VPRS-Based Knowledge Discovery Approach in Incomplete Information System

Shibao SUN[1,2]
1. Electronic Information Engineering College, Henan University of Science and Technology, Luoyang Henan 471003, China
2. National Lab. Of Software Development Environment, Beijing University of Aeronautics & Astronautics, Beijing, 100191, China
Email: sunshibao@126.com

Ruijuan ZHENG[1], Qingtao WU[1], and Tianrui LI[3]
1. Electronic Information Engineering College, Henan University of Science and Technology, Luoyang 471003, China
3. School of Information Science and Technology, Southwest Jiaotong University, Chengdu, 610031,China.
Email: rjwo@163.com, wqt8921@126.com, trli30@gmail.com

Abstract—**Through changing the equivalence relation in the incomplete information system, a new variable precision rough set model and an approach for knowledge reduction are proposed. To overcome no monotonic property of the lower approximation, a cumulative variable precision rough set model is explored, and the basic properties of cumulative lower and upper approximation operators are investigated. The example proves that the cumulative variable precision rough set model has wide range of applications and better result than variable precision rough set model.**

Index Terms—**Variable precision rough set, Incomplete information system, Decision table, Cumulative approximation**

I. Introduction

Rough set theory (RST) has been proposed by Pawlak [1] as a tool to conceptualize, organize and analyze various types of data in knowledge discovery. This method is especially useful for dealing with uncertain and vague knowledge in information systems. Many examples about applications of the rough set method to process control, economics, medical diagnosis, biochemistry, environmental science, biology, chemistry, psychology, conflict analysis and other fields can be found in [2,3]. However, the classical rough set theory is based on an equivalence relation and can not be applied in many real situations. Therefore, many extended RST models, e.g. binary relation based rough sets [4], covering based rough sets [5,6], and fuzzy rough sets [7,8] have been proposed. In order to solve classification problems with uncertain data and no functional relationship between attributes and relax the rigid boundary definition of the classical rough set model to improve the model suitability, the variable precision rough set (VPRS) model was proposed by Ziarko [9] in 1993. It is an effective mathematical tool with an error-tolerance capability to handle uncertainty problem. Basically, the VPRS is an extension of classical rough set theory [1-3], allowing for a partial classification. By setting a confidence threshold, $\beta(0 \le \beta < 0.5)$, the VPRS can allow noise data or remove error data [10]. Recently the VPRS model has been widely applied in many fields [11].

The key issues of VPRS model mainly concentrates on generalization of models and development of reduction approaches under the equivalence relation. For example, β-reduct [12], β lower (upper) distribution reduction [13] and reduction based on structure [14], etc, are reduction approaches under the equivalence relation. However, in many practical problems, the equivalence relation of objects is difficult to construct, or the equivalence relation of objects essentially does not exist. In this case, we need to generalize the VPRS model. The ideas of generalization are from two aspects. One is to generalize approximated objects from a crisp set to a fuzzy set [15]; The other is to generalize the relation on the universe from the equivalence relation to the fuzzy relation [15], binary relation [16], or covering relation [17,18]. The idea of the VPRS was introduced to fuzzy rough set and the theory and application of fuzzy rough set were discussed in [15]. The equivalence relation was generalized to a binary relation R on the universe U in the VPRS model, so that a generalized VPRS model was obtained [16]. Covering rough set model [19] has been obtained when the equivalence relation on the universe was generalized to cover on the universe in rough set model. The equivalence relation was generalized to cover

This work is supported by National Natural Science Foundation of China (Grant No. 60873108), the Special Research Funding to Doctoral Subject of Higher Education Institutes in P.R. China (Grant No. 20060613007), Supported by Natural Science Foundation of Henan Province of China(Grant No. 072300410210), Natural Science Research Foundation of Henan University of Science and Technology (Grant No.09001172).
Corresponding author:
Email: sunshibao@126.com (Shibao SUN).

doi:10.4304/jcp.5.1.110-116

on universe U in the VPRS model and two kinds of variable precision covering rough set models were obtained in [17,18]. The definition of the variable precision rough fuzzy set model under the equivalence relation was given in [20].

The classical rough set approach requires the data table to be complete, i.e., without missing values. In practice, however, the data table is often incomplete. To deal with these cases, Greco, et al [21] proposed an extension of the rough set methodology to the analysis of incomplete data tables. The extended indiscernible relation between two objects is considered as a directional statement where a subject is compared to a referent object. It requires that the referent object has no missing values. The extended rough set approach maintains all good characteristics of its original version. It also boils down to the original approach when there is no missing value. The rules induced from the rough approximations defined according to the extended relation verify a suitable property: they are robust in a sense that each rule is supported by at least one object with no missing value on the condition attributes represented in the rule. Obviously, these ideas can be used to the VPRS model.

The classical and the generalized VPRS approach based on indiscernible relations also require the data table to be complete. In this paper, two kinds of VPRS approaches for dealing with incomplete dada table are proposed. The paper is organized as follows. In Section 2, a general view of VPRS approach and incomplete information system are given. In Section 3, based on the extended indiscernible relation, we propose a new VPRS model and an approach for knowledge reduction in the incomplete information system. In Section 4, a cumulative VPRS model in the incomplete information system is discussed. They are based on the cumulative β lower (upper) approximation of X. In Section 5, we present an illustrative example which is intended to explain the concepts introduced in Section 3 and Section 4. The paper ends with conclusions and further research topics in Section 6.

II. PRELIMINARIES AND NOTATIONS

Definition 1 [1]. An information system is the 4-tuple $S=(U,Q,V,f)$, where U is a non-empty finite set of objects (universe), $Q=\{q_1,q_2,\cdots,q_m\}$ is a finite set of attributes, V_q is the domain of the attribute q, $V=\bigcup_{q\in Q}V_q$ and $f:U\times Q\to V$ is a total function such that $f(x,q)\in V_q$ for each $q\in Q$, $x\in U$, called an information function. If $Q=C\bigcup\{d\}$ and $C\bigcap\{d\}=\varnothing$, then $S=(U,C\bigcup\{d\},V,f)$ is called a decision table, where d is a decision attribute.

To every (non-empty) subset of attributes $P\subseteq C$ is associated an indiscernible relation on U, denoted by R_P:

$$R_P=\{(x,y)\in U\times U: f(x,q)=f(y,q)\ \forall q\in P\}. \qquad (1)$$

If $(x,y)\in R_P$, it is said that the objects x and y are P-indiscernible. Clearly, the indiscernible relation thus defined is an equivalence relation (reflexive, symmetric and transitive). The family of all the equivalence classes of the relation R_p is denoted by U/R_p and the equivalence class containing an element $x\in U$ by $[x]=\{y\in U:(x,y)\in R_p\}$. The equivalence classes of the relation R_p are called P-elementary sets. If $P=C$, the C-elementary sets are called atoms.

Definition 2 [9]. Let X and Y be subsets of non-empty finite universe U, if every $e\in X$ then $e\in Y$, we call Y contain X. It is described as $Y\supseteq X$. Let

$$c(X,Y)=\begin{cases}1-|X\bigcap Y|/|X|, & |X|>0,\\ 0, & |X|=0,\end{cases} \qquad (2)$$

where $|X|$ is cardinality of set X. $c(X,Y)$ is called the relative error ratio for X with regard to Y.

Definition 3 [9]. Let S be a decision table, X a nonempty subset of U, $0\le\beta<0.5$ and $\varnothing\ne P\subseteq C$. The β lower approximation and the β upper approximation of X in S are defined, respectively, by:

$$\underline{P}_\beta(X)=\{x\in U: c([x],X)\le\beta\}. \qquad (3)$$

$$\overline{P}_\beta(X)=\{x\in U: c([x],X)<1-\beta\}. \qquad (4)$$

The elements of $\underline{P}_\beta(X)$ are those objects $x\in U$ which belong to the equivalence classes generated by the indiscernible relation R_p, contained in X with the error ratio β; the elements of $\overline{P}^\beta(X)$ are all and only those objects $x\in U$ which belong to the equivalence classes generated by the indiscernible relation R_p, contained in X with the error ratio $1-\beta$.

Definition 4 [21]. An information system is called an incomplete information system if there exists $x\in U$ and $a\in C$ that satisfy that the value $f(x,a)$ is unknown, denoted as "*". It assumes here that at least one of the states of x in terms of P is certain where $P\subseteq C$, i.e. $\exists a\in P$ such that $f(x,a)$ is known. Thus, $V=V_C\bigcup V_d\bigcup\{*\}$.

Definition 5 [21]. $\forall x,y\in U$, object y is called subject and object x, referent. Subject y is indiscernible with referent x, with respect to condition attributes from $P\subseteq C$ (denoted as yI_Px), if for every

$q \in P$ the following conditions satisfy: (1) $f(x,q) \neq *$, (2) $f(x,q) = f(y,q)\ or\ f(y,q) = *$, where "*" denotes a missing value.

The above definition means that the referent object considered for indiscernible with respect to P should have no missing values on attributes from set P. The binary relation I_P is not necessarily reflexive and also not necessarily symmetric. However, I_P is transitive.

For each $P \subseteq C$, let us define a set of objects having no missing values on attributes from P:

$$U_P = \{x \in U : f(x,q) \neq *\ for\ each\ q \in P\}. \quad (5)$$

III. Variable Precision Rough Set Model in The Incomplete Information System

Definition 6. Let S be an incomplete information system, X a nonempty subset of U, $0 \le \beta < 0.5$ and $\varnothing \neq P \subseteq C$. The β lower approximation and the β upper approximation of X in S are defined, respectively, by:

$$\underline{P}^I_\beta(X) = \{x \in U_P : c(I_P(x), X) \le \beta\}. \quad (6)$$

$$\overline{P}^I_\beta(X) = \{x \in U_P : c(I_P(x), X) < 1-\beta\}. \quad (7)$$

The elements of $\underline{P}^I_\beta(X)$ are those objects $x \in U$ which belong to the equivalence classes generated by the indiscernible relation I_P, contained in X with the error ratio β; the elements of $\overline{P}^I_\beta(X)$ are all and only those objects $x \in U$ which belong to the equivalence classes generated by the indiscernible relation I_P, contained in X with the error ratio $1-\beta$.

The β boundary of X in S, denoted by $BN^I_\beta(X)$, is:

$$BN^I_\beta(X) = \{x \in U_P : \beta < c(I_P(x), X) < 1-\beta\}. \quad (8)$$

The β negative domain of X in S, denoted by $NEG_\beta(X)$, is:

$$NEG^I_\beta(X) = \{x \in U_P : c(I_P(x), X) \ge 1-\beta\}. \quad (9)$$

Corollary 1. When $\beta = 0$, the VPRS model defined above is equivalent to rough set model in incomplete information system [21].

Proof. In formula (10), $c(I_P(x), X) \le \beta$ is equivalent to $c(I_P(x), X) \le 0$, so $1 \le |I_P(x) \cap X| / |I_P(x)|$, such that $I_P(x) \subseteq X$, that is to say, $\underline{P}^I_\beta(X)$ is equivalent to $\underline{P}(X)$.

Analogously, $\overline{P}^I_\beta(X)$ is equivalent to $\overline{P}(X)$.

Corollary 2. If an information system is complete, the VPRS model defined above is equivalent to the classical VPRS model.

Proof. $\forall x, y \in U$, $P \subseteq C$, yI_Px, then for every $q \in P$, we have (1) $f(x,q) \neq *$, (2) $f(x,q) = f(y,q)\ or\ f(y,q) = *$. In a complete information system, we have $f(x,q) = f(y,q)$, so $I_P(x)$ is equal to $[x]$, such that formula (6) is equivalent to formula (3) and formula (7) is equivalent to formula (4).

Theorem 1. $\forall X \subseteq U$, $\underline{P}^I_\beta(\sim X) = NEG^I_\beta(X)$, where $(\sim X = U\text{-}X)$.

Proof. From Definition 6,

$$\underline{P}^I_\beta(\sim X) = \{x \in U_P : 1 - \frac{|I_P(x) \cap (\sim X)|}{|I_P(x)|} \le \beta\}$$

$$= \{x \in U_P : \frac{|I_P(x) \cap (\sim X)|}{|I_P(x)|} \ge 1-\beta\}$$

$$= \{x \in U_P : \frac{|I_P(x) \cap X|}{|I_P(x)|} \le \beta\}$$

$$= \{x \in U_P : 1 - \frac{|I_P(x) \cap X|}{|I_P(x)|} \ge 1-\beta\} = NEG^I_\beta(X).$$

Theorem 2. Let S be an incomplete information system, X, Y are two nonempty subsets of U, $0 \le \beta < 0.5$ and $\varnothing \neq P \subseteq C$. The rough approximations defined as above satisfy the following properties:

① $\underline{P}^I_\beta(X) \subseteq \overline{P}^I_\beta(X)$;

② $\underline{P}^I_\beta(\varnothing) = \overline{P}^I_\beta(\varnothing) = \varnothing$; $\underline{P}^I_\beta(U) = \overline{P}^I_\beta(U) = U$;

③ $X \subseteq Y \Rightarrow \underline{P}^I_\beta(X) \subseteq \underline{P}^I_\beta(Y)$;

④ $X \subseteq Y \Rightarrow \overline{P}^I_\beta(X) \subseteq \overline{P}^I_\beta(Y)$

⑤ $\underline{P}^I_\beta(X \cup Y) \supseteq \underline{P}^I_\beta(X) \cup \underline{P}^I_\beta(Y)$;

⑥ $\underline{P}^I_\beta(X \cap Y) \subseteq \underline{P}^I_\beta(X) \cap \underline{P}^I_\beta(Y)$;

⑦ $\overline{P}^I_\beta(X \cup Y) \supseteq \overline{P}^I_\beta(X) \cup \overline{P}^I_\beta(Y)$;

⑧ $\overline{P}^I_\beta(X \cap Y) \subseteq \overline{P}^I_\beta(X) \cap \overline{P}^I_\beta(Y)$;

⑨ $\underline{P}^I_\beta(\sim X) = \sim \overline{P}^I_\beta(X)$;

⑩ $\overline{P}^I_\beta(\sim X) = \sim \underline{P}^I_\beta(X)$;

Proof. ① Because of $0 \le \beta < 0.5$, x satisfies $\underline{P}^I_\beta(X)$ such that x satisfies $\overline{P}^I_\beta(X)$.

② Because of $0 \le \beta < 0.5$ and $X = \varnothing$, we have $\underline{P}^I_\beta(\varnothing) = \varnothing$ and $\overline{P}^I_\beta(\varnothing) = \varnothing$. Therefore, $\underline{P}^I_\beta(U) = \overline{P}^I_\beta(U) = U$.

③ $\forall x \in \underline{P}^I_\beta(X)$, $c(I_P(x), X) \le \beta$, when $X \subseteq Y$, we have $c(I_P(x), Y) \le \beta$, that is to say, $x \in \underline{P}^I_\beta(Y)$.

④ Similar to ③, we have $\overline{P}^I_\beta(X) \subseteq \overline{P}^I_\beta(Y)$.

⑤ From ③, we have $\underline{P}^I_\beta(X \bigcup Y) \supseteq \underline{P}^I_\beta(X) \bigcup \underline{P}^I_\beta(Y)$.

⑥ From ③, we have $\underline{P}^I_\beta(X \bigcap Y) \subseteq \underline{P}^I_\beta(X) \bigcap \underline{P}^I_\beta(Y)$.

⑦ From ④, we have $\overline{P}^I_\beta(X \bigcup Y) \supseteq \overline{P}^I_\beta(X) \bigcup \overline{P}^I_\beta(Y)$.

⑧ From ④, we have $\overline{P}^I_\beta(X \bigcap Y) \subseteq \overline{P}^I_\beta(X) \bigcap \overline{P}^I_\beta(Y)$.

⑨ From Theorem 1, $\underline{P}^I_\beta(\sim X) = \{x \in U_P : 1 - \frac{|I_P(x) \bigcap X|}{|I_P(x)|} \ge 1 - \beta\}$

$= \sim \{x \in U_P : 1 - \frac{|I_P(x) \bigcap X|}{|I_P(x)|} < 1 - \beta\} = \sim \overline{P}^I_\beta(X)$.

⑩ Similar to ③, we have $\overline{P}^I_\beta(\sim X) = \sim \underline{P}^I_\beta(X)$.

The following ratio defines a β accuracy of the approximation of $X \subseteq U$, $X \ne \varnothing$, by means of the attributes from $P \subseteq C$:

$$\alpha_P(X) = {|\underline{P}^I_\beta(X)|} \Big/ {|\overline{P}^I_\beta(X)|}. \tag{10}$$

Obviously, $0 \le \alpha_P(X) < 1$.

Another ratio defines a β quality of the approximation of X by means of the attributes from $P \subseteq C$:

$$\lambda_P(X) = {|\underline{P}^I_\beta(X)|} \Big/ {|X|}. \tag{11}$$

The quality $\lambda_P(X)$ represents the relative frequency of the objects with error ratio β correctly classified by means of the attributes from P.

A primary use of rough set theory is to reduce the number of attributes in databases thereby improving the performance of applications in a number of aspects including speed, storage, and accuracy. For a data set with discrete attribute values, this can be done by reducing the number of redundant attributes and find a subset of the original attributes that are the most informative.

Definition 7. (β dependability) Suppose that $S = (U, C \bigcup \{d\}, V, f)$ is an incomplete information system, β dependability is defined as follows:

$$\gamma(C, d, \beta) = |pos(C, d, \beta)| / |U|. \tag{12}$$

where

$$pos(C, d, \beta) = \bigcup_{Y \in U/d} \underline{C}^I_\beta(Y). \tag{13}$$

Definition 8. (β approximation reduction) Suppose that $S = (U, C \bigcup \{d\}, V, f)$ is an incomplete information system, $X \subseteq U$, a conditional attribute subset $A \subseteq C$ is called a β approximation reduction if and only if it satisfies: ① $\gamma(A, d, \beta) = \gamma(C, d, \beta)$ and ② there does not exist a conditional attribute subset $B \subseteq A$, such that $\gamma(B, d, \beta) = \gamma(C, d, \beta)$.

Based on Definition 8, through removing superfluous attributes, we can obtain a reductive database.

IV. Cumulative Variable Precision Rough Set Model in The Incomplete Information System

Let us observe that a very useful property of the lower approximation within the classical rough set theory is that if an object $x \in U$ belongs to the lower approximation of X with respect to $P \subseteq C$, then x belongs also to the lower approximation of X with respect to $R \subseteq C$ when $P \subseteq R$ (this is a kind of monotonic property). However, formula (6) does not satisfy this property of the lower approximation, because it is possible that $f(x, q) \ne *$ for all $q \in P$ but $f(x, q) = *$ for some $q \in R - P$. This is quite problematic for some key concepts of the variable precision rough set theory, like β accuracy and β quality of approximation, and β dependability.

Therefore, another definition of the lower approximation should be considered. Then the concepts of β accuracy, β quality , and β dependability of approximation can be still valid in the case of missing values.

Definition 9. Given $X \subseteq U$ and $P \subseteq C$,

$$\underline{P}^*_\beta(X) = \bigcup_{R \subseteq P} \underline{P}^I_\beta(X). \tag{14}$$

Then $\underline{P}^*_\beta(X)$ is called as the cumulative β lower approximation of X because it includes all the objects belonging to all β lower approximations of X, where $R \subseteq P$.

It can be shown that another type of the indiscernible relation, denoted by I^*_P, permits a direct definition of the cumulative β lower approximation in a usual way. For each $x, y \in U$ and for each $P \subseteq C$, $yI^*_P x$ means that $f(x,q) = f(y,q)$ *or* $f(x,q) = *$ and/or $f(y,q) = *$ for every $q \in P$. Let $I^*_P(x) = \{y \in U : yI^*_P x\}$ for each $x \in U$ and for each $P \subseteq C$. I^*_P is reflexive and symmetric but not transitive. We can prove that Definition 9 is equivalent to the following definition:

$$\underline{P}^*_\beta(X) = \{x \in U^*_P : c(I^*_P(x), X) \le \beta. \quad (15)$$

where

$$U^*_P = \{x \in U : f(x,q) \ne * \text{ for at least one } q \in P\}.$$

Using the indiscernible relation I^*_P, we can define the cumulative β upper approximation of X, complementary to $\underline{P}^*_\beta(X)$

$$\overline{P}^*_\beta(X) = \{x \in U^*_P : c(I^*_P(x), X) < 1-\beta\}. \quad (16)$$

For each $X \subseteq U$, let $X^*_P = X \bigcap U^*_P$. Let us remark that $x \in U^*_P$ if and only if there exists $R \ne \varnothing$ such that $R \subseteq P$ and $x \in U_R$.

Rough approximations $\underline{P}^*_\beta(X)$ and $\overline{P}^*_\beta(X)$ satisfies the following properties:

① For each $X \subseteq U$ and for each $P \subseteq C$: $\underline{P}^*_\beta(X) \subseteq \overline{P}^*_\beta(X)$;

② For each $X \subseteq U$ and for each $P \subseteq C$: $\underline{P}^*_\beta(X) = U^*_P - \overline{P}^*_\beta(U - X)$;

③ For each $X \subseteq U$ and for each $P, R \subseteq C$, if $P \subseteq R$, then $\underline{P}^*_\beta(X) \subseteq \overline{R}^*_\beta(X)$. Furthermore, if $U^*_P = U^*_R$, then $\overline{P}^*_\beta(X) \supseteq \overline{R}^*_\beta(X)$.

Due to the property of monotonic, when augmenting attributes set P, we get a lower approximation of X that is at least of the same cardinality. Thus, we can define analogously for the case of missing values the following key concepts of the variable precision rough sets theory: the cumulative β accuracy of approximation of X (denoted as $\alpha^*_P(X)$), the cumulative β quality $\lambda_P(X)$ of approximation of X (denoted as $\lambda^*_P(X)$), and the cumulative β dependability (denoted as $\gamma^*(C,d,\beta)$). These concepts have the same definitions as those given in Sections 3 but they use rough approximation $\underline{P}^*_\beta(X)$ and $\overline{P}^*_\beta(X)$.

V. An Example

The illustrative example presented in this section is to explain the concepts introduced in Section 3 and Section 4. The director of the school wants to make a global evaluation to some students. This evaluation should be based on the level in Mathematics, Physics and Literature. However, not all the students have passed all three exams and, therefore, there are some missing values. The director made the examples of evaluation as shown in Table 1.

TABLE I
STUDENT EVALUATIONS WITH MISSING VALUES

Student	Mathematics	Physics	Literature	Global evaluation
1	medium	bad	bad	bad
2	good	medium	*	good
3	medium	*	medium	bad
4	*	medium	medium	good
5	*	good	bad	bad
6	good	medium	bad	good

For $\beta = 0.35$, The lower and upper approximations can be calculated from Table 1:

Let $C = \{Mathematics, Physics, Literature\}$ be condition attributes and $\{Global \text{ evaluation}\}$ be decision attribute. Let $bad = \{1,3,5\}$ and $good = \{2,4,6\}$.

$U_C = \{1,6\}$, $I_C(1) = \{1\}$, $I_C(6) = \{2,6\}$, $\underline{C}^I_\beta(bad) = \{1\}$, $\overline{C}^I_\beta(bad) = \{1\}$, $\underline{C}^I_\beta(good) = \{6\}$, $\overline{C}^I_\beta(good) = \{6\}$, $\gamma(C,d,\beta) = 1/3$.

Let $L = \{Literature\}$, such that $\gamma(L,d,\beta) = 1/3$. It is easy to validate that L is reduction of condition attribute set C.

For $\beta = 0.35$, the cumulative lower and upper approximations can be calculated from Table 1:

$U^*_C = \{1,2,3,4,5,6\}$, $I^*_C(1) = \{1\}$, $I^*_C(2) = \{2,4,6\}$, $I^*_C(3) = \{3,4\}$, $I^*_C(4) = \{2,3,4\}$, $I^*_C(5) = \{5\}$, $I^*_C(6) = \{2,6\}$, $\underline{C}^*_\beta(bad) = \{1,5\}$, $\overline{C}^*_\beta(bad) = \{1,3,4,5\}$, $\underline{C}^*_\beta(good) = \{2,4,6\}$, $\overline{C}^*_\beta(good) = \{2,3,4,6\}$, $\gamma^*(C,d,\beta) = 5/6$.

Let $L = \{Literature\}$ and $P = \{Physics\}$, such that $\gamma(L,d,\beta) = 5/6$ and $\gamma(P,d,\beta) = 5/6$. It is easy to validate that L and P are cumulative β approximation reductions of condition attribute set C.

From this example, we can see that the cumulative variable precision rough set model better reflects the rough set's essence.

VI. CONCLUSIONS

In incomplete information system, a new VPRS model was obtained through defining the relation of objects. To overcome no monotonic property of the proposed VPRS model, a cumulative data relation was defined and a cumulative VPRS model was established. However, these models were limited in applications on a small database with incomplete information. Moreover, this paper only presented the basic reduction approach for a decision table. In our future work, we will focus on the development of reduction algorithms and extracting minimal exact rules from the large decision table. How to deal with fuzzy data in incomplete information system will also be one of our future research work.

ACKNOWLEDGMENT

The authors thank the editor and anonymous referees for their constructive comments and suggestions, which have improved the quality of this paper.

REFERENCES

[1] Z. Pawlak, A. Skowron, Rudiments of rough sets. Information Sciences, vol. 177, no. 1, pp. 3–27, 2007.
[2] Z. Pawlak, A. Skowron, Rough sets: some extensions. Information Sciences, vol. 177, no. 1, pp. 28–40, 2007.
[3] Z. Pawlak, A. Skowron, Rough sets and boolean reasoning. Information Sciences, vol. 177, no. 1, pp. 41–73, 2007.
[4] W. Zhu, F.-Y. Wang, Binary relation based rough sets, in: IEEE FSKD 2006, LNAI, vol. 4223, pp. 276–285, 2006.
[5] W. Zhu, Topological approaches to covering rough sets, Information Sciences, vol. 177, no. 6, pp. 1499–1508, 2007.
[6] W. Zhu, F.-Y. Wang, On three types of covering rough sets, IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 8, 2007.
[7] K.Y. Qin, Z. Pei, On the topological properties of fuzzy rough sets, Fuzzy Sets and Systems, vol. 151, no. 3, pp. 601–613, 2005.
[8] W.Z. Wu, W.X. Zhang, Constructive and axiomatic approaches of fuzzy approximation operators, Information Sciences, vol. 159, no. 3–4, pp. 233–254, 2004.
[9] W. Ziarko, Variable precision rough set model, Journal of computer system science, vol. 46, no. 1, pp. 39–59, 1993.
[10] D. Slezak, W. Ziarko, The investigation of the Bayesian rough set model, International Journal of Approximation Reason, vol. 40, pp. 81–91, 2005.
[11] Z. Tao, B. D. Xu, D. W. Wang, R. Li, Rough Rules Mining Approach Based on Variable Precision Rough Set Theory, Information and Control，vol. 33, no. 1, pp. 18-22, 2004 (in Chinese).
[12] M. Beynon, Reducts within the variable precision rough sets model: A further investigation, European journal of operational research, vol. 134, pp. 592-605, 2001.
[13] W. X. Zhang, Y. Liang, W. Z. Wu, Information System and Knowledge Discovery, Science Press, Beijing, China, 2003 (in Chinese).
[14] M. Inuiguchi, Structure-Based Approaches to Attribute Reduction in Variable Precision Rough Set Models, Proceeding of 2005 IEEE ICGC, pp. 34-39, 2005.
[15] A. Mieszkowicz-Rolka, L. Rolka, Remarks on approximation quality in variable precision fuzzy rough sets model Rough Sets and Current Trends in Computing: 4th International Conference, Springer-Verlag, pp. 402-411, 2004.
[16] Z. T. Gong, B. Z. Sun, Y. B Shao, D. G. Chen, Variable precision rough set model based on general relations, Journal of Lanzhou University (Natural Sciences), vol. 41, no. 6, pp. 110-114, 2005 (in Chinese).
[17] Y. J. Zhang, Y. P. Wang, Covering rough set model based on variable precision, Journal of Liaoning Institute of Technology, vol. 26, no. 4, pp. 274-276, 2006 (in Chinese).
[18] S. B. Sun, R. X. Liu, K. Y. Qin, Comparison of Variable Precision Covering Rough Set Models, Computer engineering, vol. 34, no. 7, pp. 10-13, 2008 (in Chinese).
[19] W. Zhu，F. Y. Wang, Reduction and axiomization of covering generalized rough sets, Information Sciences, vol. 152, pp. 217-230, 2003.
[20] A. Mieszkowicz-Rolka, L. Rolka, Fuzziness in Information Systems, Electronic Notes in Theoretical Computer Science, vol. 82, no. 4, pp. 1-10, 2003.
[21] X. B. Yang, J. Y. Yang, C. Wu, D. J. Yu, Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, Information Sciences, vol. 178, pp. 1219–1234, 2008.

Shibao SUN was born in Henan, China on 07/10/1970. He received B.S. degree in computer application from Henan University, Kaifeng, China in 1994, the M.S. degree in computer application from Zhengzhou University, Zhengzhou, China in 2004, the D.E. degree in traffic information engineering and control from Southwest Jiaotong University, Chengdu, China in 2008.

In 2008, he was a Post-Doc of Computer Science and Technology at the National Lab. Of Software Development Environment, Beijing University of Aeronautics & Astronautics, Beijing, China. He is currently an Associate Professor in Electronic Information Engineering College, Henan University of Science and Technology, Luoyang, China. His research interests can be summarized as intelligent information processing, machine learning. Particularly, he is currently interested in rough sets theory and approaches, as well as their applications in information processing.

Ruijuan ZHENG was born in Henan Province, China, on 01/20/80. She received B.S. degree in Computer Science and Technology from Henan University, Kaifeng, China in 2003, the M.S. and D.E. in bio-inspired computer network security theory and technology from Harbin Engineering University, Harbin, China in 2006 and 2008, respectively.

She works in Electronic Information and Engineering College, Henan University of Science and Technology now. Her research interests involve Computer immunology, Bio-inspired dependability, etc.

She is to join the China Computer Federation.

Qingtao WU was born in Jiangxi, China on 03/13/1975. He received M.S. degrees in computer science from Henan University of Science and Technology, Luoyang, China in 2003, the D.E. degree in computer applications from East China University of Science and Technology, Shanghai, China in 2006. Currently he is working as Associate Professor and Head in Dept. of Computer Science, Henan University of Science and Technology, Luoyang, China. His area of interest includes network & information security, software formal methods.

Tianrui LI was born in Fujian, China on 06/14/1969. He received B.S. and M.S. in mathematics from Southwest Jiaotong University, Chengdu, China in 1992 and 1995, respectively. He

received D.E. degree in traffic information engineering and control from Southwest Jiaotong University, Chengdu, China in 2002.

In 2005-2006, he was a Post-Doc of Computer Science and Engineering at the Belgian Nuclear Research Centre, Belgium. In 2008, he was a visiting professor for three months in Hasselt University, Belgium. He is currently a Professor in School of Information Science and Technology of Southwest Jiaotong University, Chengdu, China. His research interests can be summarized as developing effective and efficient data analysis techniques for novel data intensive applications. Particularly, he is currently interested in various techniques of data mining, granular computing and mathematical modeling, as well as their applications in traffic.

Prof. Li has served as ISKE2007, ISKE2008, ISKE2009 program chairs, IEEE GrC 2009 program vice chair and RSKT2008, FLINS2010 organizing chair, etc. and has been a reviewer for several leading academic journals.