HK.NCCP International Journal of Applied Mathematics and Soft Computing
Volume 3, Issue 1, June, 2017

A Rule Extraction Method for Hybrid Data

Yan LI[1], Tingting WU[2]
Key Lab. Of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding 071002, China

Abstract: In real world applications, the collected data often consists different types of attributes such as symbolic and ordinal attributes, which can be considered as hybrid data. In rough sets, the equivalence/dominance relations are used to represent and approximate data classification for symbolic data and ordinal data respectively. In order to extract decision rules in hybrid data, some scholars have proposed the mechanism to generate monotonic rules based on dominance relations. The original symbolic symbol values are also treated as preference-ordered attributes. Although this approach can generate rules with higher coverage, the efficiency is not satisfactory because all symbolic attributes need to be sorted two times as gain-type and cost-type. In this paper, we define dominance-equivalence relations on conditional attributes of hybrid data which introduces equivalence relations on symbolic attributes and dominance relations on ordinal attributes respectively. Thus, this method could preserve the original meaning of the data, and the rule generation could also be accelerated. Based on dominating and dominated classes, the upper and lower approximations and the decision rules can be computed and the matching mechanism of the rules is also established. Eleven UCI data sets are selected in the conducted experiments. The results show that the proposed method not only extracts more rules than the monotonic rules but also can reduce the running time obviously while the classification precision is also improved slightly.

Keywords: Rough set; Dominance relation; Equivalence relation; Monotonic Rules

1. Introduction

Rough set theory (RS) [1-7] proposed by Pawlak is an excellent mathematical tool to deal with imprecise, uncertain and blurred information and has been successfully applied in many fields such as data mining, neural network, pattern recognition, machine learning, decision analysis, etc [8-10]. The classical rough set theory is based on the concept of equivalence relation and different "granules of knowledge" formed by corresponding equivalence classes. Based on these different granules (i.e., equivalence classes), a target concept in the universe can be represented by two exact sets, i.e., the upper and lower approximations. The classical rough set theory are accurate because only two relations of "completely belong to" and "completely contain" need to be considered. Many scholars have improved and expanded it from different aspects. In 1993, Ziarko proposed a variable precision rough set model (VPRS)[11] to solve the problem that the lower approximation definition is too strict and the upper approximation definition is too loose. VPRS allows error classification on some extent that benefits to find potential data from seemingly unrelated data, which makes the application of RS more extensive [12-14]. On the other hand, traditional RS models can't deal with preference-ordered data such as attribute "bankruptcy risk" with two ordered values of "high" and "low". This is because that RS models are developed based on equivalence relations which only consider whether the attributes values can be distinguished or not, regardless of their preference relations. In order to solve this problem, Greco et al. firstly proposed rough set approach based on dominance relations (DRSA) [15-16]. The concept of original indiscernibility relation was replaced by dominance relation, and the data with preference relations could be represented [17-19]. In addition, there are many other extensions of RS models, including fuzzy rough set model [20,22], rough fuzzy set model [21,22], probability rough set model [23], etc.

Most of the above models are established for symbolic attributes or ordered attributes separately. However, some practical data is usually hybrid data. For example, the predictions of stock investors are often made based on the investors' gender, age and investment type. Obviously, the attribute values of age and type of investment are preference-ordered while gender is a symbolic attribute. In order to deal with such hybrid data, Greco et al. proposed a method to deal with ordinal and non-ordinal classification by using monotonic rules [19], placing both ordered and symbolic attributes in the framework of DRSA. In this method, the attributes with monotonic preference values are called criteria and other attributes are called regular attributes. These regular attributes are also considered as potential ordinal data, and then two decision tables are generated by sorting the attribute values in increasing and decreasing orders respectively. This method affects the actual meaning of the data itself and increases the time consumption in the computation process of upper and lower approximations and the extraction of decision rules. In this paper, we introduce dominance-equivalence relations to preserve the meaning of the original data and at the same time to improve the efficiency of generating rules. Considering the conditional attributes, the dominance relation is defined on criteria and the equivalence relation is still defined on the regular attributes. We can also define the dominating and dominated class and the upper (lower) approximations of decision classes. Decision rules can be correspondingly generated. This method significantly reduces the time consumption and preserves the classification ability of the original data.

2. Basic Concept

As a research foundation, this section briefly introduces several concepts which are closely related to this article.

Definition 1. (Target Information system)[27] S = (U, A, V, f) is called a target information system, where $U=\{x_1,x_2,...,x_n\}$ is a non-empty object set called universe; $A=C\cup D$, $C\cap D=\phi$ is a finite attribute set, where C is the condition attribute set and D decision attribute set; V is the set of attribute values, and $f:U\times A\to V$ assigns a value in V for every object on each attribute in U, namely, $\forall a\in A, x\in U, f(x,a)\in V_a$, V_a is the domain of attribute a. The target information system is also called decision table.

In real applications, an information system is often represented by a two-dimensional table, in which the rows are objects in the universe and columns are the attributes describing the objects. Each entry is an attribute value corresponding to each object. If the decision attributes are symbolic, then these attributes can form a partition of the universe (denoted by *U/D*) based on an equivalence relation. Each set in *U/D* is called a target concept or a decision class in *U*.

Definition 2. (Indiscernibility relation and equivalence class)[27] $S=(U,C\bigcup D)$ is a decision table, and for any non-empty subset $B\subseteq C$, we define the equivalence relation as indiscernibility relation:

$$IND(B)=\{(x,y)\in U\times U: a(x)=a(y),\forall a\in B\}.$$

The equivalence relation divides U into a set of equivalence classes, denote it as a partition of U:

$$U/IND(B)=\{[x]_B : x\in U\},$$

where $[x]_B=\{y\in U:(x,y)\in IND(B)\}$ is called the equivalence class of x based on $IND(B)$.

Rough sets are described by two exact sets, i.e. the upper and lower approximation sets.

Definition 3. (Dominance/dominated relation)[30]
S = (U, A, V, f) is a target information system, where $P\subseteq C$, $x,y\in U$, x dominates y with respect to $P\subseteq C$ (shortly, x P-dominates y), denoted by xD_Py, if for every criterion $q\in P$, $f(x,q)\geq f(y,q)$.

Definition 4. (Dominating/dominated set) Given a set of criteria $P\subseteq C$ and $x\in U$, the granules of knowledge used for approximation in DRSA are:

$D_P^+(x)=\{y\in U: yD_Px\}$ which is a set of objects dominating x called the P-dominating set of x; and

$D_P^-(x)=\{y\in U: xD_Py\}$ which is a set of objects dominated by x, called the P-dominated set of object x.

Definition 5. (Upward union and downward union of class)[24] Given the ordered classes cl1, cl2, …, clt, …cln. The upward union and downward union of these classes are defined as follows respectively:

$$cl_t^{\geq}=\bigcup_{s\geq t}cl_s, cl_t^{\leq}=\bigcup_{s\leq t}cl_s, t=1,2,......,n.$$

The statement $x\in cl_t^{\geq}$ means that "x belongs to at least class cl_t", while $x\in cl_t^{\leq}$ means "x belongs to at most class cl_t". Let us remark that $cl_1^{\geq}=cl_n^{\leq}=U$, $cl_n^{\geq}=cl_n$, $cl_1^{\geq}=cl_n^{\leq}=U$.

Definition 6. (Approximations of the upward union and downward union)[24] The P-lower approximation of $cl_t^{\geq}$, denoted by $\underline{P}(cl_t^{\geq})$, and the P-upper approximation of $cl_t^{\geq}$, denoted by $\overline{P}(cl_t^{\geq})$, are defined as follows ($t=1,2,......,n.$):

$$\underline{P}(cl_t^{\geq})=\{x\in U: D_P^+(x)\subseteq cl_t^{\geq}\},$$
$$\overline{P}(cl_t^{\geq})=\{x\in U: D_P^-(x)\cap cl_t^{\geq}\neq\phi\}.$$

Similarly, one can define the P-lower approximation and the P-upper approximation of $cl_t^{\leq}$ as follows ($t=1,2,......,n.$):

$$\underline{P}(cl_t^{\leq})=\{x\in U: D_P^-(x)\subseteq cl_t^{\leq}\},$$
$$\overline{P}(cl_t^{\leq})=\{x\in U: D_P^+(x)\bigcap cl_t^{\leq}\neq\phi\}.$$

Based on the above definitions, the monotonic rules of the decision classes can be extracted from the upper and lower approximation in the following form:

$$if\, x_{a1}\,\square\, p_1\, and\, x_{a2}\,\square\, p_2\, and ......and\, x_{an}\,\square\, p_n\; then\; x \in cl_t$$
$$or\; if\, x_{a1}\,\square\, p_1\, and\, x_{a2}\,\square\, p_2\, and ......and\, x_{an}\,\square\, p_n\; then\; x \in cl_t$$

where $a_i\in A, p_i\in P$, and this rule is denoted as rcl_t.

Definition 7. (Rule matching strategy)[25]
Classification of a new object x to cl_t (t = 1,2,......,n) using monotonic rules is based on a notion of class score coefficient associated with a set of rules covering the given object and three situations may occur in case of classification by a set of rules[25].

(1) None of the rules covers object x.

Object x is assigned to all considered decision classes.

(2) Exactly one decision rule covers object x.

The score that x assigned to cl_t is calculated as:

$$score_{rcl_t}(cl_t,x)=\frac{\left|\|\Phi_{rcl_t}\|\cap cl_t\right|^2}{\left|\|\Phi_{rcl_t}\|\right|\left|cl_t\right|}$$

where $\|\Phi_{rcl_t}\|$ denotes the set of objects satisfying the condition part of rule rcl_t, and $|*|$ denotes the cardinality of set *.

Similarly, based on rule $r\neg cl_t$, the score of assigning x to $\neg cl_t$ is calculated as:

$$score_{r\neg cl_t}(\neg cl_t,x)=\frac{\left|\|\Phi_{r\neg cl_t}\|\cap\neg cl_t\right|^2}{\left|\|\Phi_{r\neg cl_t}\|\right|\left|\neg cl_t\right|}$$

(3) Several rules cover object x.

The set of rules covering object x can be separated into two subsets: those suggest assigning x to cl_t and those to $\neg cl_t$. Then the positive and negative scores of x belonging to cl_t are defined as:

$$score^{+}_{rcl_t}(cl_t,x)=\frac{\left|(\|\Phi_1\|\cap cl_t)\cup\ldots\cup(\|\Phi_k\|\cap cl_t)\right|^2}{\left|\|\Phi_1\|\cup\ldots\cup\|\Phi_k\|\right|\left|cl_t\right|},$$

$$score^{-}_{r\neg cl_t}(cl_t,x)=\frac{\left|(\|\Phi_1\|\cap\neg cl_t)\cup\ldots\cup(\|\Phi_l\|\cap\neg cl_t)\right|^2}{\left|\|\Phi_1\|\cup\ldots\cup\|\Phi_l\|\right|\left|\neg cl_t\right|}.$$

The final score coefficient associated with class cl_t is computed as $score(cl_t,x)=score^{+}_{rcl_t}(cl_t,x)-score^{-}_{r\neg cl_t}(cl_t,x)$, and the class with the highest value of score coefficient is selected for the final assignment.

3. Rule Extraction based on Dominance-equivalence Relations

The method of extracting monotonic rules only considers the assignment of an object to a class or its complement. Using DRSA, the unions of decision class are approximated. For each t, consider two unions $cl_t^{\le}=cl_t, and$ $cl_t^{\ge}=U\setminus cl_t=\neg cl_t$. In this way, the original non-ordinal classification problem is reformulated to an ordinal classification problem with monotonicity constraints. The situation is similar to the binary ordinal classification with monotonicity constraints considered in DRSA. However, this would impact the actual meaning of the data to some extent and lead to a considerable time cost. To solve this problem, we give a method which combines the dominance relation and equivalence relation on condition attributes to handle both ordered and symbolic attributes.

Definition 8. (Dominance-equivalence relation) Given a decision table, the condition attribute set is $C=B\cup E$, where B is an ordinal attribute subset and E is a regular symbolic attribute subset. For $x,y\in U$, if x dominating y in B while x is equal to y in E, then we denote $xD_B y\cap x=_E y$; if x dominated by y in B and x is equal to y in E, then we denote $yD_B x\cap y=_E x$.

$xD_B y\cap x=_E y$ is called the dominance-equivalence relation defined in C.

On this basis, we can define the dominance /dominated set of an object based on dominance equivalence relations.

Definition 9. (Dominance /dominated set based on dominance-equivalence relation) In a decision table, $C=B\cup E$ is condition attribute set and D is decision attribute set. B and E are the sets of ordinal and non-ordinal attribute sets respectively. We define:

$$D_C^{+}(x)=\{y\in U, yD_B x\cap y=_E x\}$$

$$D_C^{-}(x)=\{y\in U, xD_B y\cap x=_E y\}$$

$D_C^{+}(x)$ and $D_C^{-}(x)$ are called the dominance set and dominated set of object x respectively.

Definition 10. (Upper/lower approximation based on dominance-equivalence relation) Assume cl_t is the t-th decision class in U/D. The upper and lower approximations of cl_t based on dominance-equivalence relation are defined as follows:

$$\underline{P}(cl_t)=\{x\in U: D_C^{+}(x)\subseteq cl_t\},$$

$$\overline{P}(cl_t)=\{x\in U: D_C^{+}(x)\cap cl_t\neq\phi\}.$$

$$\underline{P}'(cl_t)=\{x\in U: D_C^{-}(x)\subseteq cl_t\},$$

$$\overline{P}'(cl_t)=\{x\in U: D_C^{-}(x)\cap cl_t\neq\phi\}.$$

Based on these definitions, we can extract decision rules directly from the upper and lower approximations. Since the combination of dominance and equivalence relations is introduced, we can obtain decision rules in the form of "greater than or equal to" or "less than or equal to" on ordinal attributes and "equal to" on non-ordinal symbolic attributes. The consequences of the rules are decision classes rather than their upper/lower unions in monotonic rules. Therefore the decisions are more explicit and direct. Inspired by VC-DomLEM [26], we propose a rule extraction method based on the concept of rule coverage. The main idea is to avoid extracting rules from repeat objects. Once a rule has been generated from the upper/lower approximations of one decision class, the objects covered by the rule would be removed in the universe until U is empty. The forms of decision rules are as follows:

$$if\ x_{a1}\succ p_1\ and\ x_{a2}\succ p_2\ and......and\ x_{ak}\succ p_k\ and\ x_{ak+1}=p_{k+1}\ and......and\ x_{an}=p_n\ then\ x\in cl_t. \quad (1^*)$$

$$if\ x_{a1}\prec p'_1\ and\ x_{a2}\prec p'_2\ and......and\ x_{ak}\prec p'_k\ and\ x_{ak+1}=p_{k+1}\ and......and\ x_{an}=p_n\ then\ x\in cl_t. \quad (2^*)$$

where $k<n$, a_i is the i-th condition attribute, and p_i, p'_i are values of this attribute. Assume that the first k attributes are ordinal with dominance relations, and the rest (n-k) attributes are non-ordinal with equivalence relations.

After the raw rules are extracted, we can further simply and combine them by merging the attribute values to interval values on ordinal attributes when the rules have the same values on non-ordinal attributes and decision class. This could improve the efficiency of rule matching and classification. The detailed process of rule extraction and matching can be described as follows:

Attribute split. Given a target information system, split the condition attributes into ordinal and non-ordinal sets, and re-arrange the columns of the decision table so that the former p are ordinal and the rest ones are non-ordinal. Then the dominance/dominated sets, $D_C^+(x_i)$ and $D_C^-(x_i)$, are computed according to definition 9.

(1) Approximation set computation. Compute the decision classes to form partition U/D, and compute the upper and lower approximations of these decision classes: $\underline{P}(cl_t), \underline{P}(cl_t), \underline{P}'(cl_t),$ and $\underline{P}'(cl_t),$ by definition 10.

(2) Rule extraction. The original rules are directly generated from $\underline{P}(cl_t), \overline{P}(cl_t)$ in form (1*) and from $\underline{P}'(cl_t), \overline{P}'(cl_t)$ in form (2*). The process is completed using the covering method, where the objects covered by existing rules will be deleted until the universe is empty.

(3) Rule merge. For rules with the same values of same non-ordinal attributes and decision class, the final simplified rule set could be obtained by merging attributes values to interval values on ordinal attributes.

(4) Rule matching. The unseen objects are matched and classified by using the final rule set obtained from step (4). The rule matching strategy in definition 7 is applied and the corresponding scores of assigning one unseen object to each decision class are computed. The score maximization principle is used to finally assign the unseen object to one of the decision classes.

4. Experimental Results and Analysis

In this section, eleven UCI data sets are selected to verify the effectiveness of the proposed method based on dominance-equivalence relations. These data contain both ordinal and non-ordinal attributes and the decision classes are symbolic attributes. This partition of condition attributes is according to the description and explanation of original UCI data sets. In the experiment, the rows with missing data are deleted. We compare our proposed method with the monotonic rule-based method in [25] and analyze the performance in the following three aspects: (1) classification accuracy; (2) running time; (3) the number of extracted rules.

The two methods are implemented in the same experimental environment with identical rule extraction and matching strategy. Our experimental environment is Windows 7, Intel(R) Core(TM)i5-2400 CPU @ 3.10GHz 3.10GHz 8.00GB 64 bit operating system, and the used programming language is MATLAB (R2015b). Ten fold cross validation is used to obtain final results shown in Table 1. Here, “Objects” represents the number of rows and columns of the data sets; “DE” (abbreviation for Dominance-Equivalence) represents the proposed method based on dominance-equivalence relations; “MR” (abbreviation for Monotonic Rules) represents the monotonic rule based method. Accuracy, Time and No. of Rules represent the classification accuracy, running time and the number of extracted rules for the two corresponding algorithms, respectively. Certainly, higher accuracy and less running time are preferred.

Table 1. Comparisons of the Proposed Method and Monotonic Rule-based Method

Data Sets	Objects	Accuracy		Time		No. of Rules	
		DE	MR	DE	MR	DE	MR
AutoMpg	393*8	0.6260	0.6300	7.9034	12.670	46	10
balancescale	625*5	0.4624	0.4624	9.0867	21.769	25	3
haberman	306*4	0.7353	0.7319	3.1785	5.8160	62	5
postoperative	87*9	0.7272	0.7072	1.1142	1.6362	13	2
heart	270*10	0.5556	0.5556	6.0497	6.7329	28	4
hayesroth	132*6	0.3863	0.3714	1.6430	2.5830	16	5
Wholesale	440*8	0.6773	0.6758	7.5112	12.464	3	5
kohkiloyeh	100*6	0.6933	0.6803	0.8847	1.7371	4	3
tictactoe	958*10	0.6534	0.6534	35.538	54.483	9	1
Qualitative	250*7	0.5720	0.5721	3.2247	5.1714	12	1
tae	151*6	0.3246	0.3569	1.9814	2.9824	52	7
Average		0.5830	0.5815	7.1014	11.640	24	5

It can be seen from Table 1 that there is not noticeable difference in classification accuracy of the two methods, while the time consumption of the proposed method is obvious less than that of using monotonic rules. Therefore, the proposed method is more efficient without decreasing the classification quality. This is because the MR method needs to produce two copies of the original decision table to deal with both ordinal and non-ordinal attributes. Subsequently, the two decision tables were treated by the increasing preference and decreasing preference methods respectively to extract monotonic rules. Since the proposed method only needs to deal with one decision table, it saves much running time. On the other hand, DE method can extract more rules than MR method. This is because that the rules become more rigorous after the introduction of equivalence relation on non-ordinal attributes. The knowledge granules are much finer than the methods totally use dominance relations. This will preserve more information from original data sets.

Future work includes the improvement of rule matching mechanism such as introducing of coverage, confidence, and priority weights of the rules. This will help to further enhance the classification quality. Another consideration is to balance the number of rules (i.e., the generality ability of the knowledge) and the accuracy.

5. Conclusions

In this work, a novel rule extraction method is proposed to deal with hybrid data based on dominance-equivalence relations. This method can preserve the actual meaning of original attributes and extract more rules. At the same time, it can effectively reduce the running time compared with monotonic rule-based method. Experimental results on UCI data sets show the feasibility and effectiveness of this method.

6. Acknowledgements

This work is supported by the National Natural Science Foundations of China (No. 61473111), and the Natural Science Foundations of Hebei Province (No. F2014201100, A2014201003).

References

[1] Z. Pawlak, "Rough sets", International Journal of Information and Computer Sciences, 1982, vol.11, no. 3, pp. 341-356.

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

[3] D. Q. Miao, D. G. Li, Rough set theory algorithm and application. Beijing: Tsinghua university press, 2008.

[4] Q. Zhang, Q. Xie, and G. Wang, "A survey on rough set theory and its applications", CAAI Transactions on Intelligence Technology, 2016.

[5] J. T. Yao, P. Lingras, W. Z. Wu, et al., Rough Sets and Knowledge Technology. Springer Berlin Heidelberg, 2012.

[6] G. Y. Wang, Y. Y. Yao, and H. Yu, "A survey on rough set theory and applications", Chinese Journal of Computers, 2009, vol. 32, pp.1229-1246.

[7] Y. Y. Yao, "The two sides of the theory of rough sets", Knowledge-Based Systems, 2015, vol. 80, pp.67-77.

[8] Q. Cheng, Z. Qi, G. Zhang, et al., "Robust modeling and prediction of thermally induced positional error based on grey rough set theory and neural networks", International Journal of Advanced Manufacturing Technology, 2016, vol. 83, no. 5, pp.1-12.

[9] N. Pérez-Díaz, D. Ruano-Ordás, F. Fdez-Riverola, et al., "Boosting accuracy of classical machine learning antispam classifiers in real scenarios by applying rough set theory", Scientific Programming, 2016, no.3-4, pp.1-10.

[10] L. Peng, R. Niu, B. Huang, et al., "Landslide susceptibility mapping based on rough set theory and support vector machines: A case of the Three Gorges area, China", Geomorphology, 2014, vol. 204, no.1, pp. 287–301.

[11] W. Ziarko, "Variable precision rough set model", Journal of Computer and System Sciences, 1993, vol. 46, no. 1, pp. 39-59.

[12] Xie F, Lin Y, Ren W. Optimizing model for land use land cover retrieval from remote sensing imagery based on variable precision rough set. Ecological Modeling, 2011,222(2):232-240.

[13] G. Xie, W. Yue, S. Wang, et al., "Dynamic risk management in petroleum project investment based on a variable precision rough set model", Technological Forecasting & Social Change, 2010, vol. 77, pp. 891-901.

[14] H. Chen, J. Tao, and J. D. Zhang, "Intrusion detection research based on variable precision rough set", Journal of Qingdao University of Science and Technology, 2014, vol. 35, no. 2, pp.196- 199.

[15] S. Greco, B. Matarazzo, and R. Slowinski, "Rough approximation of a preference relation by dominance relations", European Journal of Operational Research, 1999, vol. 117, no. 1, pp. 63-83.

[16] S. Greco, B. Matarazzo, and R. Slowinski, "Rough approximation by dominance relations", International Journal of Intelligent Systems, 2002, vol. 17, no. 2, pp. 153–171.

[17] S. Y. Li and T. R. Li, "Incremental update of approximations in dominance -based rough sets approach under the variation of attribute values", Information Sciences, 2015, 294, pp. 348-361.

[18] S. Y. Li, T. R. Li, and D. Liu, "Incremental updating approximations in dominance-based rough sets approach under the variation of attribute set", Knowledge-Based Systems, 2013, 40, pp.17-26.

[19] W. Du and B. Hu, "Approximate distribution reducts in inconsistent interval-valued ordered decision tables", Information Sciences, 2014, 271, pp. 93-114.

[20] S. Nanda, "Fuzzy rough sets", Fuzzy Sets and Systems, 1992, 45, pp. 157-160.

[21] M. Banerjee and S. K. Pal, "Roughness of a fuzzy set", Information Sciences, 1996, 93, pp. 235-246.

[22] D. Dubois and H. Prade, "Rough fuzzy sets and fuzzy rough sets", International Journal of General System, 1990, 17, pp. 191–209.

[23] J. Y. Wang and L. M. Xu, "Probabilistic rough set model", Computer Science, 2002, 8, pp. 76-78.

[24] J. Blaszczynski, S. Greco, and R. Slowinski, "Ordinal and non-ordinal classification using monotonic rules", 8th International Conference of Modeling and Simulation-MOSIM, 2010.

[25] J. Blaszczynski, S. Greco, and R. Slowinski, "Multi-criteria classification-A new scheme for application of dominance-based decision rules", European Journal of Operational Research, 2007, vol. 181, no. 3, pp. 1030-1044.

[26] J. Blaszczynski, R. Slowinski, and M. Szelag, "Sequential covering rule induction algorithm for variable consistency rough set approaches", Information Sciences, 2011, vol. 181, no. 5, pp. 987-1002.

[27] K. Bache and M. Lichma, UCI Macine Learning Repository. (Online Available at http://archive. ics.uci.edu/ml,Irvine), CA: University of California, School of Information and Computer Science, 2013.

[28] J. Karami, A. Alimohammadi, and T. Seifouri, "Water quality analysis using a variable consistency dominance-based rough set approach", Computer Environment and Urban Systems, 2014, 43, pp. 25-33.

[29] W. X. Zhang and G. F. Qiu, Uncertain decision making based on rough sets. Beijing: Tsinghua university press, 2005.

[30] W. X. Zhang and W. H. Xu, "Distribution reduction of non coordinated target information system based on dominance relation", Fuzzy System and Mathematics, 2007, vol. 21, no. 4, pp.124-131.