2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

An Advancing Investigation on Reduct and Consistency for Decision Tables in Variable Precision Rough Set Models

James N.K. Liu
Department of Computing
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Email: jame.liu@polyu.edu.hk

Yanxing Hu
Department of Computing
The Hong kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Email: csyhu@comp.polyu.edu.hk

Jane Jia You
Department of Computing
The Hong kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
Email: csyjia@comp.polyu.edu.hk

Yulin He
College of Mathematics and Computer Science
Hebei University
Baoding 071002, Hebei, China
Email: yulinhe@ieee.org

***Abstract*—Variable Precision Rough Set (VPRS) model is one of the most important extensions of the Classical Rough Set (RS) theory. It employs a majority inclusion relation mechanism in order to make the Classical RS model become more fault tolerant, and therefore the generalization of the model is improved. This paper can be viewed as an extension of previous investigations on attribution reduction problem in VPRS model. In our investigation, we illustrated with examples that the previously proposed reduct definitions may spoil the hidden classification ability of a knowledge system by ignoring certian essential attributes in some circumstances. Consequently, by proposing a new β-consistent notion, we analyze the relationship between the structures of Decision Table (DT) and different definitions of reduct in VPRS model. Then we give a new notion of β-complement reduct that can avoid the defects of reduct notions defined in previous literatures. We also supply the method to obtain the β- complement reduct using a decision table splitting algorithm, and finally demonstrate the feasibility of our approach with sample instances.**

I. Introduction

Rough Set (RS) theory has been developed dramatically since its introduction by Pawlak in 1982 [1], [2], [3]. It provides a formal methodology aiming at data analysis problems that involve uncertain, imprecise or incomplete information, and has had widespread success in many artificial intelligence research fields [4], [5], [6]. However, when the given information contains some errors, such as missing information and classification abnormalities or the given Decision Table (DT) is derived from a relatively smaller data set, the obtained results of the classical RS model cannot always perform well and shows a poor generalization ability [7], [8]. The Variable Precision Rough Set (VPRS) [8] model was consequently proposed by Ziarko in 1993. The VPRS is one of the most important extensions of the RS which has been proved to be capable of efficiently solving the mentioned disadvantage of the RS model[9], [10] . By introducing a threshold β, the standard inclusion relation in RS is extended to majority inclusion relation in VPRS and data patterns can be analyzed from the perspective of statistics.

For both classical RS model and VPRS model, one of the core problems is to find some particular subsets of the attributes collection. Attributes out of such subsets can be deemed as redundant and removed without causing deterioration of classification quality and inducing brief decision rules inherent in the given tables [11], [4], [12]. These subsets of attributes are called the reduct of a RS or VPRS model. By calculating the reduct, necessary attributes are identified and redundant information can be removed. For classical RS model, the definition of reduct is uncontroversial [13], [14]. Accordingly, previous investigations about the reduct problem in RS model have mainly focused on the algorithm of calculating reduct [13], [15]. Unlike the classical RS model, in VPRS model, the definition of the reduct has been revised many times in previous investigations. Ziarko defined the β- reduct first [8], but unfortunately, subsequent researches reported that decision rule conflict will occur when using Ziarko's reduct definition [16], [17], [18]. Then a new definition called β-distribution reduct in VPRS is proposed [11], [17]. Different from the β-reduct that only requires the consistency of classification quality degree between the original DT and the reduct, the β-distribution reduct is a more rigorous definition that can avoid the decision rule conflict by keeping β-positive regions of the original DT consistent.

Our paper presents a further investigation on the reduct problem in VPRS model. In our investigation, we focus on the limitation of the β-distribution reduct and try to develop the reduct definition by considering the consistent property of a certain DT in VPRS model. Firstly, we use examples to demonstrate that although the β-distribution reduct can keep the decision rule consistent, it still has the risk to lose hidden information of the original DT and weaken the system's hidden classification ability. To deal with this problem, we extend the traditional consistent notion to a β-consistent notion for a DT in VPRS model, and analyze the relationship between the β-consistent notion and different definitions of reduct. Based on this analysis, a β-complete reduct notion is proposed. The β-complete reduct is shown not only can it keep the decision rule consistent but also avoid the weakening of the system's hidden classification ability. We give a decision table splitting algorithm for obtaining the proposed β-complement reduct. All of the notions and investigations in this paper can help develop the VPRS model further.

II. PRELIMINARIES

A DT is characterized by a 4-tuple $S = \langle U, A = C \cup D, V, f \rangle$, where $U = \{x_1, x_{2,} ... x_n\}$ denotes a nonempty finite set called universe, A is a nonempty finite set of attributes that contains condition attribute set $C = \{c_1, c_2, \cdots c_k\}$ and decision attribute set $D = \{d_1, d_2, \cdots d_h\}$, where $C \cap D = \emptyset$. $V = \cup_{\alpha \in C \cap D}$ and V_α is the $domain$ of the attributes α. $f : U \times A \rightarrow V$ is a total function such that $f(x, \alpha) \in V_\alpha$ for every $x \in U$ and $\alpha \in A$, called an information function. e.g. $f(x_i, \alpha_j) = v$ means that in this DT, w.r.t. a certain attribute α_j, the element x_i has the value v.

Given an arbitrary non-empty subset $B \subseteq A$, an indiscernibility relation $IND(B)$ is defined as:

$$\{(x_i, x_j) \in U \times U \,|f(x_i, \alpha) = f(x_j, \alpha) \; ,\forall \alpha \in B\} \tag{1}$$

$IND(B)$ partial U into a family of disjoint subsets $U/IND(B)$ called a quotient set of U:

$$U/IND(B) = \{\, [x]_B |\, x \in U\} \tag{2}$$

where $[x]_B$ denotes equivalence class determined by x w.r.t. $IND(B)$, i.e., :$[x]_B = \{\, y \in U |\, (y, x) \in IND(B)\}$.

Then for a DT, we can define the equivalence classes determined by condition attribute set C and equivalence classes determined by the decision attribute set D respectively as follows:

$$\{C_i, i = 1, 2, \cdots, m\} = \{\, [x]_C |\, x \in U\} \tag{3}$$

$$\{D_j, j = 1, 2, \cdots, n\} = \{\, [x]_D |\, x \in U\} \tag{4}$$

Especially, D_j is called the decision class.

In classical RS model, For C, the lower and upper approximations of D_j can be respectively defined as:

$$\underline{C}(D_j) = \{x \,|[x]_C \subseteq D_j \,\} \tag{5}$$

$$\overline{C}(D_j) = \{x \,|[x]_C \cap D_j \, \neq \emptyset\} \tag{6}$$

and the *positive region* is defined as $POS_C(D_j) = \underline{C}(D_j)$

A VPRS model was proposed as an important extension of classical RS model, which gives a less rigorous definition of the inclusion relation in eq.(5) and eq.(6) to make the classical RS model more fault tolerant [8], [19], [13]. In VPRS model, for a given precision parameter value $\beta \in (0.5, 1]$, we denote:

$$\underline{C}^{\beta}(D_j) = \bigcup \{x \in C_i \,|\omega(C_i, D_j) \geq \beta \,\} \tag{7}$$

$$\overline{C}^{\beta}(D_j) = \bigcup \{x \in C_i \,|\omega(C_i, D_j) > 1 - \beta \,\} \tag{8}$$

where ω is the inclusion degree function defined as:

$$\omega(X, Y) = \begin{cases} \frac{|X \cap Y|}{|X|}, |X| > 0, \\ 0, |X| = 0. \end{cases} \tag{9}$$

and the β-positive region of D_j w.r.t. C is $POS_C^{\beta}(D_j) = \underline{C}^{\beta}(D_j)$.

Since $\beta > 0.5$, the less rigorous inclusion relation in VPRS model is called the majority inclusion relation that is seen as the heart of the VPRS model.

When $\beta = 1$, the VPRS model is equal to a classical RS model [8].

A notion called the *classification quality degree* (or called the degree of dependence of decision attribute set D w.r.t. condition attribute set C) is applied to measure the classification quality in classical RS model. If the decision attribute set D divides U into n decision classes, the classification quality degree w.r.t a certain attribute set C is:

$$\sigma_C(D) = \frac{\left| \sum\limits_{j=1}^{n} \underline{C}(D_j) \right|}{|U|} \tag{10}$$

Similarly, in VPRS model, when a precision parameter β is given, the classification quality degree is:

$$\sigma_C^{\beta}(D) = \frac{\left| \sum\limits_{j=1}^{n} \underline{C}^{\beta}(D_j) \right|}{|U|} \tag{11}$$

Positive regions and classification quality degree reflect the knowledge of a certain data set from the perspective of quality and quantity respectively [8], [18].

III. ATTRIBUTES REDUCTION IN VPRS MODEL

This section briefly introduces the previous definitions of *reduct* in VPRS model, and more importantly, we will show the limitation of the presently used β-distribution reduct in this section.

A. The β-reduct and its limitation in VPRS model

The quality of classification is often used to measure the classification ability of a DT. Nevertheless, the final rule inference is based on the concrete objects in positive region. In classical RS model, the monotonicity of classification quality and positive region are uniform during the procedures of attribute reduction [20]. Accordingly, the same classification quality degree implies the consistency of positive region in classical RS model[13]. However, in VPRS model, these monotonicity properties will not be satisfied any more, therefore the same classification quality may result in different β-positive regions[18]. i.e., the final decision rules extracted from the reduct maybe in conflict with those extracted from the original DT under the same classification quality degree [20].

Definition 1: For a DT $S = (U, A = C \cup D, V, f)$ in VPRS model with precision parameter β, $B \subseteq C$, if:

$$\sigma_B^{\beta}(D) = \sigma_C^{\beta}(D) \;\; and \;\; \forall \alpha \in B, \; \sigma_{B - \{\alpha\}}^{\beta} \neq \sigma_C^{\beta} \tag{12}$$

TABLE I. SAMPLE DT 1

U	α_1	α_2	α_3	α_4	d
o_1	0	1	1	1	1
o_2	1	1	1	1	1
o_3	1	1	0	0	1
o_4	0	1	0	1	0
o_5	1	1	1	1	0
o_6	1	0	1	0	0
o_7	1	1	1	1	0

such a subset B is called a β-reduct of C under the precision β, defined by Ziarko [8].

However, under some circumstances, the β-reduct, based on the equalization of classification quality degree, has to face the inconsistent problem. Some conflicts will be generated between the decision rules extracted from the original DT and the decision rules extracted from the obtained reduct. An example is shown in Table I .

In sample DT 1, $o_1, o_2, \ldots, o_7$ are the elements in the universe, $C = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is the condition attribute set of this DT, and $D = \{d\}$ is the decision attribute set. The values in the table are the attribute values of the corresponding elements. We have:
$D_1 = \{o_1, o_2, o_3\}$,
$D_2 = \{o_4, o_5, o_6, o_7\}$,
as decision classes determined by d; and:
$C_1 = \{o_1\}$,
$C_2 = \{o_2, o_5, o_7\}$,
$C_4 = \{o_4\}$,
$C_5 = \{o_6\}$,
are equivalent classes determined by C. Setting $\beta = 0.59$, we can see that $B = \{\alpha_2, \alpha_4\}$ is a β-reduct of sample DT 1 since it is easy to calculate $\sigma_B^{0.59}(D) = \sigma_C^{0.59}(D) = 1$. However, we can also find the positive region inconsistency as:
$POS^{0.59}{}_C(D_{d=1}) = \underline{C}^{0.59}(D_{d=1}) = \{o_1, o_3\}$,
$POS^{0.59}{}_C(D_{d=2}) = \underline{C}^{0.59}(D_{d=2}) = \{o_2, o_5, o_7, o_4, o_6\}$,
and:
$POS^{0.59}{}_B(D_{d=1}) = \underline{B}^{0.59}(D_{d=1}) = \{o_3\}$,
$POS^{0.59}{}_B(D_{d=2}) = \underline{B}^{0.59}(D_{d=2}) = \{o_1, o_2, o_4, o_5, o_7, o_6\}$
which show that the β-positive regions are changed in the reduct $B = \{\alpha_2, \alpha_4\}$. Specifically, from the obtained β-reduct, when $\beta = 0.59$, we can extract the decision rule: $(\alpha_2, 1) \wedge (\alpha_4, 1) \rightarrow (d, 0)$, which is supported by o_1, o_2, o_4, o_5 and o_7; but in the original DT, we have: $(\alpha_1, 0) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 1)$, supported by o_1. The classification information is changed after the reduction.

B. The β-distribution reduct and its limitation

Definition 2: Given a DT $S = (U, A = C \cup D, V, f)$ in VPRS model with precision parameter β, $B \subseteq C$, B is a β-distribution reduct of C if:

$$L_B^\beta = L_C^\beta \;\; and \;\; \forall \alpha \in B, \; L_C^\beta \neq L_{B-\{\alpha\}}^\beta \tag{13}$$

where L is used to denote the collection of β-positive regions, $L_B^\beta = (\underline{B}^\beta(D_1), \underline{B}^\beta(D_2), \cdots, \underline{B}^\beta(D_n))$ and $L_C^\beta = (\underline{C}^\beta(D_1), \underline{C}^\beta(D_2), \cdots, \underline{C}^\beta(D_n))$.

Obviously, the β-distribution reduct is a more rigorous definition since it requires the consistency of both the classification quality degree and β-positive regions. From the definition, because that according to eq.(11), if eq.(12) is true, eq.(13) must be true. a β-distribution reduct is also a β-reduct, but conversely, a β-reduct is not definitely a β-distribution reduct as shown in sample DT 1. However, in our investigation, we find that the β-distribution reduct also has limitations:

TABLE II. SAMPLE DT 2

U	α_1	α_2	α_3	α_4	α_5	d
x_1	1	1	1	1	1	1
x_2	1	1	0	1	1	1
x_3	0	0	1	0	0	1
x_4	1	1	2	1	1	1
x_5	1	1	0	1	0	2
x_6	1	1	0	1	1	2
x_7	0	0	1	2	1	2
x_8	1	1	0	1	1	2
x_9	1	1	2	1	1	2
x_{10}	1	0	2	1	1	3
x_{11}	1	0	2	1	1	4

In sample DT 2, set $\beta = 0.6$, we have:
$D_1 = \{x_1, x_2, x_3, x_4\}$,
$D_2 = \{x_5, x_6, x_7, x_8, x_9\}$,
$D_3 = \{x_{10}\}$,
$D_4 = \{x_{11}\}$,
as decision classes. Meanwhile, we have equivalent classes:
$C_1 = \{x_1\}$,
$C_2 = \{x_2, x_6, x_8\}$,
$C_3 = \{x_3\}$,
$C_4 = \{x_4, x_9\}$,
$C_5 = \{x_5\}$,
$C_6 = \{x_7\}$,
$C_7 = \{x_{10}, x_{11}\}$.

For $B = \{\alpha_3, \alpha_4\}$, we can obtain equivalent classes:
$B_1 = \{x_1\}$,
$B_2 = \{x_2, x_5, x_6, x_8\}$,
$B_3 = \{x_3\}$,
$B_4 = \{x_4, x_9, x_{10}, x_{11}\}$,
$B_5 = \{x_7\}$.

Consequently, by calculating $\underline{B}^{0.6}(D_j)$ and $\underline{C}^{0.6}(D_j)$, $j = 1, 2, 3, 4$, we have:
$POS^{0.6}{}_B(D_1) = \{x_1, x_3\} = POS^{0.6}{}_C(D_1)$,
$POS^{0.6}{}_B(D_2) = \{x_2, x_5, x_7, x_6, x_8\} = POS^{0.6}{}_C(D_2)$,
$POS^{0.6}{}_B(D_3) = \emptyset = POS^{0.6}{}_C(D_3)$,
$POS^{0.6}{}_B(D_4) = \emptyset = POS^{0.6}{}_C(D_4)$,
in the meanwhile, we also have:
$POS^{0.6}{}_C(D_1) = \underline{C}^{0.6}(D_1) = \{x_1, x_3\}$,
$POS^{0.6}{}_C(D_2) = \underline{C}^{0.6}(D_2) = \{x_2, x_5, x_7, x_6, x_8\}$,
$POS^{0.6}{}_C(D_3) = \underline{C}^{0.6}(D_3) = \emptyset$,
$POS^{0.6}{}_C(D_4) = \underline{C}^{0.6}(D_4) = \emptyset$.
So $L_B^\beta = L_C^\beta$. Accordingly, $B = \{\alpha_3, \alpha_4\}$ is a β-distribution reduct of C when $\beta = 0.6$.

Now the problem comes. Observing sample DT 2, when $\beta = 0.6$, we can extract two special decision rules:
$(\alpha_1, 1) \wedge (\alpha_2, 0) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, (1 or 2))$, by x_4 and x_9;
$(\alpha_1, 1) \wedge (\alpha_2, 0) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, (3 or 4))$, by x_{10} and x_{11};
After the reduction, we can extract:
$(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (1 or 2 or 3 or 4))$, by x_4, x_9, x_{10} and x_{11}.
If the decision attribute in a decision rule can get various values, we say the conflict happens and denote the conflict

decision part of such conflict decision rule as $(d, \emptyset)$.

In both of the original DT and the β-distribution reduct, when $\beta = 0.6$, x_4, x_9, x_{10} and x_{11} all support that $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, \emptyset)$. But in the original DT, we can observe that when $f(x_i, \alpha_2) = 0$, there is $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (1 or 2))$, and when $f(x_i, \alpha_2) = 1$, the decision rule is $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (3 or 4))$, so the α_2 has some hidden classification ability that the attributes in the β-distribution reduct don't have and should not be seen as redundant.

The problem is generated in the process of calculating the β-distribution reduct: both of the two conditions, $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (1 or 2))$ and $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (3 or 4))$, are considered as $(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, \emptyset)$ without distinguishing the decision conflicts. But actually, when $f(x_i, \alpha_2) = 1$, x_i with $(\alpha_3, 2) \wedge (\alpha_4, 1)$ has 50% probability to be classified in D_1 and 50% probability to be classified in D_2, and when $f(x_i, \alpha_2) = 0$, sample point with $(\alpha_3, 2) \wedge (\alpha_4, 1)$ has 50% probability to be classified in D_3 and 50% probability to be classified in D_4. The conflicts should be differentiated and α_2 is also an essential attribute.

Therefore, although the β-distribution reduct is a more rigorous definition, it still has the limitations that it may exclude some essential attributes. This limitation is due to the "indifferentiation" of the different conflict decision rules. This may lose some hidden classification information in the original DT and weaken the classification ability.

IV. β-CONSISTENT NOTION FOR A DT IN VPRS

In this section, a β-consistent notion for a DT in VPRS is proposed to analyze the limitations mentioned in the previous section.

A. The consistent DT

Definition 3: A DT $S = (U, A = C \cup D, V, f)$ is a consistent DT [21] if $\forall x, y \in U$ where $x \neq y$, we have:

$$f(x, C) = f(y, C) \Rightarrow f(x, D) = f(y, D). \tag{14}$$

This definition can be also represented from the perspective of indiscernibility relation as $IND(C) \subseteq IND(D)$. However, in VPRS model, because the majority inclusion relation is applied, some inconsistent DT in RS can be re-evaluated. Under VPRS models with different β, a certain DT may change its consistency property. In our paper, a new notion called β-consistency is proposed for VPRS.

B. The β-consistent DT

Given a DT $S = (U, A = C \cup D, V, f)$ and a precision parameter β, $\forall x, y \in U$, $x \neq y$, for a certain equivalent class C_i, if $x \in C_i$, $y \in C_i$, $\exists D_j$ satisfies that $x \in POS_C^\beta(D_j)$, $y \in POS_C^\beta(D_j)$, then we define this DT as β-consistent under the precision β.

Definition 4: A DT $S = (U, A = C \cup D, V, f)$ is a β-consistent DT in VPRS if $\forall x, y \in U$ where $x \neq y$, we have:

$$f(x, C) = f(y, C) \Rightarrow x, y \in POS_C^\beta(D_j) \tag{15}$$

Based on this definition, given a precision parameter β , some inconsistent DT in RS becomes β-consistent in VPRS. The sample DT 1 is an example, obviously, it is an inconsistent DT: for sample points o_2 and o_5, we have $o_2 \in C_2$ and $o_5 \in C_2$, i.e., $f(o_2, C) = f(o_5, C)$, meanwhile, $o_2 \in D_1$, $o_5 \in D_2$, i.e., $f(o_2, D) \neq f(o_5, D)$, which is in conflict with eq.(14), therefore this DT is not consistent.

However, when we analyze the same DT in VPRS model, setting $\beta = 0.59$, we have equivalent classes:
$C_1 = \{o_1\}$,
$C_2 = \{o_2, o_5, o_7\}$,
$C_3 = \{o_3\}$,
$C_4 = \{o_4\}$,
$C_5 = \{o_6\}$,
and decision classes:
$D_1 = \{o_1, o_2, o_3\}$,
$D_2 = \{o_4, o_5, o_6, o_7\}$.
So we have:
$POS_C^{0.59}(D_1) = \underline{C}^{0.59}(D_1) = \{o_1, o_3\}$,
$POS_C^{0.59}(D_2) = \underline{C}^{0.59}(D_2) = \{o_2, o_5, o_7, o_4, o_6\}$.
For any two elements in U of sample DT 1, if they belong to the same C_i, we can definitely find a D_j, and both of these two elements belong to $POS_C^{0.59}(D_j)$. So, this DT is β-consistent when $\beta = 0.59$ in VPRS.

Moreover, for DT 1, if we set $\beta = 0.67$, the β-consistent property will change: the β-positive region of D_1 becomes:
$POS_C^{0.67}(D_1) = \{o_1\}$,
and the β-positive region of D_2 is :
$POS_C^{0.67}(D_2) = \{o_4, o_6, o_7\}$,
we will observe that the elements o_2, o_5 and o_7 in C_2 cannot be included by positive region of any D_j since that $\omega(C_2, D_1) = 2/3 < 0.67$ and $\omega(C_2, D_2) = 1/3 < 0.67$, i.e., for o_2, o_5 and o_7, we cannot find a D_j to satisfy eq.(15). So, when $\beta = 0.67$, this DT is **NOT** a β-consistent DT (or β-inconsistent DT for presentation simplicity).

From the perspective of indiscernibility relation, a DT is β-consistent meaning that $IND(C)$ and $IND(D)$ are not in conflict when the classification precision is not higher than β in VPRS, and DT 1 shows that a DT may change its β-consistency property when varying the value of β.

Theorem 1: for $\beta_1, \beta_2 \in (0.5, 1]$, set $\beta_2 \leq \beta_1$, a β_1-consistent DT must be a β_2-consistent DT.

Proof: if a DT is β_1-consistent, that means all the elements in each C_i can be classified into a β_1-positive region of a certain D_j. According to eq.(7) and eq.(8), $\omega(C_i, D_j) = (|C_i \cap D_j| / |C_i|) \geq \beta_1$, since $\beta_2 \leq \beta_1$, $\omega(C_i, D_j) \geq \beta_2$. Therefore, any element in C_i can also be classified into β_2-positive region of a certain D_j. So this DT is β_2-consistent.

From the definition, the 1-consistent notion in VPRS is equal to the consistent notion in RS. So the β-consistent notion can be considered as an extension of classical consistent notion, and we easily have the following deduction:

Deduction 1: a consistent DT must be a β-consistent DT, $\forall \beta \in (0.5, 1]$.

Eq.(16) and eq.(17) give two notions for further analysis:

$$G_C^\beta(x) = \{D_j \,\Big|\, x \in POS_C^\beta(D_{\mathrm{j}})\}, x \in U \tag{16}$$

$$R^{\beta}(C_i) = \{G_C^{\beta}(x) \,| x \in C_i \} \tag{17}$$

Samples in collection $R^{\beta}(C_i)$ are constituted by the decision classes. For each C_i, if any element in it also belongs to the β-positive region of a certain decision class D_j, we put the decision class D_j into a corresponding collection $R^{\beta}(C_i)$.

Theorem 2: when $\beta \in (0.5, 1]$, $\left|R^{\beta}(C_i)\right| \in \{0, 1\}$, in other words, $\left|R^{\beta}(C_i)\right| < 2$.

Proof: It is easy to find the example of $\left|R^{\beta}(C_i)\right| = 0$ and $\left|R^{\beta}(C_i)\right| = 1$. Observe sample DT 1, Set $\beta = 0.6$, we can get $R^{0.6}(C_{\{o_1\}}) = \{D_1\}$, so $\left|R^{0.6}(C_{\{o_1\}})\right| = 1$, and $R^{0.6}(C_{\{o_2,o_3\}}) = \emptyset$, so $\left|R^{0.6}(C_{\{o_2,o_3\}})\right| = 0$.

Denote $|C_i| = N$, if $\left|R^{\beta}(C_i)\right| = 2$, setting $R(C_i) = (D_{j'}, D_{j''})$, where $D_{j'} \neq D_{j''}$. From eq. (4), we have $D_{j'} \cap D_{j''} = \emptyset$. m is the number of elements in C_i belonging to $D_{j'}$, n is the number of elements in C_i belonging to $D_{j''}$, according to the definition of β-positive region, $m/N \geq \beta$, and $n/N \geq \beta$, $0.5 < \beta \leq 1$, so $m/N > 0.5$, $n/N > 0.5$, therefore $(m+n)/N > 1$, this is in conflict with $m+n = N$. The condition is similar when $\left|R^{\beta}(C_i)\right| = 3, 4, ...\infty$.

Theorem 2 gives the fact that when $\beta > 0.5$, one sample point can only support one decision rule.

From the definition of $R^{\beta}(C_i)$, given β, if a DT is β-consistent, $\forall C_i$, $\left|R^{\beta}(C_i)\right| = 1$; meanwhile, if in a DT, $\forall C_i$, we have $\left|R^{\beta}(C_i)\right| = 1$, this DT is β-consistent. Thus:

$$\left|R^{\beta}(C_i)\right| = 1, \forall C_i \tag{18}$$

is the sufficient and necessary condition for that a DT is β-consistent. If in a DT, $\exists C_i$, where $\left|R^{\beta}(C_i)\right| = 0$, we can deduce that this DT is β-inconsistent, and the equivalent class C_i with $\left|R^{\beta}(C_i)\right| = 0$ is called *undecidable equivalent class* of this DT in our investigation. Moreover, the number of undecidable equivalent classes shows the number of decision conflict conditions in the decision rules of a DT.

V. THE RELATIONSHIPS BETWEEN THE REDUCTS AND β-CONSISTENT PROPERTY OF A DT IN VPRS MODEL

In the previous sections, we have already discussed different definitions of attribute reduct in VPRS model: the β-reduct has already been proved to have limitations in previous investigations; moreover, our example in Table II has illustrated that the β-distribution reduct also has limitations. Consequently, we proposed the definition of β-consistent property of a DT in VPRS model, which is an extension of the consistent notion in classical RS model. In this section, we will analyze the relationship between different definitions of reduct and the β-consistent property of a certain DT in VPRS.

If a DT is 1-consistent, the β-reduct is equal to the β-distribution reduct, where $\beta \in (0.5, 1]$.

If a DT is 1-consistent, that means for each equivalent class C_i, $\exists D_j$, where $(C_i \cap D_j)/D_j = 1$. In Theorem 2, we have proved that when $\beta \in (0.5, 1]$, $\left|R^{\beta}(C_i)\right| < 2$. Therefore, $\forall \beta \in (0.5, 1)$, $POS_C^{\beta}(D_j) = POS_C^{1}(D_j)$. So the VPRS evolves a classical RS when the DT is 1-consistent.

In classical RS, the monotonicity of classification quality and positive regions are always uniform. The same classification quality degree means the consistency of positive regions. Accordingly, the β-reduct is equal to the β-distribution reduct, and therefore the decision rule conflict can be avoided.

If an inconsistent DT is β-consistent for $\beta \in (0.5, 1)$, the β-reduct may lead to the decision rule conflict, while the β-distribution reduct can keep the decision rule consistent. Moreover, the β-distribution reduct can also avoid losing hidden classification information.

If a DT is inconsistent but β-consistent where $\beta \neq 1$, that means there are decision conflicts in this DT initially, but these decision conflicts can be eliminated in VPRS when the classification precision requirement is lower than β as the majority inclusion relation is employed.

For example, in sample DT 1, a decision rule conflict is:
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 1)$, by o_2;
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 0)$, by o_5 and o_7.

This DT is obviously an inconsistent DT, But in VPRS with $\beta = 0.59$, since two of the three sample points have supported $(d, 0)$, so the classification precision is $2/3 \geq 0.59$, thus the majority inclusion relation requirement is satisfied, which means that in this conflict, the majority (o_5 and o_7) of the three samples support the latter decision rule So this conflict can be eliminated and we use the decision rule:$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 0)$ in VPRS model. Furthermore, in general, for an inconsistent DT, if it is β-consistent, all the decision conflicts can be eliminated under a classification precision lower than β. However, for such DT, only calculating the β-reduct may cause the decision rule inconsistent as shown in sample DT 1.

On the other hand, for a β-consistent DT, the β-distribution reduct can not only keep the classification quality degrees consistent but also keep the decision rules consistent: decision rules of a DT is determined by the positive regions[20], the β-distribution reduct requires the β-positive regions to be kept consistent, so the decision rules are also kept consistent after reduction. Additionally, if a DT is β-consistent, the β-distribution reduct of this DT can avoid losing hidden information. The loss of hidden information is caused by neglecting the differences among different decision conflicts. From the definition of β-consistency, in a β-consistent DT, for each equivalent class C_i, $\exists D_j$, where $C_i \subseteq POS_C^{\beta}(D_j)$. Thus, $\forall C_i$, a definite decision rule: $\bigwedge(c, f(C_i, c)) \rightarrow (d, f(D_j, d))$ without any decision conflict can be extracted. Therefore the β-distribution reduct is able to avoid losing hidden classification information.

If a DT is β-inconsistent, $\beta \in (0.5, 1]$, use N to denote the number of undecidable equivalent classes, if $N = 1$, the β-distribution reduct can keep the decision rule consistent and will not lose information in the original DT.

The number of undecidable equivalent classes reflects the number of decision conflict conditions of a DT, moreover, the problem of losing hidden information of β-distribution reduct is led by ignoring the differences among various conflicts in the procedure of calculating the β-distribution reduct. For a DT, if $N = 1$, there is only one undecidable equivalent class, i.e., there is only one decision conflict in the decision rules.

TABLE III. SAMPLE DT 3

U	α_1	α_2	α_3	α_4	α_5	d
x_1	1	1	1	1	1	1
x_2	1	1	0	1	1	1
x_3	0	0	1	0	0	1
x_4	1	1	2	1	1	1
x_5	1	1	0	1	0	2
x_6	1	1	0	1	1	2
x_7	0	0	1	2	1	2
x_8	1	1	0	1	1	2
x_9	1	1	2	1	1	2

Setting this undecidable equivalent class as C_i, according to eq.(3), $\forall x \in C_i$, x can support $\bigwedge(c, f(C_i, c)) \rightarrow (d, \emptyset)$; moreover, as there is only one decision conflict in the original DT, differentiation of various conflicts is not necessary. When calculating the reduct, we can consider $\emptyset$ as a special decision class $D_\emptyset$ and set that $\forall D_j$, $D_j \neq D_\emptyset$. The sample DT 3 is an example.

In this DT, we have decision classes:
$D_1 = \{x_1, x_2, x_3, x_4\}$,
$D_2 = \{x_5, x_6, x_7, x_8, x_9\}$.
Also, we can obtain:
$C_1 = \{x_1\}$,
$C_2 = \{x_2, x_6, x_8\}$,
$C_3 = \{x_3\}$,
$C_4 = \{x_4, x_9\}$,
$C_5 = \{x_5\}$,
$C_6 = \{x_7\}$.

In sample DT 3, set $\beta = 0.6$, $C_{\{x_4,x_9\}}$ is the only undecidable equivalent class, so $N = 1$. $\{\alpha_3, \alpha_4\}$ is a β-distribution reduct for this DT. In sample DT 3, the decision rules are:
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, 1)$, by x_1;
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 0) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, 2)$, by x_2, x_6 and x_8;
$(\alpha_1, 0) \wedge (\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 0) \wedge (\alpha_5, 0) \rightarrow (d, 1)$, by x_3;
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 0) \wedge (\alpha_4, 1) \wedge (\alpha_5, 0) \rightarrow (d, 2)$, by x_5;
$(\alpha_1, 0) \wedge (\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 2) \wedge (\alpha_5, 1) \rightarrow (d, 2)$, by x_7;
$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, (1 or 2))$, by x_4 and x_9.

After the reduction, the decision rules are:
$(\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 1)$, by x_1;
$(\alpha_3, 0) \wedge (\alpha_4, 1) \rightarrow (d, 2)$, by x_2, x_5, x_6 and x_8;
$(\alpha_3, 1) \wedge (\alpha_4, 0) \rightarrow (d, 1)$, by x_3;
$(\alpha_3, 1) \wedge (\alpha_4, 2) \rightarrow (d, 2)$, by x_7;
$(\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (1 or 2))$,by x_4 and x_9.
all decision rules are kept consistent after reduction and no essential information lost.

If a DT is β-inconsistent, $\beta \in (0.5, 1]$, use N to denote the number of undecidable equivalent classes, if $N > 1$, the β-distribution reduct can keep the decision rule consistent but may lead to loss of hidden information. DT If a DT is β-inconsistent, and there are more than one undecidable equivalent classes, the β-distribution reduct can avoid the decision rule conflict by keeping the β-positive regions consistent. But the β-distribution reduct deals with all the undecidable equivalent classes by excluding them from all the β-positive regions. In so doing the β-distribution reduct neglects the differences among different undecidable equivalent classes, so that the β-distribution reduct may lose some essential information in DT and weaken the classification ability of the original DT. The sample DT 2 discussed above demonstrates this problem in β-inconsistent table with more than one undecidable equivalent classes.

VI. β-COMPLETE REDUCT CALCULATION

We have already discussed the relationship between the reduct and the β-consistent property of a DT in VPRS model. We see that in a VRPS model, set $\beta \in (0.5, 1]$, if a DT is 1-consistent, we only need to calculate the β-reduct of the DT; for a β-consistent DT, we have to calculate the β-distribution reduct to make sure the reduct not bringing in the decision rule conflicts; if a DT is β-inconsistent, the condition becomes complicated: if there is only one undecidable equivalent class, we only need to calculate the β-distribution reduct; Otherwise, only calculating the β-distribution reduct may lead to loss of some important information and lower the classification ability of the original DT. Accordingly, in this section, a new method will be proposed to deal with the reduct problem for the β-inconsistent DT with more than one undecidable equivalent classes.

In this section, we will first give a DT splitting method to deal with the reduct problem for β-inconsistent DT.

Theorem 3: In VPRS model, given $\beta \in (0.5, 1]$, if a DT is β-inconsistent, it can be split into two DTs, one is a β-consistent DT, the other one is a complete β-inconsistent DT in which all equivalent classes are undecidable.

Proof: If a DT $S = (U, A = C \cup D, V, f)$ is β-inconsistent, from eq.(17), $\exists C_i$, $\left|R^\beta(C_i)\right| = 0$. Since we have proved that $\left|R^\beta(C_i)\right| \in \{0, 1\}$ in Theorem 2, we can split the DT into two DTs S_1 and S_2. Elements in S_1 satisfy $\{x \,|x \in C_i, \left|R^\beta(C_i)\right| = 1\}$. Thus, $\forall C_i \in S_1$, $\left|R^\beta(C_i)\right| = 1$. Accordingly, S_1 is β-consistent. In the meanwhile, elements in S_2 satisfy $\{x \,|x \in C_i, \left|R^\beta(C_i)\right| = 0\}$, that means all equivalent classes are undecidable.

Then a β-inconsistent DT can be split as Figure 1, and Figure 2 uses sample DT 2 to demonstrate this DT splitting method with $\beta = 0.6$. The sample DT 2 is split into a β-consistent DT (white background) formed by the consistent equivalent classes; and a complete β-inconsistent DT (dark background) formed by elements that are all in conflict.

A method for all the reduct (denoted as RED) of a β-inconsistent DT can be given as following steps:

Input: A DT $S = (U, A = C \cup D, V, f)$, β

Output: the Reduct of the input DT RED

Step 1: Find all the equivalent classes C_i in this DT, where $\{C_i, i = 1, 2, \cdots, m\} = U/IND(C)$;

Step 2: For each C_i, calculate $\left|R^\beta(C_i)\right|$, and use N to denote the number of C_i, whose $\left|R^\beta(C_i)\right| = 0$;

Step 3: If $N = 1$, consider that the only undecidable equivalent class can be classified into $POS_c^\beta(D_\emptyset)$, $D_\emptyset \neq D_j$,

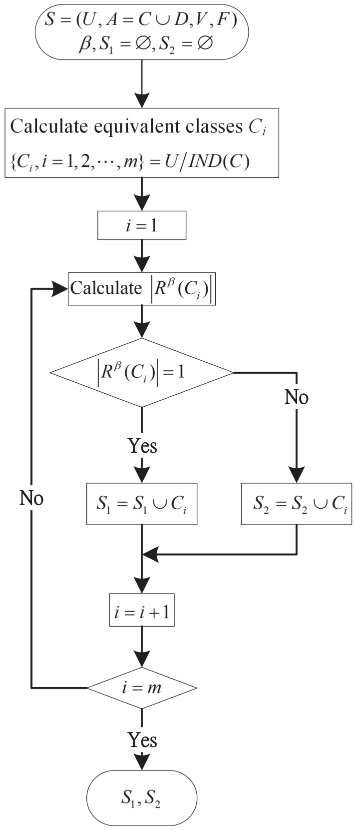

Fig. 1. The β-inconsistent DT splitting approach

U	a_1	a_2	a_3	a_4	a_5	d
x_1	1	1	1	1	1	1
x_2	1	1	0	1	1	1
x_3	0	0	1	0	0	1
x_5	1	1	0	1	0	2
x_6	1	1	0	1	1	2
x_7	0	0	1	2	1	2
x_8	1	1	0	1	1	2

$|R^\beta(C_i)|=1$

U	a_1	a_2	a_3	a_4	a_5	d
x_1	1	1	1	1	1	1
x_2	1	1	0	1	1	1
x_3	0	0	1	0	0	1
x_4	1	1	2	1	1	1
x_5	1	1	0	1	0	2
x_6	1	1	0	1	1	2
x_7	0	0	1	2	1	2
x_8	1	1	0	1	1	2
x_9	1	1	2	1	1	2
x_{10}	1	0	2	1	1	3
x_{11}	1	0	2	1	1	4

$|R^\beta(C_i)|=0$

U	a_1	a_2	a_3	a_4	a_5	d
x_4	1	1	2	1	1	1
x_9	1	1	2	1	1	2
x_{10}	1	0	2	1	1	3
x_{11}	1	0	2	1	1	4

Fig. 2. The DT splitting approach on sample DT 2

where $\{D_j, j = 1, 2, \cdots, n\} = U/IND(C)$. Then calculate the β-distribution reduct of this DT, a detailed algorithm for β-distribution reduct is given by [17] or [22], return the result as RED;

Step 4: If $N > 1$, split the DT into a β-consistent DT S_1 and a complete β-inconsistent DT S_2;

Step 5: Calculate a β-distribution reduct for S_1, and denote the reduct as RED_{S_1};

Step 6: For DT S_2, only consider the condition attributes, a DT without decision part is called Attribute-value Table (AT) [1], convert S_2 into an AT by deleting its decision attributes and calculate the reduct of this AT according to [15], denoted as RED_{S_2};

Step 7: return $RED = RED_{S_1} \cup RED_{S_2}$.

For sample DT 2, we split it into two DTs as shown in Figure 2. For the β-consistent DT, we can obtain its reduct $RED_{S_1} = \{\alpha_3, \alpha_4\}$; meanwhile, the reduct of the complete β-inconsistent DT is $RED_{S_2} = \{\alpha_2\}$, accordingly, $RED = RED_{S_1} \cup RED_{S_2} = \{\alpha_2, \alpha_3, \alpha_4\}$.The total decision rules obtained in the original sample DT 2 are:

$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, 1)$, by x_1;

$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 0) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, 2)$, by x_2, x_6 and x_8;

$(\alpha_1, 0) \wedge (\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 0) \wedge (\alpha_5, 0) \rightarrow (d, 1)$, by x_3;

$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 0) \wedge (\alpha_4, 1) \wedge (\alpha_5, 0) \rightarrow (d, 2)$, by x_5;

$(\alpha_1, 0) \wedge (\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 2) \wedge (\alpha_5, 1) \rightarrow (d, 2)$, by x_7;

$(\alpha_1, 1) \wedge (\alpha_2, 1) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, (1or2))$, by x_4 and x_9;

$(\alpha_1, 1) \wedge (\alpha_2, 0) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \wedge (\alpha_5, 1) \rightarrow (d, (3or4))$, by x_{10} and x_{11}.

the total decision rules extracted from the obtained reduct $RED = \{\alpha_2, \alpha_3, \alpha_4\}$ are:

$(\alpha_2, 1) \wedge (\alpha_3, 1) \wedge (\alpha_4, 1) \rightarrow (d, 1)$, by x_1;

$(\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 0) \rightarrow (d, 1)$, by x_3;

$(\alpha_2, 1) \wedge (\alpha_3, 0) \wedge (\alpha_4, 1) \rightarrow (d, 2)$, by x_2, x_6 ,x_5 and x_8;

$(\alpha_2, 0) \wedge (\alpha_3, 1) \wedge (\alpha_4, 2) \rightarrow (d, 2)$, by x_7;

$(\alpha_2, 1) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (1or2))$, by x_4 and x_9;

$(\alpha_2, 0) \wedge (\alpha_3, 2) \wedge (\alpha_4, 1) \rightarrow (d, (3or4))$, by x_{10} and x_{11}.

In the previous sections, we have already discussed different definitions of attribute reduct in VPRS model: the β-reduct has already been proved to have limitations in previous investigations; moreover, our example in Table III has illustrated that the β-distribution reduct also has limitations. Consequently, we proposed the definition of β-consistent property of a DT in VPRS model, which is an extension of the consistent notion in classical RS model. In this section, we will analyze the relationship between different definitions of reduct and the β-consistent property of a certain DT in VPRS. The comparison of the decision rules shows that the obtained reduct $RED = \{\alpha_2, \alpha_3, \alpha_4\}$ cannot only keep the decision rule consistent, but also discern the two different undecidable equivalent classes $C_4 = \{x_4, x_9\}$ and $C_7 = \{x_{10}, x_{11}\}$. It conserves the hidden information that, given the condition $(\alpha_3, 2) \wedge (\alpha_4, 1)$, a sample point x with $f(x, \alpha_2) = 1$ has 50% probability to be classified in D_1 and 50% probability to be classified in D_2, and a sample point with $f(x, \alpha_2) = 0$ has

50% probability to be classified in D_3 and 50% probability to be classified in D_4.

The obtained reduct by this approach is called as the β-complete reduct in our paper. DT S_1 contains all the consistent equivalent classes in the original DT, accordingly, the β-distribution reduct of S_1 can keep all decision rules in the original DT consistent. Meanwhile, reduct of S_2 reflects the hidden classification information of the original DT. Then we get the union set of the two reducts as the β-complete reduct. It not only makes sure the β-positive regions are consistent in the reduct, but also considers the hidden information in the inconsistent equivalent classes of the original DT in order to avoid the deterioration of classification ability. Obviously, the requirement of the β-complete reduct is more rigorous than that of the β-distribution reduct. We may observe that the β-distribution reduct is a subset of the β-complete reduct, that means a β-complete reduct must be a β-distribution reduct, but the inverse proposition is not aways true.

VII. CONCLUSION

The β-distribution reduct is seen as a modified version of the β-reduct in the VPRS model. However, in our investigation, from some sample instances, we find that the β-distribution reduct also has limitations. It may neglect the differences among the different conflicts of a DT. Accordingly, a β-complement reduct is proposed in this paper. By splitting the β-inconsistent DT into two DTs and combining the reduct of the two DTs together, we can obtain the β-complement reduct, and this β-complement reduct can avoid the loss of hidden classification information of the original DT.

We also analyze the hierarchical relationship between the proposed β-consistent notion and different definitions of reduct in VPRS model. For 1-consistent DT, we only need to calculate the β-reduct, if $\beta \in (0.5, 1)$, the β-distribution reduct can perform well for a β-consistent DT. For β-inconsistent DT, if there is only one undecidable class, we need to calculate the β-distribution reduct; otherwise, we need to split the β-inconsistent DT to get the β-complete reduct. Based on this investigation, for a given DT and VPRS with certain β, we can analyze the β-consistency property of this DT first, and then choose the proper definition of reduct for this DT, and calculate it with a suitable algorithm in [15] or [17].

About the further work, some more general algorithms to β-complete reduct will be introduced and evaluated; what is more, the reduct for some other important RS extension models will be investigated [23]; finally, we will work further on our investigation from the probabilistic aspect and fuzzy logic aspect[24].

ACKNOWLEDGMENT

The authors would like to acknowledge the partial support of the CRG grants G-YL14 and G-YM07 of The Hong Kong Polytechnic University.

REFERENCES

[1] Z. Pawlak, "Rough sets," *International Journal of Parallel Programming*, vol. 11, no. 5, pp. 341–356, 1982.

[2] Z. Pawlak and A. Skowron, "Rudiments of rough sets," *Information Sciences*, vol. 177, no. 1, pp. 3–27, 2007.

[3] Z. Pawlak, J. Grzymala-Busse, R. Slowinski, and W. Ziarko, "Rough sets," *Communications of the ACM*, vol. 38, no. 11, 1995.

[4] K. Thangavel and A. Pethalakshmi, "Dimensionality reduction based on rough set theory: A review," *Applied Soft Computing*, vol. 9, no. 1, pp. 1–12, 2009.

[5] K. Kaneiwa, "A rough set approach to multiple dataset analysis," *Applied Soft Computing*, vol. 11, no. 2, pp. 2538 – 2547, 2011.

[6] Y. Yao, "Three-way decisions with probabilistic rough sets," *Information Sciences*, vol. 180, no. 3, pp. 341 – 353, 2010.

[7] M. J. Beynon and M. J. Peel, "Variable precision rough set theory and data discretisation: an application to corporate failure prediction," *Omega*, vol. 29, no. 6, pp. 561 – 576, 2001.

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

[9] A. An, N. Shan, C. equivalent class ristine Chan, N. Cercone, and W. Ziarko, "Discovering rules for water demand prediction: An enhanced rough-set approach," *Engineering Applications of Artificial Intelligence*, vol. 9, no. 6, pp. 645–653, 1996.

[10] W. Li, S. Chen, and B. Wang, "A variable precision rough set based modeling method for pulsed gtaw," *The International Journal of Advanced Manufacturing Technology*, vol. 36, no. 11, pp. 1072–1079, 2008.

[11] M. Inuiguchi, "Attribute reduction in variable precision rough set model," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 14, no. 04, pp. 461–479, 2006.

[12] X. Zhang, Z. Mo, F. Xiong, and W. Cheng, "Comparative study of variable precision rough set model and graded rough set model," *International Journal of Approximate Reasoning*, vol. 53, no. 1, pp. 104–116, 2012.

[13] J. Wang and D. Miao, "Analysis on attribute reduction strategies of rough set," *Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 189–192, 1998.

[14] R. W. Swiniarski and A. Skowron, "Rough set methods in feature selection and recognition," *Pattern Recognition Letters*, vol. 24, no. 6, pp. 833 – 849, 2003.

[15] C. Rauszer and A. Skowron, "The discernibility matrices and functions in information systems," in *Intelligent Decision Support-Handbook of Applications and Advances of the Rough Sets Theory, Knowledge Engineering and Problem Solving*, vol. 11. Norwell, MA, USA: Kluwer Academic Publishers, 1992, pp. 331–362.

[16] M. Beynon, "Reducts within the variable precision rough sets model: A further investigation," *European Journal of Operational Research*, vol. 134, no. 3, pp. 592 – 605, 2001.

[17] J. Mi, W. Wu, and W. Zhang, "Approaches to knowledge reduction based on variable precision rough set model," *Information Sciences*, vol. 159, no. 3, pp. 255–272, 2004.

[18] J. Wang and J. Zhou, "Research of reduct features in the variable precision rough set model," *Neurocomputing*, vol. 72, no. 10, pp. 2643–2648, 2009.

[19] J. Katzberg and W. Ziarko, "Variable precision rough sets with asymmetric bounds," in *Rough Sets, Fuzzy Sets and Knowledge Discovery*, ser. Workshops in Computing, W. Ziarko, Ed. Springer London, 1994, pp. 167–177.

[20] J. Zhou and D. Miao, "β-interval attribute reduction in variable precision rough set model," *Soft Computing*, vol. 15, pp. 1643–1656, 2011.

[21] Z. Pawlak, "Rough sets and decision tables," in *Computation Theory*, ser. Lecture Notes in Computer Science, A. Skowron, Ed. Springer Berlin Heidelberg, 1985, vol. 208, pp. 187–196.

[22] J. N. Liu, Y. Hu, and Y. He, "A set covering based approach to find the reduct of variable precision rough set," *Information Sciences*, 2014.

[23] W. Z. Wu, J. S. Mi, and W. X. Zhang, "Generalized fuzzy rough sets," *Information Sciences*, vol. 151, no. 0, pp. 263 – 282, 2003.

[24] Y.Y.Yao, "A comparative study of fuzzy sets and rough sets," *Information Sciences*, vol. 109, no. 14, pp. 227 – 242, 1998.