

About a Fuzzy Distance between Two Fuzzy Partitions and Application in Attribute Reduction Problem

Cao Chinh Nghia¹, Demetrovics Janos², Nguyen Long Giang³,
Vu Duc Thi⁴

¹People's Police Academy, Viet Nam

²Institute for Computer and Control (SZTAKI) Hungarian Academy of Sciences

³Institute of Information Technology, VAST, Viet Nam

⁴Institute of Information Technology, VNU, Viet Nam

Emails: ccnghia@gmail.com demetrovics@sztaki.mta.hu nlgiang@ioit.ac.vn vdthi@vnu.edu.vn

Abstract: According to traditional rough set theory approach, attribute reduction methods are performed on the decision tables with the discretized value domain, which are decision tables obtained by discretized data methods. In recent years, researches have proposed methods based on fuzzy rough set approach to solve the problem of attribute reduction in decision tables with numerical value domain. In this paper, we propose a fuzzy distance between two partitions and an attribute reduction method in numerical decision tables based on proposed fuzzy distance. Experiments on data sets show that the classification accuracy of proposed method is more efficient than the ones based fuzzy entropy.

Keywords: Fuzzy rough set, fuzzy equivalence relation, fuzzy distance, decision table, attribute reduction, reduct.

1. Introduction

Attribute reduction is an important issue of data preprocessing steps in data mining and knowledge discovery. The aim of attribute reduction is eliminated redundant attributes to enhance the effectiveness of data mining algorithms. Rough set theory of Pawlak is an effective tool to solve the attribute reduction problem in decision tables and the one that researchers has performed for a long time. Rough set based attribute reduction methods are performed on decision tables with discretized value domain [2-4]. In fact, the attribute value domain of decision tables often contains numerical values and continuous values. For example, the attribute of body weight and blood pressure in patient data tables is usually numerical value, continuous value. When performing attribute reduction methods based on rough set, data needs to be discretized. However, these discretized methods do not preserve the initial difference

between objects in the original data, so it reduces the classification accuracy after attribute reduction. To solve the issue of attribute reduction in decision tables with numerical value and continuous value, in recent years, researches have proposed new methods based on fuzzy rough set approach.

Dubois and Prade [1] proposed fuzzy rough set theory is a combination of rough set theory and fuzzy set theory in order to approximate fuzzy sets based on fuzzy equivalence relation. The fuzzy equivalence is determined by the attribute value domain. Traditional rough set based on similarity relation to approximate sets. In rough set theory, two objects are called equivalent on the attribute set (the similarity is 1) if their attribute values are equal on all attributes. Conversely, they are not equal (the similarity is 0). The fuzzy rough set theory has used the fuzzy equivalence relation to replace the equivalence relation. The value similarity in the range $[0, 1]$ shows the close or similar properties of two objects. Therefore, the fuzzy equivalence preserves the difference or the similarity of objects. Attribute reduction methods based on fuzzy rough set approach has the potential to preserve the classification accuracy after implementing attribute reduction methods.

In recent years, the topic of the attribute reduction based on fuzzy rough set has attracted many researchers [5-14]. With attribute reduction issue directly on the decision table based on fuzzy rough set, these related researches are concerning in two main directions: fuzzy positive region approach and fuzzy entropy approach. Based on fuzzy positive region, Hu, Xie and Yu [8] proposed FAR-VPFRS algorithm to find one fuzzy positive region reduct which use the fuzzy membership function. The experimental data sets show that the classification accuracy of FAR-VPFRS algorithm is better than the one of algorithm which use the membership function according to traditional rough set. Qian et al. [14] proposed FA_FPR algorithm, which is an improvement of FAR-VPFRS [8] in terms of executed time. According to fuzzy entropy approach, Hu, Yu and Xie [7] proposed the fuzzy entropy which is based on entropy Shannon and introduces FSCE to find one reduct using fuzzy entropy. Dai and Xu [6] proposed fuzzy gain ratio based on fuzzy entropy and introduces GAIN_RATIO_AS_FRS to find one reduct using fuzzy gain ration. The experimental data sets show that the classification of FSCE, GAIN_RATIO_AS_FRS algorithms are better than the ones based on traditional rough set. Qian et al. [14] who proposed FA_FSCE algorithm, is an improvement of FSCE algorithm [7] in terms of executed time. In both direction approaches, authors in [14] have never evaluated the classification accuracy after implementing improved algorithms FA_FPR, FA_FSCE. With direct reduction attribute on the decision table based on fuzzy rough set, the aim of this paper is to propose a new method which improves the classification accuracy more than the previous ones.

In this paper, we propose the heuristic algorithm to find a best reduction of the decision table with numerical attribute value domain using a fuzzy distance. The fuzzy distance is constructed between two partitions. The experimental results in data sets from UCI [17] show that the classification accuracy of proposed algorithm is better than FA_FSCE and FA_FPR algorithms [14]. The structure of this paper is as follows. Section 2 presents some basic concepts of fuzzy rough set theory. Section 3 presents the method of constructing the fuzzy distance between two attribute sets.

Section 4 presents an attribute reduction method using fuzzy distance measure. Section 5 presents experimental results. Finally, Section 6 gives a conclusion of this paper and subsequent developments.

2. Basic concepts

In this section, we introduce some concepts in rough set theory, fuzzy rough set theory and some related concepts in fuzzy partition space.

A decision table is a pair $DS = (U, C \cup D)$, where U be a non-empty finite set; C is called conditional attribute set, D is called decision attribute set with $C \cap D \neq \emptyset$. DS is called the numerical decision table where the value domain of $c \in C$ is the numerical for any $c \in C$.

Pawlak's traditional rough set theory [15] used an equivalence relation to approximate sets. A subset $P \subseteq C$ determines equivalence relation on attribute value domain, denote by $IND(P)$,

$$IND(P) = \{(u, v) \in U \times U \mid \forall a \in P, a(u) = a(v)\},$$

$a(v)$ is denoted as the value attribute a in object v ; $IND(P)$ determines the partition on U , denoted by $U / IND(P)$ and the equivalence class of u , denoted by $[u]_P$. The lower approximation set and the upper approximate set of $X \subseteq U$ related to $P \subseteq C$ is defined as $\underline{P}X = \{u \in U \mid [u]_P \subseteq X\}$ and $\overline{P}X = \{u \in U \mid [u]_P \cap X \neq \emptyset\}$.

D. Dubois and others proposed the fuzzy rough set which used fuzzy equivalent to approximate the fuzzy sets. The decision table with numerical attribute domain $DS = (U, C \cup D)$, the relation R defined on U is called fuzzy similarity relation if it satisfies the following conditions [14]:

- 1) Reflectivity: $R(x, x) = 1$.
- 2) Symmetry: $R(x, y) = R(y, x)$.
- 3) Max-min transitivity: $R(x, z) \geq \min\{R(x, y), R(y, z)\}$, for any $x, y, z \in U$.

Let U be a non-empty finite set and R_P và R_Q be a fuzzy equivalence relation on U if for any $x, y \in U$, we have:

- 1) $R_P = R_Q \Leftrightarrow R_P(x, y) = R_Q(x, y)$,
- 2) $R = R_P \cup R_Q \Leftrightarrow R(x, y) = \max\{R_P(x, y), R_Q(x, y)\}$,
- 3) $R = R_P \cap R_Q \Leftrightarrow R(x, y) = \min\{R_P(x, y), R_Q(x, y)\}$,
- 4) $R_P \subseteq R_Q \Leftrightarrow R_P(x, y) \leq R_Q(x, y)$.

The relation matrix of R_P denoted by $M(R_P) = [p_{ij}]_{n \times n}$ is defined as

$$M(R_P) = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix},$$

where $p_{ij} = R_P(x_i, x_j)$ is the fuzzy relation value of x_i and x_j on P , $p_{ij} \in [0, 1]$.

Let $DS = (U, C \cup D)$ be a decision table with numerical attributes and $P, Q \subseteq C$. According to [11] we have $R_P = \bigcap_{a \in P} R_a$ and $R_{P \cup Q} = R_P \cap R_Q$, it means that $R_{P \cup Q}(x, y) = \min\{R_P(x, y), R_Q(x, y)\}$ for any $x, y \in U$. Suppose that $M(R_P) = [p_{ij}]_{n \times n}$ and $M(R_Q) = [q_{ij}]_{n \times n}$ are relational matrices of R_P, R_Q corresponding, then the relational matrix on the attribute sets $S = P \cup Q$ is defined as

$$M(R_S) = M(R_{P \cup Q}) = [s_{ij}]_{n \times n} \text{ with } s_{ij} = \min\{p_{ij}, q_{ij}\}.$$

For $P \subseteq C$, $U = \{x_1, \dots, x_n\}$, the fuzzy partition $\pi(P) = U / R_P$ on U can be generated from the fuzzy equivalence relation R_P :

$$\pi(R_P) = U / R_P = \left\{ [x_i]_{R_P} \right\}_{i=1}^n = \left\{ [x_1]_{R_P}, \dots, [x_n]_{R_P} \right\},$$

where $[x_i]_{R_P} = p_{i1}/x_1 + p_{i2}/x_2 + \dots + p_{in}/x_n$ is a fuzzy set, is called a fuzzy equivalence of object x_i . The membership function of objects is determined by $\mu_{[x_i]_{R_P}}(x_j) = \mu_{R_P}(x_i, x_j) = R_P(x_i, x_j) = p_{ij}$ for any $x_j \in U$. Then, the cardinality of fuzzy equivalence $[x_i]_{R_P}$ is calculated [11] as

$$|[x_i]_{R_P}| = \sum_{j=1}^n p_{ij}.$$

Let \mathcal{P} is called a set of all of fuzzy partitions on U which determined by fuzzy equivalence on attribute sets. Then \mathcal{P} is called a *fuzzy partition space on U* . Thus, the fuzzy partition space is determined by fuzzy equivalence relation which chose from the attribute value domain. Let $\pi(R_P) = \{[x_1]_{R_P}, \dots, [x_n]_{R_P}\}$ be a fuzzy partition where $[x_i]_{R_P} = p_{i1}/x_1 + \dots + p_{in}/x_n$. Specially, if $p_{ij} = 0$, $i, j \leq n$, then $[x_i]_{R_P} = 0$ and the fuzzy partition $\pi(R_P)$ is called the finest one, write as $\pi(\omega)$. Then, $\pi(\omega) = \{[x_1]_{\omega}, \dots, [x_n]_{\omega}\}$ where $[x_i]_{\omega} = \sum_{j=1}^n \omega_{ij}/x_j$, $i, j \leq n$, $\omega_{ij} = 0$. If

$p_{ij} = 1, i, j \leq n$, then $|[x_i]_{R_P}| = |U|, i \leq n$, and the fuzzy partition $\pi(R_P)$ is called the coarsest one, write as $\pi(\delta)$. Then, $\pi(\delta) = ([x_1]_\delta, \dots, [x_n]_\delta)$ where $[x_i]_\delta = \sum_{j=1}^n \delta_{ij} / x_j, i, j \leq n, \delta_{ij} = 1$.

\mathcal{P} is a fuzzy partition space on U , for $\pi(R_P), \pi(R_Q) \in \mathcal{P}$, the partial order relation

$$\preceq: \pi(R_P) \preceq \pi(R_Q) \Leftrightarrow [x_i]_{R_P} \subseteq [x_i]_{R_Q}, i \leq n \Leftrightarrow p_{ij} \leq q_{ij}, i, j \leq n,$$

denoted by $R_P \preceq R_Q$. Furthermore,

$$\pi(R_P) = \pi(R_Q) \Leftrightarrow [x_i]_{R_P} = [x_i]_{R_Q}, i \leq n \Leftrightarrow p_{ij} = q_{ij}, i, j \leq n,$$

denoted by $R_P = R_Q$;

$$\pi(R_P) \prec \pi(R_Q) \Leftrightarrow \pi(R_P) \preceq \pi(R_Q) \text{ and } \pi(R_P) \neq \pi(R_Q),$$

denoted by $R_P \prec R_Q$.

Example 1. Let

$$\begin{aligned} U &= \{x_1, x_2\}, \pi(R_P) = ([x_1]_{R_P}, [x_2]_{R_P}), \pi(R_Q) = ([x_1]_{R_Q}, [x_2]_{R_Q}), \\ \pi(R_S) &= ([x_1]_{R_S}, [x_2]_{R_S}), \text{ where } [x_1]_{R_P} = 0.1/x_1 + 0.2/x_2, \\ [x_2]_{R_P} &= 0.2/x_1 + 0.3/x_2, [x_1]_{R_Q} = 0.2/x_1 + 0.3/x_2, [x_2]_{R_Q} = 0.3/x_1 + 0.4/x_2, \\ [x_1]_{R_S} &= 0.3/x_1 + 0.4/x_2, [x_2]_{R_S} = 0.4/x_1 + 0.6/x_2. \end{aligned}$$

Then we have

$$\begin{aligned} |[x_1]_{R_P}| &= 0.1 + 0.2 = 0.3, |[x_2]_{R_P}| = 0.2 + 0.3 = 0.5, |[x_1]_{R_Q}| = 0.2 + 0.3 = 0.5, \\ |[x_2]_{R_Q}| &= 0.3 + 0.4 = 0.7, |[x_1]_{R_S}| = 0.3 + 0.4 = 0.7, |[x_2]_{R_S}| = 0.4 + 0.6 = 1, \\ |[x_1]_{R_P} \cap [x_1]_{R_Q}| &= 0.3, |[x_2]_{R_P} \cap [x_2]_{R_Q}| = 0.5, |[x_1]_{R_Q} \cap [x_1]_{R_S}| = 0.5, \\ |[x_2]_{R_Q} \cap [x_2]_{R_S}| &= 0.7, |[x_1]_{R_P} \cap [x_1]_{R_S}| = 0.3, |[x_2]_{R_P} \cap [x_2]_{R_S}| = 0.5. \end{aligned}$$

3. Fuzzy distance between two fuzzy partitions and its properties

3.1. Fuzzy distance between two fuzzy sets

Firstly, we have proposed a distance measure between two fuzzy sets, called a fuzzy distance.

Lemma 1. Let a, b, m be three real numbers with $a \geq b$. Then, we have $a - b \geq \min(a, m) - \min(b, m)$.

Proof: It is easy to see that $a - b \geq \min(a, m) - \min(b, m)$ satisfies: $m \geq a, b \leq m < a, m < b$. This completes the proof.

Lemma 2. Let A, B, C be three fuzzy sets on the same universe U . Then, we have:

$$1) \text{ If } A \subseteq B \text{ then } |B| - |B \cap C| \geq |A| - |A \cap C|.$$

$$2) \text{ If } A \subseteq B \text{ then } |C| - |C \cap A| \geq |C| - |C \cap B|.$$

$$3) |A| - |A \cap B| + |C| - |C \cap A| \geq |C| - |C \cap B|.$$

Proof:

1) Because of $A \subseteq B$, for any $x_i \in U$ we have $\mu_B(x_i) \geq \mu_A(x_i)$. According to Lemma 1, we have

$$\begin{aligned} \mu_B(x_i) - \mu_A(x_i) &\geq \min(\mu_B(x_i), \mu_C(x_i)) - \min(\mu_A(x_i), \mu_C(x_i)) \Leftrightarrow \\ &\Leftrightarrow \sum_{i=1}^{|U|} \mu_B(x_i) - \sum_{i=1}^{|U|} \mu_A(x_i) \geq \sum_{i=1}^{|U|} \min(\mu_B(x_i), \mu_C(x_i)) - \\ &\quad - \sum_{i=1}^{|U|} \min(\mu_A(x_i), \mu_C(x_i)) \Leftrightarrow \\ &\Leftrightarrow |B| - |A| \geq |B \cap C| - |A \cap C| \Leftrightarrow |B| - |B \cap C| \geq |A| - |A \cap C|. \end{aligned}$$

2) Because of $A \subseteq B$, for any $x_i \in U$, we have

$$\begin{aligned} \mu_B(x_i) \geq \mu_A(x_i) &\Leftrightarrow \min(\mu_B(x_i), \mu_C(x_i)) \geq \min(\mu_A(x_i), \mu_C(x_i)) \Leftrightarrow \\ &\Leftrightarrow \mu_C(x_i) - \min(\mu_A(x_i), \mu_C(x_i)) \geq \mu_C(x_i) - \min(\mu_B(x_i), \mu_C(x_i)) \Leftrightarrow \\ &\Leftrightarrow \sum_{i=1}^{|U|} \mu_C(x_i) - \sum_{i=1}^{|U|} \min(\mu_A(x_i), \mu_C(x_i)) \geq \\ &\geq \sum_{i=1}^{|U|} \mu_C(x_i) - \sum_{i=1}^{|U|} \min(\mu_B(x_i), \mu_C(x_i)) \Leftrightarrow |C| - |C \cap A| \geq |C| - |C \cap B|. \end{aligned}$$

3) From $A \cap C \subseteq A$, according to property 1) we have

$$(*) \quad |A| - |A \cap B| \geq |A \cap C| - |A \cap C \cap B|.$$

Furthermore, from $A \cap B \subseteq A$, according to property 2) we have

$$(**) \quad |C| - |C \cap A \cap B| \geq |C| - |C \cap A|.$$

From (*) and (**) we have

$$\begin{aligned} |A| - |A \cap B| + |C| - |C \cap A| &\geq |A \cap C| - |A \cap C \cap B| + |C| - |C \cap A| = \\ &= |C| - |A \cap B \cap C| \geq |C| - |C \cap A|. \end{aligned}$$

Proposition 1. Let A, B be two fuzzy sets on the same universe U . Then, $d(A, B) = |A| + |B| - 2|A \cap B|$ is a distance measure between A and B .

Proof: Obviously, from $|A| \geq |A \cap B|$ and $|B| \geq |A \cap B|$ to $d(A, B) \geq 0$. Furthermore, $d(A, B) = d(B, A)$. We need to prove the triangle inequality; without loss of generality, one needs to prove $d(A, B) + d(A, C) \geq d(B, C)$. According to Lemma 2 (Part 3), we have:

$$(***) \quad |A| - |A \cap B| + |C| - |C \cap A| \geq |C| - |C \cap B|,$$

$$(****) \quad |A| - |A \cap C| + |B| - |B \cap A| \geq |B| - |B \cap C|.$$

It is inferred from (***) and (****), we have

$$\begin{aligned} (|A| + |B| - 2|A \cap B|) + (|A| + |C| - 2|A \cap C|) &\geq |B| + |C| - 2|B \cap C|, \text{ similarly} \\ d(A, B) + d(A, C) &\geq d(B, C). \end{aligned}$$

Therefore, $d(A, B)$ is a distance measure between fuzzy set A and fuzzy set B , called fuzzy distance. We have proposed a distance between two fuzzy partitions based on fuzzy distance.

3.2. Fuzzy distance between two fuzzy partitions and its properties

Theorem 1. Let $DS = (U, C \cup D)$ be a decision table, where $U = \{x_1, x_2, \dots, x_n\}$ and $\pi(R_P), \pi(R_Q)$ be two fuzzy partitions induced by two fuzzy equivalence R_P, R_Q on $P, Q \subseteq C$. Then

$$(1) \quad D(\pi(R_P), \pi(R_Q)) = \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_P}| + |[x_i]_{R_Q}| - 2|[x_i]_{R_P} \cap [x_i]_{R_Q}|}{n} \right)$$

is a fuzzy distance between $\pi(R_P)$, and $\pi(R_Q)$.

Proof: Obviously, $D(\pi(R_P), \pi(R_Q)) \geq 0$ and

$$D(\pi(R_P), \pi(R_Q)) = D(\pi(R_Q), \pi(R_P)).$$

We need to prove the triangle inequality, without loss of generality, for any $\pi(R_P), \pi(R_Q), \pi(R_S) \in \mathcal{P}$, and we prove

$$D(\pi(R_P), \pi(R_Q)) + D(\pi(R_P), \pi(R_S)) \geq D(\pi(R_Q), \pi(R_S)).$$

It is inferred from Proposition 1, for any $x_i \in U$ we have

$$d\left([x_i]_{R_P}, [x_i]_{R_Q}\right) + d\left([x_i]_{R_P}, [x_i]_{R_S}\right) \geq d\left([x_i]_{R_Q}, [x_i]_{R_S}\right).$$

Then

$$\begin{aligned} & D\left(\pi\left(R_P\right), \pi\left(R_Q\right)\right) + D\left(\pi\left(R_P\right), \pi\left(R_S\right)\right) = \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\left| [x_i]_{R_P} \right| + \left| [x_i]_{R_Q} \right| - 2 \left| [x_i]_{R_P} \cap [x_i]_{R_Q} \right|}{n} \right) + \\ &+ \frac{1}{n} \sum_{i=1}^n \left(\frac{\left| [x_i]_{R_P} \right| + \left| [x_i]_{R_S} \right| - 2 \left| [x_i]_{R_P} \cap [x_i]_{R_S} \right|}{n} \right) = \\ &= \frac{1}{n} \sum_{i=1}^n \frac{d\left([x_i]_{R_P}, [x_i]_{R_Q}\right)}{n} + \frac{1}{n} \sum_{i=1}^n \frac{d\left([x_i]_{R_P}, [x_i]_{R_S}\right)}{n} \geq \\ &\geq \frac{1}{n} \sum_{i=1}^n \frac{d\left([x_i]_{R_Q}, [x_i]_{R_S}\right)}{n} = D\left(\pi\left(R_Q\right), \pi\left(R_S\right)\right). \end{aligned}$$

It is easy to see that

$$D\left(\pi\left(R_P\right), \pi\left(R_Q\right)\right) = 0 \Leftrightarrow \pi\left(R_P\right) = \pi\left(R_Q\right);$$

$D\left(\pi\left(R_P\right), \pi\left(R_Q\right)\right) = 1 \Leftrightarrow \pi\left(R_P\right) = \pi(\omega)$ and $\pi\left(R_Q\right) = \pi(\delta)$, (or $\pi\left(R_P\right) = \pi(\delta)$, and $\pi\left(R_Q\right) = \pi(\omega)$). Therefore, $0 \leq D\left(\pi\left(R_P\right), \pi\left(R_Q\right)\right) \leq 1$.

Proposition 2. Let $\pi\left(R_P\right) \in \mathcal{P}$ be a fuzzy partition on \mathcal{P} . Then, we have

$$D\left(\pi\left(R_P\right), \pi(\delta)\right) + D\left(\pi\left(R_P\right), \pi(\omega)\right) = 1.$$

Proof: Suppose that $\pi\left(R_P\right) = \left\{ [x_1]_{R_P}, [x_2]_{R_P}, \dots, [x_n]_{R_P} \right\}$. Then

$$D\left(\pi\left(R_P\right), \pi(\omega)\right) = \frac{1}{n^2} \sum_{i=1}^n \left| [x_i]_{R_P} \right|, \quad D\left(\pi\left(R_P\right), \pi(\delta)\right) = \frac{1}{n^2} \sum_{i=1}^n \left(n - \left| [x_i]_{R_P} \right| \right).$$

we have $D\left(\pi\left(R_P\right), \pi(\delta)\right) + D\left(\pi\left(R_P\right), \pi(\omega)\right) = 1$.

Example 2. Continuing from Example 1. According to Theorem 1, we have

$$D\left(\pi\left(R_P\right), \pi\left(R_Q\right)\right) = 0.1, \quad D\left(\pi\left(R_Q\right), \pi\left(R_S\right)\right) = 0.125,$$

$$D\left(\pi\left(R_P\right), \pi\left(R_S\right)\right) = 0.225.$$

Therefore:

$$\begin{aligned} D(\pi(R_P), \pi(R_Q)) + D(\pi(R_Q), \pi(R_S)) &= D(\pi(R_P), \pi(R_S)), \\ D(\pi(R_P), \pi(R_Q)) + D(\pi(R_P), \pi(R_S)) &> D(\pi(R_Q), \pi(R_S)), \\ D(\pi(R_Q), \pi(R_S)) + D(\pi(R_P), \pi(R_S)) &> D(\pi(R_P), \pi(R_Q)). \end{aligned}$$

4. A fuzzy distance based attribute reduction method in decision tables with numerical attributes

In this section, we introduce a fuzzy distance based attribute reduction method which performs directly on the decision tables with numerical attributes. The new fuzzy distance is defined between two fuzzy partitions (see Section 3).

Let $DS = (U, C \cup D)$ be a decision table with numerical attributes, $U = \{x_1, x_2, \dots, x_n\}$. We use a fuzzy equivalence relation defined on conditional attributes. For any $p \in C$, the following fuzzy equivalence relation R_p is often used to construct relational matrix $M(R_p) = [p_{ij}]_{n \times n}$ [6]

$$(2) \quad p_{ij} = \begin{cases} 1 - 4 \cdot \frac{|p(x_i) - p(x_j)|}{|p_{\max} - p_{\min}|}, & \frac{|p(x_i) - p(x_j)|}{|p_{\max} - p_{\min}|} \leq 0.25, \\ 0 & \text{otherwise,} \end{cases}$$

where $p(x_i)$ is the value of the attribute p in object x_i ; p_{\max}, p_{\min} are maximum value, minimum value of the attribute p , corresponding.

We use an equivalence relation $\text{IND}(D)$ and an equivalence matrix $M(\text{IND}(D)) = [d_{ij}]_{n \times n}$ on the decision attribute set, $d_{ij} = 1$ if $x_j \in [x_i]_D$ and $d_{ij} = 0$ if $x_j \notin [x_i]_D$. In other words, an equivalence class $[x_i]_D$ can be seen as a fuzzy equivalence class, denoted by $[x_i]_D$, the membership function $\mu_{[x_i]_D}(x_j) = 1$ if

$$x_j \in [x_i]_D \text{ and } \mu_{[x_i]_D}(x_j) = 0 \text{ if } x_j \notin [x_i]_D.$$

Then, the fuzzy partition denoted by

$$\pi(D) = \{[x_i]_D\}_{i=1}^n = \{[x_1]_D, \dots, [x_n]_D\}.$$

Based on the fuzzy equivalence relation, we propose a fuzzy distance between the conditional attribute set and the decision attribute set. In Section 3, the attribute set $P \subseteq C$ determined a fuzzy partition $\pi(R_P)$. Thus, for simplicity we replace the concept fuzzy distance between two partitions with the concept fuzzy distance between two attribute sets.

Definition 1. Let $DS=(U, C \cup D)$ be a decision table with numerical attributes, $\pi(R_P)$, $\pi(R_Q)$ be two fuzzy partitions induced by two fuzzy equivalence relations R_P , R_Q on $P, Q \subseteq C$. Then, fuzzy distance between P and Q , denote by $d_{NF}(P, Q)$, is defined a fuzzy distance between two fuzzy partitions $\pi(R_P)$ and $\pi(R_Q)$, it means that $d_{NF}(P, Q) = D(\pi(R_P), \pi(R_Q))$.

Proposition 3. Let $DS=(U, C \cup D)$ be a decision table with numerical attributes, where $U = \{x_1, x_2, \dots, x_n\}$ and R be a fuzzy equivalence relation determined on conditional attributes. Then, fuzzy distance between two attribute sets C and $C \cup D$ which is determined as

$$(2) \quad d_{NF}(C, C \cup D) = \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_{R_D}|}{n} \right).$$

Proof: From Definition 1 and Theorem 1, we have:

$$\begin{aligned} d_{NF}(C, C \cup D) &= D(\pi(R_C), \pi(R_{C \cup D})) = \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| + |[x_i]_{R_{C \cup D}}| - 2|[x_i]_{R_C} \cap [x_i]_{R_{C \cup D}}|}{n} \right) = \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_{R_D}|}{n} \right) = \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| + |[x_i]_{R_C} \cap [x_i]_{R_D}| - 2|[x_i]_{R_C} \cap [x_i]_{R_D}|}{n} \right) = \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_{R_D}|}{n} \right). \end{aligned}$$

It is easy to see that $0 \leq d_{NF}(C, C \cup D) \leq 1 - \frac{1}{n}$; $d_{NF}(C, C \cup D) = 0 \Leftrightarrow \pi(R_C) \preceq \pi(R_D)$ and $d_{NF}(C, C \cup D) = 1 - \frac{1}{n} \Leftrightarrow \pi(R_C) = \pi(\delta)$ and $[x_i]_{R_D} = \{x_i\}$ for $1 \leq i \leq n$.

Proposition 4. Let $DS=(U, C \cup D)$ be a decision table with numerical attributes, where $U = \{x_1, x_2, \dots, x_n\}$, $B \subseteq C$ and R be a fuzzy equivalence relation determined on conditional attributes. Then, $d_{NF}(B, B \cup D) \geq d_{NF}(C, C \cup D)$.

Proof. From $B \subseteq C$, according to [14] we have $\pi(R_C) \preceq \pi(R_B)$, which means that it can be inferred from $[x_i]_{R_C} \subseteq [x_i]_{R_B}$ that $|[x_i]_{R_C}| \leq |[x_i]_{R_B}|$ for $1 \leq i \leq n$.

For any $x_i \in U$, we have:

$$|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_D| = \sum_{j=1}^n \mu_{[x_i]_{R_C}}(x_j) - \sum_{j=1}^n \min\{\mu_{[x_i]_{R_C}}(x_j), \mu_{[x_i]_D}(x_j)\},$$

$$|[x_i]_{R_B}| - |[x_i]_{R_B} \cap [x_i]_D| = \sum_{j=1}^n \mu_{[x_i]_{R_B}}(x_j) - \sum_{j=1}^n \min\{\mu_{[x_i]_{R_B}}(x_j), \mu_{[x_i]_D}(x_j)\}.$$

(1) For any $x_j \in [x_i]_D$ we have $\mu_{[x_i]_D}(x_j) = 1$, therefore

$$|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_D| = 0 = |[x_i]_{R_B}| - |[x_i]_{R_B} \cap [x_i]_D|.$$

(2) For any $x_j \notin [x_i]_D$ we have $\mu_{[x_i]_D}(x_j) = 0$, therefore

$$|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_D| = |[x_i]_{R_C}| \leq |[x_i]_{R_B}| = |[x_i]_{R_B}| - |[x_i]_{R_B} \cap [x_i]_D|.$$

From (1), (2) we have:

$$\begin{aligned} & |[x_i]_{R_B}| - |[x_i]_{R_B} \cap [x_i]_D| \geq |[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_D| \Leftrightarrow \\ \Leftrightarrow & \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_B}| - |[x_i]_{R_B} \cap [x_i]_D|}{n} \right) \geq \frac{1}{n} \sum_{i=1}^n \left(\frac{|[x_i]_{R_C}| - |[x_i]_{R_C} \cap [x_i]_D|}{n} \right) \Leftrightarrow \\ & \Leftrightarrow d_{\text{NF}}(B, B \cup D) \geq d_{\text{NF}}(C, C \cup D). \end{aligned}$$

It is easy to see that $d_{\text{NF}}(B, B \cup D) = d_{\text{NF}}(C, C \cup D) \Leftrightarrow |[x_i]_{R_B}| = |[x_i]_{R_C}|$ for any $x_i \in U$.

In next part, we present an attribute reduction method of the decision table using the fuzzy distance measure in Proposition 3. Our method includes: defining the reduct based on fuzzy distance, defining the importance of the attribute and designing a heuristic algorithm to find the best reduct based on the importance of the attribute.

Definition 2. Let $DS = (U, C \cup D)$ be a decision table with numerical attributes, $B \subseteq C$ and R be a fuzzy equivalence relation determined on conditional attributes. If

- 1) $d_{\text{NF}}(B, B \cup D) = d_{\text{NF}}(C, C \cup D)$,
- 2) $\forall b \in B, d_{\text{NF}}(\{B - \{b\}\}, \{B - \{b\}\} \cup D) \neq d_{\text{NF}}(C, C \cup D)$,

then B is a reduct of C based on fuzzy distance.

Definition 3. Let $DS = (U, C \cup D)$ be a decision table with numerical attributes, $B \subset C$ and $b \in C - B$. The importance of attribute b with respect to B is defined as

$$\text{SIG}_B(b) = d_{\text{NF}}(B, B \cup D) - d_{\text{NF}}(B \cup \{b\}, B \cup \{b\} \cup D).$$

From Proposition 4, we have $SIG_B(b) \geq 0$. The importance of $SIG_B(b)$ characterizes the classification accuracy of attribute b with respect to decision attribute D . It is used as the attribute selection criterial for heuristic algorithms to find the best reduct.

Algorithm F_DBAR (Fuzzy distance based on attribute reduction): A heuristic algorithm to find the best reduct by using fuzzy distance.

Input: Decision table with numerical attributes $DS = (U, C \cup D)$, fuzzy equivalence relation R .

Output: The best reduct B

Step 1. $B \leftarrow \emptyset$; $M(R_B) = [1]_{n \times n}$;

Step 2. Calculate relation matrix $M(R_C)$, calculate equivalence matrix $M(\text{IND}(D))$, calculate fuzzy distance $d_{\text{NF}}(C, C \cup D)$;

//Adding gradually to B an attribute having the greatest importance

Step 3. While $d_{\text{NF}}(B, B \cup D) \neq d_{\text{NF}}(C, C \cup D)$ do

Step 4. Begin

Step 5. For each $a \in C - B$ calculate

$$SIG_B(a) = d_{\text{NF}}(B, B \cup D) - d_{\text{NF}}(B \cup \{a\}, B \cup \{a\} \cup D);$$

Step 6. Select $a_m \in C - B$ so that $SIG_B(a_m) = \text{Max}_{a \in C - B} \{SIG_B(a)\}$;

Step 7. $B = B \cup \{a_m\}$;

Step 8. End;

//Remove redundant attribute in B

Step 9. For each $a \in B$

Step 10. Begin

Step 11. Calculate $d_{\text{NF}}(B - \{a\}, B - \{a\} \cup D)$;

Step 12. If $d_{\text{NF}}(B - \{a\}, B - \{a\} \cup D) = d_{\text{NF}}(C, C \cup D)$ then $B = B - \{a\}$;

Step 13. End;

Step 14. Return B ;

Example 3. Let $DS = (U, C \cup D)$ be a decision table with numerical attributes (Table 1) where $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$, $C = \{c_1, c_2, c_3, c_4, c_5, c_6\}$, the fuzzy equivalence relation R is defined in the Formula (2).

Table 1. The decision table with numerical attributes

$c_i \backslash u_i$	C_1	C_2	C_3	C_4	C_5	C_6	D
u_1	0.8	0.2	0.6	0.4	1	0	0
u_2	0.8	0.2	0	0.6	0.2	0.8	1
u_3	0.6	0.4	0.8	0.2	0.6	0.4	0
u_4	0	0.4	0.6	0.4	0	1	1
u_5	0	0.6	0.6	0.4	0	1	1
u_6	0	0.6	0	1	0	1	0

By using steps of F_DBAR algorithm to find the best reduct, we have

$$B \leftarrow \emptyset; M(R_B) = [1]_{n \times n}; d_{NF}(\emptyset, \emptyset \cup D) = 0.375,$$

and we calculate some relation matrices

$$M(R_{c_1}), M(R_{c_2}), M(R_{c_3}), M(R_{c_4}), M(R_{c_5}), M(R_{c_6}), M(R_C),$$

and the equivalence matrix $M(\text{IND}(D))$:

$$M(R_{c_1}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}, M(R_{c_2}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix},$$

$$M(R_{c_3}) = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}, M(R_{c_4}) = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$M(R_{c_5}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \end{bmatrix}, M(R_{c_6}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \\ 0 & 0.2 & 0 & 1 & 1 & 1 \end{bmatrix},$$

$$M(R_C) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, M(\text{IND}(D)) = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Calculate

$$d_{NF}(C, C \cup D) = 0, \quad d_{NF}(\{c_1\}, \{c_1\} \cup D) = 0.166667,$$

$$d_{NF}(\{c_2\}, \{c_2\} \cup D) = 0.166667, \quad d_{NF}(\{c_3\}, \{c_3\} \cup D) = 0.166667,$$

$$d_{NF}(\{c_4\}, \{c_4\} \cup D) = 0.111111, \quad d_{NF}(\{c_5\}, \{c_5\} \cup D) = 0.122222,$$

$$d_{NF}(\{c_6\}, \{c_6\} \cup D) = 0.122222; \quad \text{SIG}_B(c_1) = 0.208333333, \quad \text{SIG}_B(c_2) = 0.208333333,$$

$$\text{SIG}_B(c_3)=0.208333333, \text{SIG}_B(c_4)=0.263888889, \text{SIG}_B(c_5)=0.252777778,$$

$$\text{SIG}_B(c_6)=0.252777778.$$

Attribute $\{c_4\}$ is selected.

Similarity, $d_{\text{NF}}(\{c_4, c_1\}, \{c_4, c_1\} \cup D) = 0$, checked

$$d_{\text{NF}}(\{c_4, c_1\}, \{c_4, c_1\} \cup D) = d_F(C, C \cup D) = 0,$$

the algorithm finished and $B = \{c_4, c_1\}$. Consequently, $B = \{c_4, c_1\}$ is the best reduct of DS.

5. Experiments

We select eight data sets with numerical attributes from the UCI repository [17] to test proposed algorithm in Table 2. Environmental testing is PC with Pentium Core i3, 2.4 GHz CPU, 2 GB RAM, using Windows 10 operating system.

Table 2. Data sets in the experimental analysis

Id	Data sets	Number of conditional attributes	Number of objects
1	Ecoli	7	336
2	Ionosphere	34	351
3	Wdbc (Breast Cancer Wisconsin)	30	569
4	Wpbc (Breast Cancer Wisconsin)	32	198
5	Wine	13	178
6	Glass	9	214
7	Sonar (Connectionist Bench)	60	208
8	Heart	13	270

We select FA_FPR algorithm (Finding Reduct based on Fuzzy Positive Region) and FA_FSCE algorithm (finding reduct based on fuzzy entropy) in [14] to compare with F_DBAR proposed algorithm on the classification accuracy of reduct. The FA_FPR algorithm is an improvement of FAR-VPFRS algorithm in [8] on executed time, the FA_FSCE is an improvement of FSCE algorithm in [7] on executed time. According to fuzzy rough set approach, the classification accuracy of FAR-VPFRS algorithm [8], FSCE algorithm [7] are almost higher than the ones in rough set approach after discretized data. However, authors [11] have not evaluated the classification accuracy for algorithms FA_FSCE, FA_FPR. For testing, we perform the following tasks:

- 1) Code FA_FPR, FA_FSCE and F_DBAR algorithms by program C#. Algorithms used the fuzzy equivalence relation defined by the formula (2).
- 2) Execute three algorithms on eight data sets by environment testing.
- 3) Use C4.5 algorithm in WEKA [18] to evaluate the classification accuracy of three algorithms by selecting 2/3 first objects as training set and the remainder objects as testing set.

Table 3 shows the testing results of eight data sets where $|U|$ is the number of objects, $|C|$ is the number of the conditional attribute, $|R|$ is the number of attribute of the reduct for each algorithm.

Table 3. The experimental result of three algorithms FA_FSCE, FA_FPR, F_DBAR

Id	Data set	$ U $	$ C $	FA_FSCE Algorithm		FA_FPR Algorithm		F_DBAR Algorithm	
				$ R $	Classification accuracy of C4.5 (%)	$ R $	Classification accuracy of C4.5 (%)	$ R $	Classification accuracy of C4.5 (%)
1	Ecoli	336	7	6	81.50	7	82.45	7	82.45
2	Ionosphere	351	34	11	88.72	13	91.52	15	94.25
3	Wdbc	569	30	16	95.2	17	90.46	19	92.84
4	Wpbc	198	32	16	65.32	17	73.60	18	74.60
5	Wine	178	13	5	88.72	9	91.57	10	89.25
6	Glass	214	9	6	80.15	7	81.56	7	81.56
7	Sonar	208	60	8	75.40	12	70.60	13	76.25
8	Heart	270	13	8	74.62	9	76.95	10	78.65
The average classification accuracy of C4.5					81.2		82.33		83.73

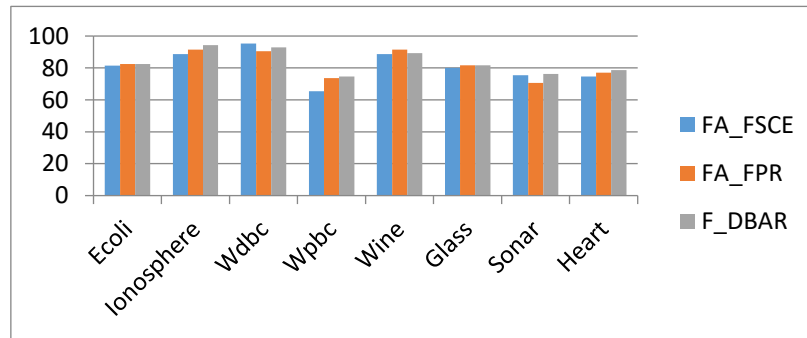


Fig. 1. The classification accuracy C4.5 of FA_FSCE, FA_FPR and F_DBAR

The experimental results in Table 3 and Fig. 1 show that the average classification accuracy of F_DBAR (used the fuzzy distance) is highest, next to FA_FPR (used fuzzy positive region) and FA_FSCE is lowest (used fuzzy entropy). For each data set, the classification accuracy of three algorithms are different. Consequently, the classification accuracy of algorithm F_DBAR is the best one of three algorithms.

6. Conclusion

The aim of attribute reduction in the decision table is to improve the accuracy of classification model. Among attribute reductions in the decision table with numerical value domain, related researches show that attribute reduction methods based on fuzzy rough set approach have the classification accuracy more higher than one based on traditional rough set. In this paper, we propose an attribute reduction method on the decision table with numerical attribute value which uses fuzzy distance based on fuzzy rough set. Our research includes: proposing a new

fuzzy distance between two fuzzy partitions, defining reduct and importance of attributes based on fuzzy distance, proposing a heuristic algorithm to find the best reduct. The experimental results from data sets show that the classification accuracy of fuzzy distance method is higher than that of the ones using fuzzy positive region and fuzzy entropy. Our further research approach issue is finding the relation between reduct obtained by different methods to subgroup and overall evaluation of methods based on fuzzy rough set approach.

Acknowledgements: This research has been funded by the Research Project, VAST 01.08/16-17. Vietnam Academy of Science and Technology.

References

1. Dubois, D., H. Prade. Rough Fuzzy Sets and Fuzzy Rough Sets. – International Journal of General Systems, Vol. **17**, 1990, pp. 191-209.
2. Demetrovics, J., V. D. Thi, N. L. Giang. An Efficient Algorithm for Determining the Set of All Reductive Attributes in Incomplete Decision Tables. – Cybernetics and Information Technologies, Vol. **13**, 2013, No 4, pp. 118-126.
3. Demetrovics, J., V. D. Thi, N. L. Giang. On Finding All Reducts of Consistent Decision Tables. – Cybernetics and Information Technologies, Vol. **14**, 2014, No 4, pp. 3-10.
4. Demetrovics, J., N. Thi, L. Huong, V. D. Thi, N. L. Giang. Metric Based Attribute Reduction Method in Dynamic Decision Tables. – Cybernetics and Information Technologies, Vol. **16**, 2016, No 2, pp. 3-15.
5. Tsang, E. C. C., D. G. Chen, D. S. Yeung, X. Z. Wang, J. W. T. Lee. Attributes Reduction Using Fuzzy Rough Sets. – IEEE Trans. Fuzzy Syst., Vol. **16**, 2008, pp. 1130-1141.
6. Dai, J., Q. Xu. Attribute Selection Based on Information Gain Ratio in Fuzzy Rough Set Theory with Application to Tumor Classification. – Applied Soft Computing, Vol. **13**, 2013, pp. 211-221.
7. Hu, Q., D. R. Yu, Z. X. Xie. Information-Preserving Hybrid Data Reduction Based on Fuzzy-Rough Techniques. – Pattern Recognit. Lett., Vol. **27**, 2006, No 5, pp. 414-423.
8. Hu, Q., Z. X. Xie, D. R. Yu. Hybrid Attribute Reduction Based on a Novel Fuzzy-Rough Model and Information Granulation. – Pattern Recognit., Vol. **40**, 2007, pp. 3509-3521.
9. Jensen, R., Q. Shen. Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough-Based Approaches. – IEEE Trans. Knowl. Data Eng., Vol. **16**, 2004, No 12, pp. 1457-1471.
10. Jensen, R., Q. Shen. Fuzzy-Rough Attribute Reduction with Application to Web Categorization. – Fuzzy Sets Syst., Vol. **141**, 2004, pp. 469-485.
11. Jensen, R., Q. Shen. Fuzzy-Rough Sets Assisted Attribute Reduction. – IEEE Trans. Fuzzy Syst., Vol. **15**, 2007, No 1, pp. 73-89.
12. Jensen, R., Q. Shen. New Approaches to Fuzzy-Rough Feature Selection. – IEEE Trans. Fuzzy Syst., Vol. **17**, 2009, No 4, pp. 824-838.
13. Bhatt, R. B., M. Gopal. On Fuzzy-Rough Sets Approach to Feature Selection. – Pattern Recognit. Lett., Vol. **26**, 2005, pp. 965-975.
14. Qian, Y. H., Q. Wang, H. H. Cheng, J. Y. Liang, C. Y. Dang. Fuzzy-Rough Feature Selection Accelerator. – Fuzzy Sets and Systems, Vol. **258**, 2015, pp. 61-78.
15. Pawlak, Z. Rough Sets: Theoretical Aspects of Reasoning about Data. London, Kluwer Academic Publisher, 1991.
16. Pawlak, Z., J. W. Grzymala-Busse, R. Slowinski, W. Ziako. Rough Sets. Commun. – ACM, Vol. **38**, 1995, No 11, pp. 89-95.
17. The UCI Machine Learning Repository.
<http://archive.ics.uci.edu/ml/datasets.html>
18. <https://sourceforge.net/projects/weka/>