Journal of Computer Science and Cybernetics, V.40, N.1 (2024), 37–51
DOI no. 10.15625/1813-9663/19218

A NOVEL EXTENSION METHOD OF VPFRS MODE FOR ATTRIBUTE REDUCTION PROBLEM IN NUMERICAL DECISION TABLES

PHAM MINH NGOC HA[1], TRAN THANH DAI[2,*], NGUYEN MANH HUNG[3], HOANG TUAN DUNG[4]

[1] *Academy of Finance, Ha Noi, Viet Nam*
[2] *University of Economics and Industrial Technology, Ha Noi, Viet Nam*
[3] *Military Technical Academy, Ha Noi, Viet Nam*
[4] *VNU University of Science, Ha Noi, Viet Nam*

Abstract. Attribute reduction is an essential application of the Rough Set (RS) theory that has been receiving the attention of many researchers. Up to now, attribute reduction methods to improve classification accuracy on noisy datasets following the IFRS approach still have many limitations in terms of computation time. In this paper, we use the variable precision method on approximate operators of the FRS model to expand measures to effectively evaluate attributes on noisy datasets. The main contributions of this paper include: 1) proposing new approximation operations to extend VPFRS to VPOFRS, 2) proposing some measures based on VPOFRS to evaluate the consistency degree of the decision table and the dependence degree of the attribute, 3) proposing an attribute reduction algorithm VPOFRS_AR. Experimental results on noisy datasets from UCI show that the proposed method not only improves the noise for the reduct but the algorithm's execution time is faster than other algorithms.

Keywords. Rough set; Variable precision rough set; Fuzzy rough set; Variable precision fuzzy rough set; Attribute reduction.

1. INTRODUCTION

This section introduces the problem of attribute reduction using the rough set approach, presents some related research, points out limited issues of existing noise improvement studies, and defines the research goals.

Rough set (RS) is an extended theoretical model of set theory, effectively applied in handling all uncertain, incomplete, and inconsistent data. Attribute reduction or attribute selection is an essential application of RS, currently receiving the attention of many researchers. Attribute reduction is an essential data preprocessing step, widely applied in machine learning and data mining. Some specific applications include data classification [1,2], handwriting recognition [3,4], speech recognition [5,6], spam detection and classification [7,8], and decision support [9,10]. According to the RS approach, attribute reduction is building measures

Corresponding author.
E-mail addresses: phamminhngocha@hvtc.edu.vn (P.M.N.Ha); ttdaiuneti@gmail.com (T.T.Dai); manhhungk12@mta.edu.vn (N.M.Hung); tuandunghg01@gmail.com (H.T.Dung).

to evaluate and select attributes. The metrics are often based on information granularity and RS's positive domain (Positive - POS).

- The information granularity by the RS approach is an equivalence class according to a predefined equivalence relationship. Then, the equivalence relation divides a set into equivalence classes constituting a partition, meaning that the equivalence classes do not intersect. Then, granularity computation [11, 12] and information entropy computation [13–15] methods are used to construct measures. These measures often work directly on partitions, so they have the advantage of execution time, but the accuracy depends on the relational formulas.

- The positive domain of RS is the set of objects of the lower approximation sets. This is a critical concept applied to construct metrics to evaluate the consistency of the decision table and the dependency of attributes [16]. Based on the attribute dependency measure, several attribute reduction methods for decision tables are proposed [13, 17]. From the viewpoint of granularity computing, RS's approximation operations are the process of smoothing partitions into a set of certain objects (lower set) and a set of uncertain objects (upper set). Therefore, measurements built on lower sets give higher accuracy than measurements on partitions, but execution time may be slower on covers. However, the traditional RS approach measures have ineffective accuracy for data sets with high misclassification rates (noise). Therefore, Ziarko first proposed the variable precision rough set (VPRS) model to adjust for objects that can be misclassified in the lower and upper approximation sets [18].

The metrics constructed using the traditional RS approach are only suitable for categorical value decision tables. The neighbor relations are proposed to construct equivalence classes for attribute reduction on the numerical value decision tables. Through neighbor relationships, equivalence classes constitute covers [17]. The disadvantage of the neighbor relationships is that it is difficult to determine an appropriate neighbor threshold for all objects [19, 20]. Therefore, the concept of similarity relationship is proposed. Most similarity relationships between objects are normalized to the interval [0, 1], so the equivalence class of an object is extended into a fuzzy equivalence class that constitutes fuzzy partitions. Based on fuzzy partitions, fuzzy information granular computation [11, 21], fuzzy entropy [22], and fuzzy positive domain computation methods [23] are proposed to construct measures. These measures are called the measures by the fuzzy rough set (FRS) [24] approach.

The current FRS approach measures also face many challenges on numerical decision tables with low consistency [25]. Therefore, Zhao and colleagues extended the FRS to the VPFRS model [26]. Zhang extends the FRS model to the Intuitionistic Fuzzy Rough Set (IFRS) model [25]. Research results show that these extended methods improve classification accuracy well on noisy datasets [26–28], but the execution time could be improved. Due to intuitionistic fuzzy sets (IFS) characteristics, the IFRS model is twice as complex as the FRS and VPRS models. Furthermore, the FRS and VPFRS models still calculate the membership of all objects in the decision table. Recently, Yi and Yu proposed the VPNRS model [29], but the approximation operations are still the same compared to VPRS. Therefore, we propose a variable precision method based on the objects of each decision class in the decision table. At that time, the number of variable precision objects is significantly small. We denoted this extended model as the VPOFRS. The main contributions of this study include: 1) proposing new approximation operations to extend VPFRS to VPOFRS, 2) proposing some measures based on VPOFRS to evaluate the consistency degree of the decision table and the dependence degree of the attribute, 3) proposing an attribute reduction algorithm

Table 1: The numeric decision table

U	a	b	c	d	e	f	D
u_1	1.0	0.4	0.8	0.2	1.0	0.0	0
u_2	1.0	0.4	0.2	0.4	0.2	0.8	1
u_3	0.8	0.6	1.0	0.0	0.6	0.4	0
u_4	0.2	0.6	0.8	0.2	0.0	1.0	1
u_5	0.2	0.8	0.8	0.2	0.0	1.0	1
u_6	0.2	0.8	0.2	0.8	0.0	1.0	0

VPOFRS_AR. Experimental results show that the proposed VPOFRS_AR is efficient in computation time while the reduct has the same size and accuracy as published algorithms.

The rest of this paper is organized as follows. Section II reviews VPRS and VPFRS models. Section III proposes a novel VPOFRS extended from the VPOFRS model, invests some properties, and constructs some metrics to evaluate the decision table's consistency and the attributes' dependency. In Section IV, we make the VPOFRS_AR algorithm to find the reducts using the filter method. Section V presents some notable results of the proposed method compared to other methods. The final section concludes this paper.

2. PRELIMINARIES

This section recalls some basic concepts of the VPRS [18] and the VPFRS model [26]. First, we recall the concept of a numerical decision table.

The numeric decision Table 1 is represented by the tuple $DT = (U, C, D, f)$, where $C \cap D = \emptyset$, U represents a non-empty collection of objects, C represents a non-empty set of conditional attributes, and D represents a decision attribute. The function f_c determines a value in V_c for each $u \in U$ and $c \in C$, where V_c is the value domain of the attribute c and V_c belongs to $\mathbb{R}$. The function f_D determines a value in V_D for each $u \in U$, where V_D is the value domain value of D and V_D belongs to $\mathbb{N}$.

2.1. VPRS model

In this section, we recall some basic concepts of the VPRS [18] model expanded from the RS model.

Given $X, Y \subseteq U$, with U non-empty, then X is said to belong to Y if, for every $x \in X$, we have $x \in Y$. The following formula calculates the ratio $X \subseteq Y$.

$$c(X,Y) = \begin{cases} 1 - \frac{card(X \cap Y)}{card(X)} \Leftrightarrow card(X) > 0 \\ 0 \Leftrightarrow card(X) = 0, \end{cases} \tag{1}$$

where card function denotes set cardinality.

Based on the measure of relative misclassification one can define the inclusion relationship between X and Y without explicitly using a general quantifier

$$X \subseteq Y \Leftrightarrow c(X,Y) = 0. \tag{2}$$

Based on this assumption the majority inclusion relation is defined as

$$X \overset{\beta}{\subseteq} Y \Leftrightarrow c(X,Y) \leqslant \beta. \tag{3}$$

Then, the β-lower approximation of $X \subseteq U$ is defined as

$$\underline{R}_\beta X = \bigcup \{E \in U/R : c(E, X) \leqslant \beta\}. \tag{4}$$

The β-upper approximation of $X \subseteq U$ is defined as

$$\bar{R}_\beta X = \bigcup \{E \in U/R : c(E, X) < 1 - \beta\}, \tag{5}$$

where, U/R is a partition according to the equivalence relation R.

2.2. VPFRS model

In this section, we recall some basic concepts of fuzzy equivalence relations and the VPFRS model [26].

Consider $\mathcal{R}$ to be a relation on the non-empty set U, then $\mathcal{R}$ is called a fuzzy equivalence relation if the following requirements are satisfied:

1. Reflexivity $\mathcal{R}(x, x) = 1$,
2. Symmetry $\mathcal{R}(x, y) = \mathcal{R}(y, x)$,
3. min-transitivity $\mathcal{R}(x, y) \geqslant \min(\mathcal{R}(x, z), \mathcal{R}(z, y))$, for every $x, y, z \in U$.

Then lower approximation of element $x \in U$ with respect to $A \subseteq U$ is defined as

$$\underline{\mathcal{R}}A(x) = \inf_{y \in U} \mathcal{I}(\mathcal{R}(x, y), A(y)). \tag{6}$$

The upper approximation of element $x \in U$ with respect to $A \subseteq U$ is defined as

$$\overline{\mathcal{R}}A(x) = \sup_{y \in U} \min(\mathcal{R}(x, y), A(y)). \tag{7}$$

Based on the VP concept of the VPRS model, Zhao proposed the following approximation operations:

The β-lower approximation of element $x \in U$ with respect to $A \subseteq U$ is defined as

$$\underline{\mathcal{R}_\beta}A(x) = \inf_{A(y) \leq \beta} \mathcal{I}(\mathcal{R}(x, y), \beta) \wedge \inf_{A(y) > \beta} \mathcal{I}(\mathcal{R}(x, y), A(y)). \tag{8}$$

The β-upper approximation of element $x \in U$ with respect to $A \subseteq U$ is defined as

$$\overline{\mathcal{R}_\beta}A(x) = \sup_{A(y) \geq 1-\beta} \min(\mathcal{R}(x, y), 1 - \beta) \vee \sup_{A(y) < 1-\beta} \min(\mathcal{R}(x, y), A(y)). \tag{9}$$

3. VARIABLE PRECISION DECISION-BASED FUZZY ROUGH SETS

In this section, we present the method of adjusting the lower approximation set of FRS based on the objects of the decision set (D) in the decision table denoted by VPOFRS.

3.1. Fuzzy approximation space

To construct the fuzzy approximation space, we first propose a fuzzy equivalence relation for two objects x, y as the formula follows

$$\mathcal{R}(x,y) = 1 - |x - y|\,. \tag{10}$$

Given the decision table $DT = (U, C, D)$ and the fuzzy equivalence relation $\mathcal{R}$ on U. Based on the information of the condition attribute C, $\mathcal{R}$ will divide U into $|U|$ fuzzy equivalence class $[u]_{\mathcal{C}}$, which is called the fuzzy partition of U over C, denoted U/C or $\mathcal{C}$. This article uses the symbols $\mathcal{C}$ and U/C interchangeably. Then, the pair $(U, \mathcal{R})$ is called the fuzzy approximation space.

Definition 1. Let $DT = (U, C, D)$ and $\mathcal{P}, \mathcal{Q}$ are fuzzy partitions on U with respect to $P, Q \subseteq C$, then the fuzzy partition of $P \cup Q$ on U denoted by $\mathcal{K}$, which is determined as follows

$$\mathcal{K} = \mathcal{P} \cap \mathcal{Q} = \{[u]_{\mathcal{K}} : u \in U, [u]_{\mathcal{K}} = [u]_{\mathcal{P}} \cap [u]_{\mathcal{Q}}\}. \tag{11}$$

Proposition 1. *Let $DT = (U, C, D)$ and $\mathcal{P}, \mathcal{Q}$ are fuzzy partitions on U with respect to $P, Q \subseteq C$, then $\mathcal{P} \subseteq \mathcal{Q}$ if $Q \subseteq P$.*

Proof. Based on Definition 1, we see that the partition becomes smoother as the number of attributes increases. This proposition has been proven. ■

Definition 2. $\mathcal{P}$ is said to be the smoothest partition on U if for all $u \in U$, $[u]_{\mathcal{P}} = \{u\}$.

Definition 3. $\mathcal{Q}$ is said to be the coarsest partition on U if for all $u \in U$, $[u]_{\mathcal{Q}} = U$.

3.2. VPOFRS extended from the VPOFRS model

Based on the proposed fuzzy approximation space, we extend the VPRS model on the fuzzy approximation space based on expanding the approximation operations (VPOFRS) as follows.

Definition 4. Let $\mathcal{R}$ be a fuzzy equivalence relation, then the β membership (β-lower) of $x \in U$ with $A \subseteq U$ is determined by

$$\underline{\mathcal{R}}_{\beta} A(x) = \min\{\inf_{y \in U} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\}, \tag{12}$$

and, the β membership capabilities (β-upper) of $x \in U$ with $A \subseteq U$ is determined by

$$\overline{\mathcal{R}_{\beta}} A(x) = \max\{\sup_{y \in U} \min(\mathcal{R}(x,y), A(y)), \beta\}. \tag{13}$$

Then, the Then, VPOFRS is determined by the pair $(\underline{\mathcal{R}}_{\beta} A(x), \overline{\mathcal{R}_{\beta}} A(x))$.

Proposition 2. *Let $\mathcal{R}$ be a fuzzy equivalence relation, $A \subseteq U$, then for all $y \in U$*

$$\min\{\inf_{y \in U} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\} = \min\{\inf_{y \in A} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\}, \tag{14}$$

$$\max\{\sup_{y \in U} \min(\mathcal{R}(x,y), A(y)), \beta\} = \max\{\sup_{y \in A} \min(\mathcal{R}(x,y), A(y)), \beta\}. \tag{15}$$

Proof

(1) For formula 14: If $y \in A$ then $A(y) = 1$. So

$$\min\{\inf_{y\in U} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\} = \min\{\inf_{y\in A} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\} = \min\{1, \beta\}.$$

If $y \notin A$ then $A(y) = 0$. So

$$\min\{\inf_{y\in U} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\} = \min\{\inf_{y\in A} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\} = \min\{0, \beta\}.$$

(2) For formula 15: If $y \in A$ then $A(y) = 1$. So

$$\max\{\sup_{y\in U} \min(\mathcal{R}(x,y), A(y)), \beta\} = \max\{\sup_{y\in A} \min(\mathcal{R}(x,y), A(y)), \beta\} = \max\{\mathcal{R}(x,y), \beta\}.$$

If $y \notin A$ then $A(y) = 0$. So

$$\max\{\sup_{y\in U} \min(\mathcal{R}(x,y), A(y)), \beta\} = \max\{\sup_{y\in A} \min(\mathcal{R}(x,y), A(y)), \beta\} = \beta.$$

From (1) and (2), the proposition has been proven. ■

Based on Proposition 2, we redefine the approximation operations of VPOFRS as follows.

Definition 5. Let $\mathcal{R}$ be a fuzzy equivalence relation, then the β membership (β-lower) of $x \in U$ with $A \subseteq U$ is determined by

$$\underline{\mathcal{R}}_\beta A(x) = \min\{\inf_{y\in A} \mathcal{I}(\mathcal{R}(x,y), A(y)), \beta\}, \tag{16}$$

$$\overline{\mathcal{R}_\beta} A(x) = \max\{\sup_{y\in A} \min(\mathcal{R}(x,y), A(y)), \beta\}. \tag{17}$$

Then, the β-lower approximation of A according to $\mathcal{R}$ is a fuzzy set defined by

$$\underline{\mathcal{R}}_\beta A = \bigcup_{x\in A} \left\{\underline{\mathcal{R}}_\beta A(x)\right\}. \tag{18}$$

The β-upper approximation of A according to $\mathcal{R}$ is a fuzzy set defined by

$$\overline{\mathcal{R}_\beta} A = \bigcup_{x\in A} \left\{\overline{\mathcal{R}_\beta} A(x)\right\}. \tag{19}$$

Obviously, $\underline{\mathcal{R}}_\beta A$ is always less than or equal to $\overline{\mathcal{R}_\beta} A$. If $\underline{\mathcal{R}}_\beta A$ is equal to $\overline{\mathcal{R}_\beta} A$ we call it β exact set:

The β-boundary region of A is given by

$$BN_\beta A = \overline{\mathcal{R}_\beta} A - \underline{\mathcal{R}}_\beta A. \tag{20}$$

The β-negative region of A is given by

$$NEG_\beta A = U - \overline{\mathcal{R}_\beta} A. \tag{21}$$

The β lower approximation of the set A can be understood as the collection of all those elements of $F(U)$ which can be classified into A with a classification error not greater than β. Similarly, the β-negative region of A is the collection of all those elements of $F(U)$ which can be classified into the complement of A with the classification error not greater than β.

3.3. Some metrics on VPOFRS

In this section, we build several measures to evaluate the consistency of the decision table, and the dependency of attributes. First, we define the positive region POS of the decision table as follows.

Definition 6. Let decision table $DT = (U, C, D)$ with $D = \{D_1, D_2, ...D_n\}$ and $\mathcal{C}$ is the fuzzy partition of C on U. Then the positive region of D according to the VPOFRS approach is defined as follows

$$POS_\beta(C, D) = \bigcup \left\{ \underline{\mathcal{C}}_\beta X : X \in D \right\}. \tag{22}$$

Then, the decision table is said to be β-consistent if

$$POS_\beta(C, D) = U. \tag{23}$$

Definition 7. Let decision table $DT = (U, C, D)$, the β dependence of D on C is determined as follows

$$\gamma_\beta(C, D) = \frac{POS_\beta(C, D)}{|U|}. \tag{24}$$

Proposition 3. *Let decision table $DT = (U, C, D)$ and $P, Q \subseteq C$, then $\gamma_\beta(P, D) \leqslant \gamma_\beta(Q, D)$ if $P \subseteq Q$.*

Proof. Indeed, according to proposition 1, we have that if $P \subseteq Q$, then $\mathcal{Q} \subseteq \mathcal{P}$. Therefore, for every $x \in U$ and $A \subseteq U$, if $[x]_P$ belongs to A, then $[x]_Q$ also belongs to A, then $\underline{\mathcal{P}}_\beta A(x) \subseteq \underline{\mathcal{Q}}_\beta A(x)$ leads to $POS_\beta(P, A) \subseteq POS_\beta(Q, A)$. The proposition has been proven. ∎

4. ATTRIBUTE REDUCTION METHOD BASED ON VPOFRS

This section presents the attribute reduction method according to the VPOFRS. Before building the attribute reduction algorithm, we define the significance of the attribute and the criteria of the reduct according to VPOFRS.

4.1. The reduct and attribute selection method based on VPOFRS

Based on the anti-monotonic property of Proposition 3. We propose a measure to evaluate the significance of attributes as follows.

Definition 8. Given a decision table $DT = (U, C, D)$ and $B \subseteq C$. Then, attribute $c \in \{C - B\}$ is said to be β significant to B if

$$Big_\beta(c, B) = \gamma_\beta(B \cup \{c\}, D) - \gamma_\beta(B, D). \tag{25}$$

We can see that the formula (25) is always a non-negative value. On that basis, we define the reduct as follows.

Definition 9. Given a decision table $DT = (U, C, D)$, then $B \subseteq C$ is called a β reduct if the following conditions are satisfied

$$\gamma_\beta(D, B) = \gamma_\beta(D, C), \tag{26}$$

$\forall b \in B, \ \gamma_\beta(D, B - \{b\}) \neq \gamma_\beta(D, B)$.

4.2. Attribute reduction algorithm based on VPRS

Based on the definitions of attribute significance according to formula 25 and the reduct according to Definition 9, we assemble the attribute filter algorithm VPOFRS_AR to find the reduct. The main design method of the algorithm follows the Heuristic approach of selecting each best attribute in the decision table to add to the reduct. This is the traditional approach used in most attribute reduction algorithms following the RS approach. Algorithm details are presented in Algorithm 1 below.

Algorithm 1 Attribute filtering algorithm according to VPOFRS_AR approach

Input Decision table $DT = (U, C, D)$, $\beta \in [0, 1]$
Output The reduct B_β

1: $B = \emptyset$;
2: **while** $\gamma_\beta(D, B) \neq \gamma_\beta(D, C)$ **do**
3:    $SIG = 0$;
4:    $b = \emptyset$;
5:    **for all** $c \in \{C - B\}$ **do**
6:       **if** $Sig_\beta(c, B) > SIG$ **then**
7:          $SIG = Sig_\beta(c, B)$;
8:          $b = \{c\}$;
9:       **end if**
10:    **end for**
11:    $B = B \cup b$;
12: **end while**
13: **return** B;

Next, we proceed to evaluate Algorithm 1.

Let $|U|$ be the number of objects, $|C|$ be the number of condition attributes, and $|D|$ be the number of classes of the data table. Then based on Definition 1 and Definition 5 we have:

1) Steps 5-10 is $\mathcal{O}(|D||U||C|)$;

2) Steps 2-12 is $\mathcal{O}(|D||U||C|^2)$.

From 1) and 2) we have the complexity of the algorithm is $\mathcal{O}(|D||U||C|^2)$.

Most of the current attribute reduction algorithms based on the FRS and VPFRS have a complexity of a 3rd-degree polynomial. Based on Definition 4, we can see that the complexity of the VPFRS algorithm is $\mathcal{O}(|U|^2|C|^2)$. In reality, $|D|$ is always much smaller than $|U|$; therefore, $|D||U|$ is much smaller than $|U|^2$. Therefore, according to the VPOFRS approach, the proposed algorithm has an efficient calculation time.

5. SOME EXPERIMENTAL RESULTS

5.1. Experimental scenarios

To evaluate the effectiveness of the proposed algorithm VPOFRS_AR, we construct an experimental scenario as follows:

1) First, we evaluate the impact of the β values on the reduct obtained at each different hop. From there, determine the trend of the reduct in terms of size and classification accuracy and the appropriate β value range for each dataset.

2) Compare the proposed VPOFRS_AR with the VPFRS algorithm [26], in which the VPFRS algorithm improves noise for reducts and has a faster execution time than algorithms following the IFRS [27] approach.

5.2. Experimental environment

Table 2: Datasets description

ID	Data	Describe	$\|U\|$	$\|C\|$	$\|D\|$
1	UFDC	Ultrasonic flowmeter diagnostics (C)	181	43	4
2	UFDD	Ultrasonic flowmeter diagnostics (D)	181	43	4
3	SHDC	SPECTF heart data set	267	44	2
4	SONAR	Connectionist bench	208	60	2
5	VRB	Voice rehabilitation(Binary)	126	310	2
6	VRG	Voice rehabilitation(Gender)	126	310	2

Test datasets were downloaded from the UC Irvine machine learning repository (UCI) [30]. All experimental datasets have low classification accuracy ($< 70\%$), especially those with only 40% accuracy, such as UFDC. Details of the experimental datasets are in Table 2, where $|U|$ is the number of rows, $|C|$ is the number of condition attributes, and $|D|$ is the number of classes for each dataset. These datasets have real number attribute value domains normalized to the value domain [0, 1]. Because all datasets have a value range of [0, 1], so we use SVM and kNN ($k = |D|$) models to evaluate the reduct of the algorithms. The 10-fold cross-evaluation method combines with (recall, precision, f_1) scores to determine the average accuracy of the reducts. We use Python programming languages for all algorithms on the Anaconda 3.6 platform running on a Windows 10 operating system with an i5 Processor and 8GB of RAM. Before comparing with the VPFRS algorithm, we evaluate the reduct of the VPOFRS_AR algorithm at each β leave difference.

5.3. Evaluate the VPOFRS_AR algorithm

This section analyzes the proposed algorithm's performance on each dataset. We found that each dataset had a different β range and jump step. Therefore, we choose the most suitable β range to get the best result for each dataset.

Figures 1, 2 and 3 present the trends of the reducts obtained from the VPOFRS_AR algorithm at each β level. In particular, Figure 1 shows the trend in the size of the reducts; Figure 2 shows the trend in classification accuracy of the reducts with the SVM model; Figure 3 shows the trend in classification accuracy with the KNN model. Figure 2 shows that most reducts tend to decrease in size when increasing the β level, in which VRB and VRG datasets reduce the most attributes compared to other datasets. UFDC dataset has the opposite trend compared to other datasets, in which the reduct size increases when the β level is increased. Figure 1 shows that the classification accuracy of most reducts decreases when the β level increases. Combining with Figure 1, we can see that the number of attributes decreases, the accuracy also decreases. However, UFDC and SHDC do not decrease in accuracy. In

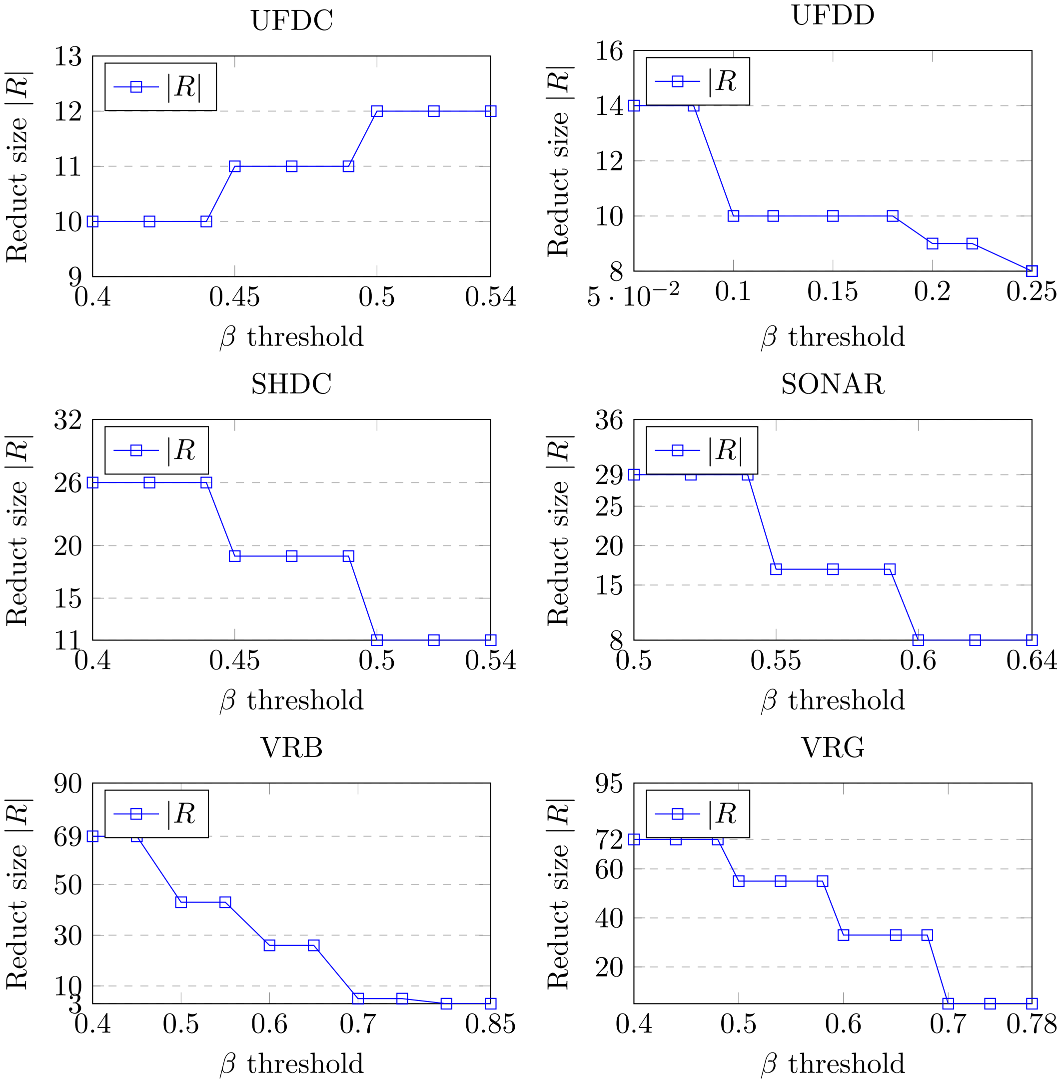

Figure 1: The reduct size trend when adjusting β threshold

particular, the UFDC data set improves classification accuracy by more than 10% on the SVM model. Figure 2 and 3 show that most reducts decrease classification accuracy when increasing the β level. Combining observations with Figure 1, we can see that as the number of attributes decreases, the accuracy also decreases on the KNN classification model.

5.4. Compare the VPOFRS_AR algorithm with the VPFRS algorithm

According to the FRS approach, the VPFRS model has been evaluated as the most effective anti-noise model among the VP methods. Therefore, we use the algorithm on the VPFRS model [26], referred to as Alg_2 for short, and compare it with the proposed model VPOFRS_AR, referred to as Alg_1.

Table 3 shows that both algorithms are good in terms of execution time. Furthermore, the size and classification accuracy of the result on both SVM and KNN models are equal on

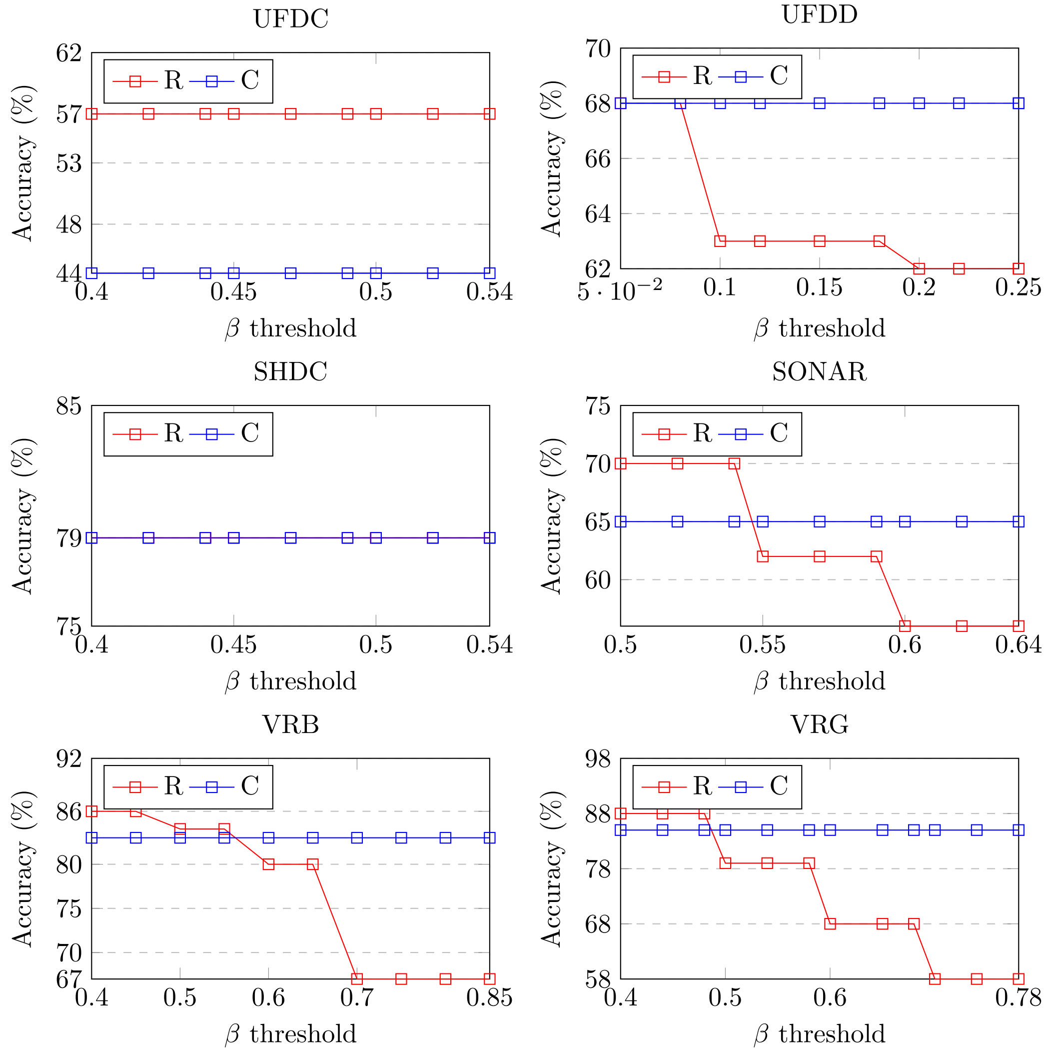

Figure 2: The accuracy classification trend when adjusting the threshold on the SVM model

Table 3: Comparison between VPFRS and VPOFRS_AR algorithms

ID	Datasets	$\|R\|$		Accuracy (%) SVM		Accuracy (%) KNN		Time (s)	
		Alg_1	Alg_2	Alg_1	Alg_2	Alg_1	Alg_2	Alg_1	Alg_2
1	UFDC	**12**	26	**0.57**	0.52	0.86	0.88	0.32	1.22
2	UFDD	14	20	0.68	0.68	0.78	0.68	0.54	1.7
3	SHDC	19	15	0.79	0.8	0.71	0.79	0.95	1.5
4	SONAR	29	20	0.7	0.65	0.7	0.65	1.73	2.3
5	VRB	69	45	0.86	0.88	0.71	0.72	3.65	5.2
6	VRG	72	83	0.88	0.82	0.74	0.7	5.8	8.8
	$\| - \|$	35.83	34.83	0.75	0.73	0.75	0.74	**2.17**	3.45

average $| - |$. However, looking at the UFDC dataset, we can see that the reduct obtained from the Alg_1 algorithm can improve noise by 5% and the size of the reduct is only half that

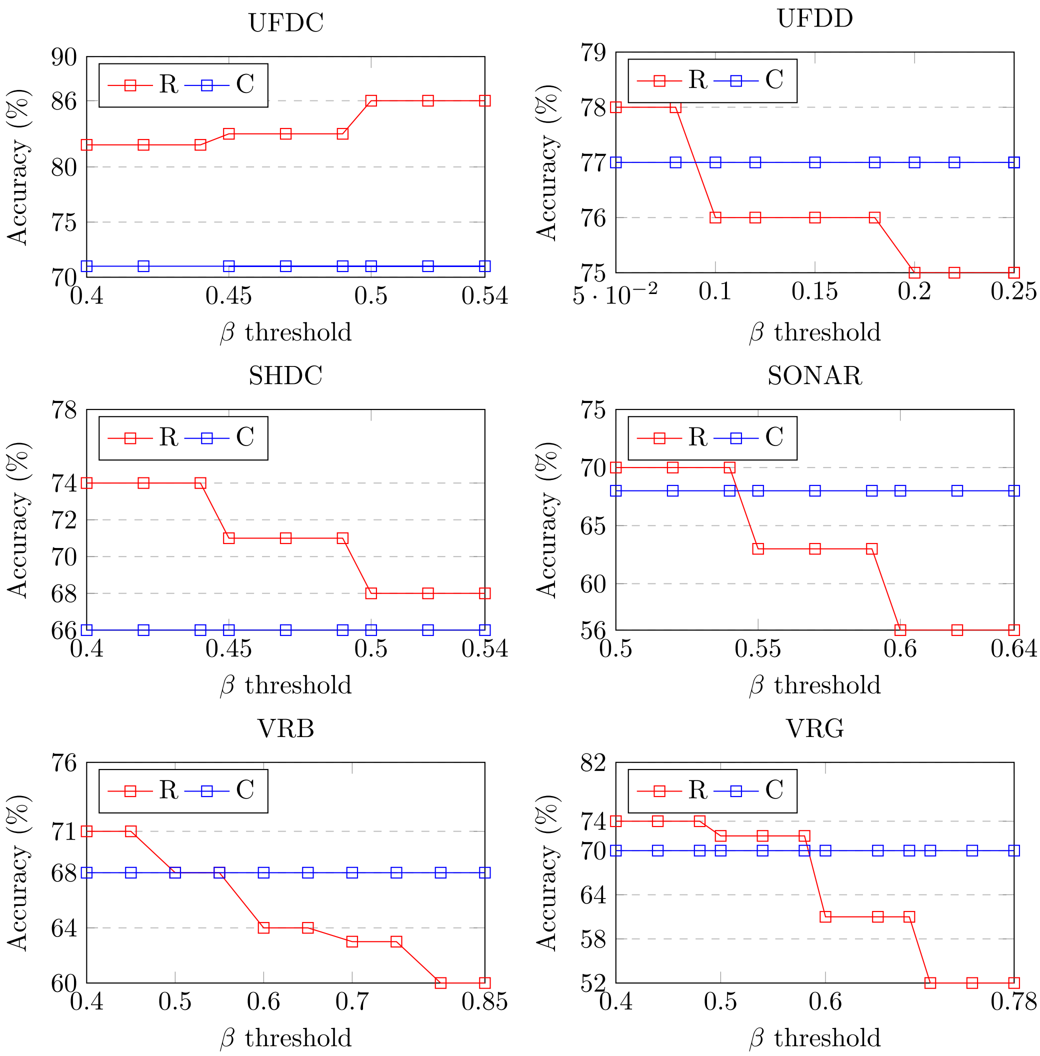

Figure 3: The accuracy classification trend when adjusting the threshold on the KNN model

of the Alg_2 algorithm. The average execution time of Alg_1 is only 2.17 seconds compared to the Alg_2 algorithm, which is 3.45 seconds. That confirms that the proposed VPFRS approach algorithm has efficient calculation time.

6. CONCLUSIONS

In this paper, we have improved the VP method for approximation operations of the FRS model and constructed the VPOFRS model. Based on the VPOFRS model, we propose several measures to evaluate the decision table's consistency and the attributes' dependency. These measures are used to build the VPOFRS_AR attribute reduction algorithm, applied to noisy datasets. Experimental and comparative results show that the proposed method is efficient in computation time while the classification accuracy and the size of the reduct are similar to the VPFRS algorithm. In the future, we will extend the VPOFRS model to construct incremental computation methods for variable decision tables.

ACKNOWLEDGMENT

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2021.10.

REFERENCES

[1] S. Bashir, I. U. Khattak, A. Khan, F. H. Khan, A. Gani, and M. Shiraz, "A novel feature selection method for classification of medical data using filters, wrappers, and embedded approaches," *Complexity*, vol. 2022, pp. 1–12, 2022.

[2] L. Meenachi and S. Ramakrishnan, "Differential evolution and ACO based global optimal feature selection with fuzzy rough set for cancer data classification," *Soft Computing*, vol. 24, no. 24, pp. 18 463–18 475, 2020.

[3] S. Ahlawat and R. Rishi, "A genetic algorithm based feature selection for handwritten digit recognition," *Recent Patents on Computer Science*, vol. 12, no. 4, pp. 304–316, 2018.

[4] H. huang Zhao, H. Liu, H. huang Zhao, and H. Liu, "Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition," *Granular Computing*, vol. 5, no. 3, pp. 411–418, 2020.

[5] L. Sun, S. Fu, and F. Wang, "Decision tree SVM model with Fisher feature selection for speech emotion recognition," *Eurasip Journal on Audio, Speech, and Music Processing*, vol. 2019, no. 1, 2019.

[6] S. Yildirim, Y. Kaya, and F. Kılıç, "A modified feature selection method based on metaheuristic algorithms for speech emotion recognition," *Applied Acoustics*, vol. 173, 2021.

[7] G. Ansari, T. Ahmad, and M. N. Doja, "Spam review classification using ensemble of global and local feature selectors," *Cybernetics and Information Technologies*, vol. 18, no. 4, pp. 29–42, 2018.

[8] H. Mohammadzadeh and F. S. Gharehchopogh, "A novel hybrid whale optimization algorithm with flower pollination algorithm for feature selection: Case study Email spam detection," *Computational Intelligence*, vol. 37, no. 1, pp. 176–209, 2021.

[9] A. J. Fernández-García, L. Iribarne, A. Corral, J. Criado, and J. Z. Wang, "A recommender system for component-based applications using machine learning techniques," *Knowledge-Based Systems*, vol. 164, pp. 68–84, 2019.

[10] B. Saravanan, V. Mohanraj, and J. Senthilkumar, "A fuzzy entropy technique for dimensionality reduction in recommender systems using deep learning," *Soft Computing*, vol. 23, no. 8, pp. 2575–2583, 2019.

[11] N. L. Giang, L. H. Son, T. T. Ngan, T. M. Tuan, H. T. Phuong, M. Abdel-Basset, A. R. L. De Macedo, and V. H. C. De Albuquerque, "Novel incremental algorithms for attribute reduction from dynamic decision tables using hybrid filter-wrapper with fuzzy partition distance," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 5, pp. 858–873, 2020.

[12] T. T. Nguyen, N. L. Giang, D. T. Tran, T. T. Nguyen, H. Q. Nguyen, A. V. Pham, and T. D. Vu, “A novel filter-wrapper algorithm on intuitionistic fuzzy set for attribute reduction from decision tables,” *International Journal of Data Warehousing and Mining*, vol. 17, no. 4, pp. 67–100, 2021.

[13] Y. Li, M. Cai, J. Zhou, and Q. Li, “Accelerated multi-granularity reduction based on neighborhood rough sets,” *Applied Intelligence*, vol. 52, no. 15, pp. 17 636–17 651, 2022.

[14] X. Yang, H. Chen, T. Li, J. Wan, and B. Sang, “Neighborhood rough sets with distance metric learning for feature selection,” *Knowledge-Based Systems*, vol. 224, no. 107076, p. 107076, 2021.

[15] J. Liu, Y. Lin, J. Du, H. Zhang, Z. Chen, and J. Zhang, “ASFS: A novel streaming feature selection for multi-label data based on neighborhood rough set,” *Applied Intelligence*, 2022.

[16] Z. Pawlak, “Rough sets,” *International Journal of Computer & Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.

[17] D. Liu and J. Li, “Safety monitoring data classification method based on wireless rough network of neighborhood rough sets,” *Safety Science*, vol. 118, pp. 282–296, 2019.

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

[19] J. Ye, J. Zhan, W. Ding, and H. Fujita, “A novel fuzzy rough set model with fuzzy neighborhood operators,” *Information Sciences*, vol. 544, pp. 266–297, 2021.

[20] P. Zhang, T. Li, G. Wang, C. Luo, H. Chen, J. Zhang, D. Wang, and Z. Yu, “Multi-source information fusion based on rough set theory: A review,” pp. 100–107, 2021.

[21] P. Jain, A. K. Tiwari, and T. Som, “A fitting model based intuitionistic fuzzy rough feature selection,” *Engineering Applications of Artificial Intelligence*, vol. 89, no. 103421, p. 103421, 2020.

[22] B. Liang, L. Wang, and Y. Liu, “Attribute reduction based on improved information entropy,” *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 1, pp. 709–718, 2019.

[23] Z. Qiu and H. Zhao, “A fuzzy rough set approach to hierarchical feature selection based on Hausdorff distance,” *Applied Intelligence*, vol. 52, no. 10, pp. 11 089–11 102, 2022.

[24] D. Dubois and H. Prade, “Rough fuzzy sets and fuzzy rough sets,” *International Journal of General Systems*, vol. 17, no. 2-3, pp. 191–209, 1990.

[25] Z. Zhang, “Attributes reduction based on intuitionistic fuzzy rough sets,” *Journal of Intelligent and Fuzzy Systems*, vol. 30, no. 2, pp. 1127–1137, 2016.

[26] S. Zhao, E. C. Tsang, and D. Chen, “The model of fuzzy variable precision rough sets,” *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 2, 2009.

[27] A. Tan, W. Z. Wu, Y. Qian, J. Liang, J. Chen, and J. Li, “Intuitionistic fuzzy rough set-based granular structures and attribute subset selection,” *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 3, pp. 527–539, 2019.

[28] A. Tan, S. Shi, W. Z. Wu, J. Li, and W. Pedrycz, “Granularity and entropy of intuitionistic fuzzy information and their applications,” *IEEE Transactions on Cybernetics*, vol. 52, no. 1, pp. 192–204, 2022.

[29] Y. Chen and Y. Chen, "Feature subset selection based on variable precision neighborhood rough sets," *International Journal of Computational Intelligence Systems*, vol. 14, no. 1, 2021.

[30] UCI, "UC Irvine machine learning repository," 2023, accessed on 26 March 2023. [Online]. Available: http://https://archive.ics.uci.edu/

Received on December 19, 2023
Revised on February 28, 2024