;

(4) $\overline{R}_\alpha^{\leq}(X\cup Y)=\overline{R}_\alpha^{\leq}(X)\cup\overline{R}_\alpha^{\leq}(Y),\overline{R}_\alpha^{\geq}(X\cup Y)=\overline{R}_\alpha^{\geq}(X)\cup\overline{R}_\alpha^{\geq}(Y)$;

(5) $\underline{R}_\alpha^{\leq}(\sim X)=\sim\overline{R}_\alpha^{\leq}(X),\underline{R}_\alpha^{\geq}(\sim X)=\sim\overline{R}_\alpha^{\geq}(X)$;

(6) $\underline{R}_\alpha^{\leq}([x_i]_P^{\leq})=[x_i]_P^{\leq},\underline{R}_\alpha^{\geq}([x_i]_P^{\geq})=[x_i]_P^{\geq}$.

These results can be directly obtained according to Definitions 4-6, and we omitted the proofs.

5. Conclusions

In dynamic environment, information is constantly updated, and how to effectively deal with this kind of information system is an important topic. In this paper, we proposed an incremental approach for updating the approximations of VPRS model based on dominance relations under the variation of the object set. We gave detailed theoretical results with proofs and a numerical example to support our incremental method. One of our future work is to conduct some experiments with real datasets and consider the variations of attribute sets.

Acknowledgements

This work is supported by NSFC (No. 61473111), NSF of Hebei Province (No. F2014201100, A2014201003), and the Science Research Project from Hebei Education Department (No. Z2012147).

References

[1] Pawlak Z. Rough sets: theoretical aspects of reasoning about data [M]. Boston, Kluwer Academic Publishers, 1991.

[2] Greco S, Matarazzo B, Slowingski R. Rough approximation of a preference relation by dominance relation[J]. European Journal of Operation Research, 1991, 117 : 63-83.

[3] Greco S., Matarazzo B., Slowinski R. A new rough set approach to evaluation of bankruptcy risk. In: Zopounidis C.(Ed), Operational Tool in the Management of Financial Risk[M]. 1998, 121-136.

[4] Greco S., Matarazzo B., Slowinski R. Dominance-based rough set approach to rating analysis[J]. Fuzzy Economic Review. 1999.

[5] Greco S., Matarazzo B., Slowinski R. A new rough set approach to multicriteria and multiattribute classification. In: Polkowski, L., Skowron, A. (EdS), Rough sets and Current Trends in Computing[M]. 1998, 60-67.

[6] Greco S., Matarazzo B., Slowinski R. Rough sets theory for multicriteria decision analysis [J]. European Journal of Operational Research, 2001, 129:1-47.

[7] Qinghua Hu, Renda Yu. Application of rough calculation [M]. Science Press, 2012. 6

[8] Peng Song, Jiye Liang, and Yuhua Qian. A two-grade approach to ranking interval data[J], Knowledge-Based Systems, 2012, 27: 234-244.