Vietnam Journal of Science and Technology ***59*** *(2) (2021) 261-274*

doi:10.15625/2525-2518/59/2/15698

FUZZY DISTANCE BASED FILTER-WRAPPER INCREMENTAL ALGORITHMS FOR ATTRIBUTE REDUCTION WHEN ADDING OR DELETING ATTRIBUTE SET

Ho Thị Phuong[1], **Nguyen Long Giang**[2, *]

[1]*Tay Nguyen University, 567 Le Duan, Buon Ma Thuot , DakLak, Viet Nam,*
km 10 Nguyen Trai, Thanh Xuan, Ha Noi, Viet Nam

[2]*Institute of Information Technology, Vietnam Academy of Science and Technology,*
18 Hoang Quoc Viet, Cau Giay, Ha Noi, Viet Nam

[*]Email: *nlgiang@ioit.ac.vn*

Received: 24 November 2020; Accepted for publication: 25 January 2021

Abstract: Attribute reduction is a critical problem in the data preprocessing step with the aim of minimizing redundant attributes to improve the efficiency of data mining models. The fuzzy rough set theory is considered an effective tool to solve the attribute reduction problem directly on the original decision system, without data preprocessing. With the current digital transformation trend, decision systems are larger in size and updated. To solve the attribute reduction problem directly on change decision systems, a number of recent studies have proposed incremental algorithms to find reducts according to fuzzy rough set approach to reduce execution time. However, the proposed algorithms follow the traditional filter approach. Therefore, the obtained reduct is not optimal in both criteria: the number of attribute of the reducts and the accuracy of classification model. In this paper, we propose incremental algorithms that find reducts by filter-wrapper approach using fuzzy distance measure in the case of adding and deleting attribute set. The experimental results show that the proposed algorithms significantly reduce the number of attributes in reduct and improve the classification accuracy compared to other algorithms using filter approach.

Keywords: decision system, fuzzy distance, incremental algorithm, attribute reduction, reduct.

Classification numbers: 4.7.3, 4.7.4, 4.8.3.

1. INTRODUCTION

Attribute reduction is a crucial problem in the data preprocessing step of data mining and knowledge discovery. The goal of attribute reduction is to remove redundant attributes as much as possible to improve the efficiency of data mining models [1 - 3]. The fuzzy rough set theory proposed by Dubois *et al*. [4] is an effective tool to solve the attribute reduction problem directly on the original decision system without data preprocessing step to effectively improve accuracy of classification model. Up to now, many methods of attribute reduction according to fuzzy rough set approach have been proposed, typically fuzzy membership functions methods [5, 6], fuzzy positive domain methods [7, 8], fuzzy entropy methods [9], fuzzy distance methods [10, 11]. In the current digital transformation trend, decision systems have an increasingly large number of attributes, for example, data tables in the bioinformatics field have millions of attributes. Furthermore, decision systems are always changing, updating with scenarios such as adding and deleting object sets, adding and deleting attribute sets, object set values, and changing attribute sets. To build an effective classification model, we need to solve the attribute reduction problem on change and large decision systems. Attribute reduction method in traditional approach applied on such decision systems often faces two challenges. Firstly, with large size decision systems, the implementation of the reduct finding algorithms has difficulty in storage space and computation speed. Secondly, with updated and changed decision systems, these algorithms have to recalculate the reduct on the entire decision system after the change, so the computation time cost increases significantly. To solve the above two challenges, the researchers propose an incremental computational approach to find the reduct. The incremental algorithms only update the reduct on the changed data part and not recalculate the reduct on the entire original decision system. Therefore, they are much shorter time.

According to the fuzzy rough set approach [4], in recent years, a number of incremental algorithms for calculating reduct have been proposed according to the following cases: adding and deleting an object set [12 - 17], adding and deleting an attribute set [18]. For the case of adding and deleting an object set, Liu *et al*. [12] proposed the algorithm FIAT to find reduct based on fuzzy dependence. Yang *et al*. [13] built the incremental algorithm IARM to calculate reduct based on fuzzy indiscernibility relation. Yang *et al*. [14] proposed the algorithms IV-FS-FRS-1 and IV-FS-FRS-2 to get reduct based on fuzzy indiscernibility relation. Zhang *et al*. [16] proposed the incremental algorithm AIFWAR to find reduct using extended conditional entropy. Ni *et al.* [17] proposed two incremental algorithms to find reduct based on the key instance set: the algorithm DIAR uses fuzzy membership function and the algorithm PIAR uses fuzzy positive domain. In the case of adding and deleting an attribute set, the research results of the incremental algorithms for calculating reduct based on fuzzy rough set approach are still limitations. Zeng *et al*. [18] constructed incremental formulas for updating fuzzy dependency in hybrid information system (HIS), on that basis they proposed two incremental algorithms for calculating reduct using fuzzy dependency: the algorithm FRSA-IFS-HIS (AA) in the case of adding an attribute set and the algorithm FRSA-IFS-HIS (AD) in the case of deleting an attribute set. The experimental results in the above works show that the incremental algorithms significantly decrease the time execution compared to non-incremental algorithms. As a result, they can be effectively performed on large, changeable, up-to-date decision systems. However, most of the proposed algorithms follow the traditional filter approach. With this filter approach, the obtained reduct is the minimal set of conditional attributes that preserve original measure. The classification accuracy is calculated after obtaining reduct. Therefore, the obtained reduct is not the best choice on two criteria: the number of attribute of reduct and the classification accuracy.

In order to decrease the cardinality of obtained reduct, in cases of adding and deleting object sets, Giang *et al*. [15] proposed filter-wrapper incremental algorithms for finding reduct of complete decision tables by fuzzy rough set approach. Quang *et al*. [19] proposed filter-wrapper incremental algorithms for finding reduct of incomplete decision tables by tolerance rough set. In this proposed filter-wrapper algorithms, the filter phase finds the candidate for reduct when adding the most important attribute, while the wrapper phase obtains reduct with the highest classification accuracy. The experimental results show that the reduct of filter-wrapper approach minimizes the number of attributes and improves the classification accuracy compared to the filter approach. Therefore, the research motivation of the paper is to apply the filter-wrapper approach to develop incremental attribute reduction algorithms in case of adding and deleting an attribute set in order to minimize the number of attribute of the reduct and improve the classification accuracy.

In this paper, we propose two filter-wrapper incremental algorithms to find reduct of the decision system using fuzzy distance measure in the paper [15]: the algorithm IFW_FDAR_AA in case of adding an attribute set and the algorithm IFW_FDAR_DA in case of deleting an attribute set. The structure of the paper is as follows: Section 2 presents some basic concepts. Section 3 proposes the incremental filter-wrapper algorithm IFW_FDAR_AA to find reduct when adding an attribute set. Section 4 proposes the filter-wrapper incremental algorithm IFW_FDAR_DA that updates reduct when deleting an attribute set. Section 5 presents the experimental results of proposed algorithms. Finally, is conclusion and research direction.

2. PRELIMINARY

This section presents some concepts of fuzzy rough set model [4]. The decision system is a pair of $DS=(U,C\cup D)$ in which *U* is a finite, non-empty set of objects; *C* is set of conditional attributes, *D* is decisive attribute set with $C\cap D=\varnothing$. The fuzzy rough set theory proposed by D. Dubois *et al*. [4] uses fuzzy equivalence relations to approximate fuzzy sets. Let us consider the decision system $DS=(U,C\cup D)$, a relation defined on the attribute value domain is called fuzzy equivalence relation if the following conditions are met:

1) Reflexive: $\tilde{R}(x,x)=1$;

2) Symmetric: $\tilde{R}(x,y)=\tilde{R}(y,x)$;

3) Max-min transitive: $\tilde{R}(x,z)\geq\min\left\{\tilde{R}(x,y),\tilde{R}(y,z)\right\})$ with every $x,y,z\in U$.

Given two fuzzy equivalent relations R_P and R_Q defined on $P,Q\subseteq C$, then for all $x,y\in U$ we have [18]:

1) $R_P=R_Q\Leftrightarrow R_P(x,y)=R_Q(x,y)$

2) $R=R_P\cup R_Q\Leftrightarrow R(x,y)=\max\left\{R_P(x,y),R_Q(x,y)\right\}$

3) $R=R_P\cap R_Q\Leftrightarrow R(x,y)=\min\left\{R_P(x,y),R_Q(x,y)\right\}$

4) $R_P\subseteq R_Q\Leftrightarrow R_P(x,y)\leq R_Q(x,y)$

The relation R_P is presented by the fuzzy equivalent matrix $M(R_P)=\left[p_{\text{ij}}\right]_{n\times n}$ as follows:

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

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

For any $P,Q\subseteq C$, we have $R_P=\cap_{a\in P}R_a$ và $R_{P\cup Q}=R_P\cap R_Q$, that is for any $x,y\in U$, $R_{P\cup Q}(x,y)=\min\left\{R_P(x,y),R_Q(x,y)\right\}$. Suppose that $M(R_P)=\left[p_{\text{ij}}\right]_{n\times n}$ and

$M(R_Q)=\left[q_{\mathrm{ij}}\right]_{n\times n}$ are two fuzzy equivalent matrices on R_P, R_Q, then the fuzzy equivalent matrix on $S=P\cup Q$ is as follows:

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

For any $P\subseteq C$, $U=\{x_1,x_2,...,x_n\}$, fuzzy equivalent relation R_P determines an fuzzy partition $\pi\left(R_P\right)=U/R_P$ on U as $\pi\left(R_P\right)=U/R_P=\left\{[x_i]_P\right\}_{i=1}^{n}=\left\{[x_1]_P,...,[x_n]_P\right\}$ where $[x_i]_P=p_{i1}/x_1+p_{i2}/x_2+...+p_{in}/x_n$ is a fuzzy set as fuzzy equivalent class of object x_i. Membership function of object x_i is defined as $\mu_{[x_i]_P}\left(x_j\right)=\mu_{R_P}\left(x_i,x_j\right)=R_P\left(x_i,x_j\right)=p_{ij}$ where $x_j\in U$. Then, the cardinality of the fuzzy equivalent class $[x_i]_P$ is calculated as $\left|[x_i]_P\right|=\sum_{j=1}^{n}p_{ij}$.

3. FILTER-WRAPPER INCREMENTAL ALGORITHMS FOR CALCULATING REDUCT WHEN ADDING AN ATTRIBUTE SET

In this section, we construct an incremental algorithm for calculating reduct of decision system using fuzzy distance in the paper [15] when adding an attribute set. The proposed algorithm follows a filter-wrapper hybrid approach. First of all, we construct a formula to calculate the fuzzy distance when adding an attribute set.

Given an decision system $DS=(U,C\cup D)$ where $U=\{u_1,u_2,...,u_n\}$. Then, the fuzzy distance between two attribute sets of C and D in the paper [15] is determined as follows:

$$D(C,C\cup D)=\frac{1}{n^2}\sum_{i=1}^{n}\left(\left|[u_i]_C\right|-\left|[u_i]_C\cap[u_i]_D\right|\right) \tag{1}$$

Proposition 1. *Given an decision system* $DS=(U,C\cup D)$ *where* $U=\{u_1,u_2,...,u_n\}$. *Assume that the conditional attribute set B is added into C where* $B\cap C=\varnothing$. *Assume that* $M(R_B)=\left[b_{\mathrm{ij}}\right]_{n\times n}$, $M(R_C)=\left[c_{\mathrm{ij}}\right]_{n\times n}$, $M(R_D)=\left[d_{\mathrm{ij}}\right]_{n\times n}$ *are fuzzy equivalent matrices of* R_B,R_C,R_D *on B, C, D; respectively. Then we have:*

1) If $c_{ij}\le d_{ij}$ *for any* $1\le i,j\le n$ *then* $D(C\cup B,C\cup B\cup D)=0$

2) If $b_{ij}\ge c_{ij}$ *for any* $1\le i,j\le n$ *then*

$$D(C\cup B,C\cup B\cup D)=D(C,C\cup D)=\frac{1}{n^2}.\sum_{i=1}^{n}\sum_{j=1}^{n}\left(c_{ij}-\min\left(c_{ij},d_{ij}\right)\right)$$

3) If $b_{ij}<c_{ij}$ *for any* $1\le i,j\le n$ *then*

$$D(C\cup B,C\cup B\cup D)=D(B,B\cup D)=\frac{1}{n^2}.\sum_{i=1}^{n}\sum_{j=1}^{n}\left(b_{ij}-\min\left(b_{ij},d_{ij}\right)\right)$$

Proof. When adding *B* into *C*, the fuzzy distance is defined as [15]:

$$D(C\cup B,C\cup B\cup D)=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_{C\cup B}\right|-\left|[u_i]_{C\cup B}\cap[u_i]_D\right|\right)$$

$$=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\cap[u_i]_B\right|-\left|[u_i]_C\cap[u_i]_B\cap[u_i]_D\right|\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\min\left(c_{ij},b_{ij}\right)-\min\left(c_{ij},b_{ij},d_{ij}\right)\right)$$

1) If $c_{ij}\le d_{ij}$ for any $1\le i,j\le n$ then $[u_i]_C\subseteq[u_i]_D$ and $[u_i]_C\cap[u_i]_B\cap[u_i]_D=[u_i]_C\cap[u_i]_B$. Then we have:

$$D\left(C\cup B,C\cup B\cup D\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_{C\cup B}\right|-\left|[u_i]_{C\cup B}\cap[u_i]_D\right|\right)$$

$$=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\cap[u_i]_B\right|-\left|[u_i]_C\cap[u_i]_B\cap[u_i]_D\right|\right)=0$$

2) From $b_{ij}\ge c_{ij}$ we have $[u_i]_C\subseteq[u_i]_B$ và $[u_i]_C\cap[u_i]_B=[u_i]_C$ for any $u_i\in U$. Then we have:

$$D\left(C\cup B,C\cup B\cup D\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\cap[u_i]_B\right|-\left|[u_i]_C\cap[u_i]_B\cap[u_i]_D\right|\right)$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}\left(\left|[u_i]_C\right|-\left|[u_i]_C\cap[u_i]_D\right|\right)=D\left(C,C\cup\{d\}\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\sum_{j=1}^{n}\left(c_{ij}-\min\left(c_{ij},d_{ij}\right)\right)$$

3) From $b_{ij}<c_{ij}$, $[u_i]_B\subset[u_i]_C$ and $[u_i]_C\cap[u_i]_B=[u_i]_B$ for any $u_i\in U$. Then we have:

$$D\left(C\cup B,C\cup B\cup D\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\cap[u_i]_B\right|-\left|[u_i]_C\cap[u_i]_B\cap[u_i]_D\right|\right)$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}\left(\left|[u_i]_B\right|-\left|[u_i]_B\cap[u_i]_D\right|\right)=D\left(B,B\cup D\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\sum_{j=1}^{n}\left(b_{ij}-\min\left(b_{ij},d_{ij}\right)\right)$$

Based on the formula for calculating fuzzy distance in Proposition 1, we next construct an incremental algorithm for calculating reduct using fuzzy distance according to the filter-wrapper hybrid approach. First of all, we define reduct and the importance of attribute based on the fuzzy distance.

Definition 1. [15] Given a decision system $DS=\left(U,C\cup D\right)$ where $B\subseteq C$ and R_B,R_C are fuzzy equivalent relations on B,C respectively. If

1) $D\left(B,B\cup D\right)=D\left(C,C\cup D\right)$

2) $\forall b\in B,\ D\left(B-\{b\},\{B-\{b\}\}\cup D\right)\ne D\left(C\cup D\right)$

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

Definition 2. [15] Given a decision system $DS=\left(U,C\cup D\right)$ where $B\subset C$ and $b\in C-B$. the importance of attribute b with respect to B is defined as

$$SIG_B\left(b\right)=D\left(B,B\cup D\right)-D\left(B\cup\{b\},B\cup\{b\}\cup D\right)$$

Importance $SIG_B\left(b\right)\ge 0$ characterize the classification quality of attribute b for decision attribute set D and it is used as the attribute selection criterion for the algorithm finding reduct.

From the formula for calculating fuzzy distance in Proposition 1 and Definition 1, we have the following Proposition 2:

Proposition 2. *Given a decision system* $DS=(U,C\cup D)$ *where* $U=\{u_1,u_2,...,u_n\}$ *and* $R\subseteq C$ *is a reduct base on fuzzy distance. Assume that the conditional attribute set B is added into C where* $B\cap C=\varnothing$. $M(R_B)=\left[b_{\text{ij}}\right]_{n\times n}$, $M(R_C)=\left[c_{\text{ij}}\right]_{n\times n}$, $M(R_D)=\left[d_{\text{ij}}\right]_{n\times n}$ *are fuzzy equivalent matrices of* R_B,R_C,R_D *on B, C, D, respectively. Then we have:*

1) If $b_{ij}\geq c_{ij}$ *for any* $1\leq i\leq n, 1\leq j\leq n$ *then R is a reduct of* $DS_1=(U,C\cup B\cup D)$.

2) If $b_{ij}<c_{ij}$ *for any* $1\leq i\leq n, 1\leq j\leq n$ *then B contains a reduct of* $DS_1=(U,C\cup B\cup D)$.

Proof.

1) Accrding to Proposition 1, if $b_{ij}\geq c_{ij}$ for $1\leq i\leq n, 1\leq j\leq n$ then $D(C\cup B, C\cup B\cup D)=D(C,C\cup D)$. By R is a reduct of DS then $D(R,R\cup D)=D(C,C\cup D)=D(C\cup B,C\cup B\cup D)$ and $\forall r\in R, D(R-\{r\},(R-\{r\})\cup D)\neq D(C,C\cup D)$. According to Definition 1, R is a reduct of $DS_1=(U,C\cup B\cup D)$.

2) According to Proposition 1, if $b_{ij}\geq c_{ij}$ for $1\leq i\leq n, 1\leq j\leq n$ then $D(C\cup B,C\cup B\cup D)=D(B,B\cup D)$, that is there exists $B_1\subseteq B$ such that B_1 satisfies Definition 1 about reduct of $DS_1=(U,C\cup B\cup D)$.

Next, we construct a filter-wrapper incremental algorithm that finds the reduct of a decision system using fuzzy distance when adding attribute set. The proposed algorithm consists of two phases: the filter phase finds candidates for the reduct when adding the attribute with the highest importance, the wrapper phase finds the reduct with the highest classification accuracy. The algorithm is described as follows:

Algorithm IFW_FDAR_AA (Incremental Filter-Wrapper Fuzzy Distance-based Attribute Reduction Algorithm when Adding Attributes).

```
Input:
1) Decision system $DS=(U,C\cup D)$ where $U=\{u_1,u_2,...,u_n\}$, reduct $R\subseteq C$, fuzzy
equivalent matrices $M(R_C)=\left[c_{\text{ij}}\right]_{n\times n}$, $M(R_D)=\left[d_{\text{ij}}\right]_{n\times n}$ of $R_C,R_D$, fuzzy distance $D(C,C\cup D)$;
2) The adding attribute set B where $B\cap C=\varnothing$;
Output: Reduct $R_1$ of $DS_1=(U,C\cup B\cup D)$
Step 1: Initialization and checking the adding attribute set
    1. $T:=\varnothing$; // Contains candidates for reduct
    2. Calculating fuzzy equivalent matrix $M(R_B)=\left[b_{\text{ij}}\right]_{n\times n}$;
    3. If $b_{ij}\geq c_{ij}$ for any $1\leq i\leq n, 1\leq j\leq n$ then Return $R$;
    4. If $b_{ij}<c_{ij}$ for any $1\leq i\leq n, 1\leq j\leq n$ then $R=\varnothing$; //finds the reduct in B
Step 2: Perform the algorithm for finding reduct
    // Filter stage, finding candidates for reduct from R.
    5. While $D(R,R\cup D)\neq D(C\cup B,C\cup B\cup D)$ do
    6. Begin
```

```
7.    For each $a \in B$, calculate $SIG_R(a) = D(R, R \cup D) - D(R \cup \{a\}, R \cup \{a\} \cup D)$
      where $D(R \cup \{a\}, R \cup \{a\} \cup D)$ is calculated by Proposition 1;
8.    Select $a_m \in B$ such that $SIG_R(a_m) = \underset{a \in B}{Max}\{SIG_R(a)\}$;
9.    $R := R \cup \{a_m\}$;
10.   $T := T \cup R$;
11. End;
// Wrapper stage, finding the reduct with the highest classification accuracy
12. Set $t := |T|$ //t is the number of T, T contains selected candidates, that is
      $T = \{R \cup \{a_{i_1}\}, R \cup \{a_{i_1}, a_{i_2}\}, ..., R \cup \{a_{i_1}, a_{i_2}, ..., a_{i_t}\}\}$;
13. Set $T_1 := R \cup \{a_{i_1}\}; T_2 := R \cup \{a_{i_1}, a_{i_2}\}; ...; T_t := R \cup \{a_{i_1}, a_{i_2}, ..., a_{i_t}\}$;
14. For $j = 1$ to $t$ calculate the classification accuracy on $T_j$ by a classifier;
15. $R_1 := T_{jo}$ where $T_{jo}$ has the highest classification accuracy;
Return $R_1$;
```

Finally, we evaluate the complexity of the algorithm IFW_FDAR_AA. Denoted by $|C|, |U|, |B|$ are number of conditional attributes, the number of objects, and the number of additional conditional attributes, respectively. In command line 2, the complexity for calculating the fuzzy equivalence relation $M(R_B)$ is $O(|B||U|^2)$. The complexity of the While loop is $O(|B|^2|U|^2)$ and the complexity of the filter phase is $O(|B|^2|U|^2)$. Assume that the classifier complexity is $O(T)$, then the complexity of the wrapper phase is $O(|B| * T)$. So, the complexity of the algorithm IFW_FDAR_AA is $O(|B|^2|U|^2) + O(|B| * T)$. If the non-incremental filter-wrapper algorithm FW_FDAR [15] is implemented directly on the decision system, the complexity is

$$O\left((|C| + |B|)^2 * |U|^2\right) + O\left((|C| + |B|) * T\right).$$

Consequently, the proposed incremental algorithm IFW_FDAR_AA significantly decreases the time complexity, especially in the case of small $|B|$.

4. FILTER-WRAPPER ALGORITHMS FOR UPDATING REDUCT WHEN DELETING ATTRIBUTE SET

In this section, we construct a filter-wrapper algorithm that updates reduct of a decision system when deleting an attribute set using fuzzy distance in paper [15]. First of all, we construct a formula for updating the distance when removing an attribute set by the following Proposition 3.

Proposition 3. *Given a decision system* $DS = (U, C \cup D)$ *where* $U = \{u_1, u_2, ..., u_n\}$. *Assume that the conditional attribute set B is deleted from C where* $B \subset C$ *and* $A = C - B$ *is the remaining* *attribute set.* $M(R_B)=\left[b_{\mathrm{ij}}\right]_{n\times n}$ *,* $M(R_C)=\left[c_{\mathrm{ij}}\right]_{n\times n}$ *,* $M(R_A)=\left[a_{\mathrm{ij}}\right]_{n\times n}$*,* $M(R_D)=\left[d_{\mathrm{ij}}\right]_{n\times n}$ *are fuzzy equivalent matrices of* R_B,R_C,R_A,R_D *respectively. Then we have*

$$D\left(A,A\cup\{d\}\right)=D\left(C,C\cup\{d\}\right)+\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[a_{ij}-c_{ij}+\min\left(c_{ij},d_{ij}\right)-\min\left(a_{ij},d_{ij}\right)\right]$$

Proof. We have:

$$D\left(A,A\cup D\right)=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_A\right|-\left|[u_i]_A\cap[u_i]_D\right|\right)$$

$$=\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\right|-\left|[u_i]_C\cap[u_i]_D\right|\right)+\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_A-[u_i]_C\right|\right)+\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_C\cap[u_i]_D\right|\right)-\frac{1}{n^2}.\sum_{i=1}^{n}\left(\left|[u_i]_A\cap[u_i]_D\right|\right)$$

$$=D\left(C,C\cup\{d\}\right)+\frac{1}{n^2}.\sum_{i=1}^{n}\left(a_{ij}-c_{ij}\right)+\frac{1}{n^2}.\sum_{i=1}^{n}\left(\min\left(c_{ij},d_{ij}\right)\right)-\frac{1}{n^2}.\sum_{i=1}^{n}\left(\min\left(a_{ij},d_{ij}\right)\right)$$

$$=D\left(C,C\cup\{d\}\right)+\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[a_{ij}-c_{ij}+\min\left(c_{ij},d_{ij}\right)-\min\left(a_{ij},d_{ij}\right)\right]$$

From Proposition 3, the filter-wrapper algorithm for updating reduct using fuzzy distance when deleting attribute set is as follows:

Algorithm IFW_FDAR_DA (Incremental Filter-Wrapper Fuzzy Distance-based Attribute Reduction Algorithm when Deleting Attributes).

Input:

1) Decision system $DS=\left(U,C\cup D\right)$ where $U=\{u_1,u_2,...,u_n\}$, reduct $R\subseteq C$, fuzzy equivalent matrices $M(R_C)=\left[c_{\mathrm{ij}}\right]_{n\times n}$, $M(R_D)=\left[d_{\mathrm{ij}}\right]_{n\times n}$, fuzzy distance $D\left(C,C\cup D\right)$;

2) Attribute set *B* is deleted from *C* where $B\subset C$;

Output: Reduct R_1 of $DS_1=\left(U,\left(C-B\right)\cup D\right)$;

1) ***Case 1:*** If $B\subseteq C-R$ then Retturn (*R*);
2) ***Case 2:*** If $R\subseteq B$ then perform non-incremental filter-wrapper algorithm based on fuzzy distance FW_FDAR in paper [15]**.**
3) ***Case 3:*** If $R\cap B\neq\varnothing$ then perform steps of attribute reduction algorithm.

Step 1: Initialization

1. *Set* $T:=\varnothing$; $A:=C-B$ *; // Contains candidates for reduct*
2. Calculating fuzzy equivalent matrices $M(R_B)=\left[b_{\mathrm{ij}}\right]_{n\times n}$, $M(R_A)=\left[a_{\mathrm{ij}}\right]_{n\times n}$
3. Set $R:=R-B$ *//check attributes in reduct*

Step 2*:* Perform the algorithm for finding reduct

// Filter stage, finding candidates for reduct from R

4. While $D\left(R,R\cup D\right)\neq D\left(A,A\cup D\right)$ do
5. Begin
6. For each $a\in R$ calculate $SIG_R\left(a\right)=D\left(R-\{a\},\{R-\{a\}\}\cup D\right)-D\left(R,R\cup D\right)$ where $D\left(R-\{a\},\{R-\{a\}\}\cup D\right)$ is calculated by Proposition 3;
7. Select $a_m\in R$ such that $SIG_R\left(a_m\right)=\underset{a\in R}{Min}\left\{SIG_R\left(a\right)\right\}$;

```
8.        $R := R - \{a_m\}$;
9.        $T := T \cup R$;
10.   End;
// Wrapper stage, finding the reduct with the highest classification accuracy
11.    Set $t := |T|$  // t is the number of T, T contains selected candidates, that is
          $T = \{R - \{a_{i_1}\}, R - \{a_{i_1}, a_{i_2}\}, ..., R - \{a_{i_1}, ..., a_{i_t}\}\}$;
12.    Set $T_1 = R - \{a_{i_1}\}, T_2 = R - \{a_{i_1}, a_{i_2}\}, ..., T_t = R - \{a_{i_1}, ..., a_{i_t}\}$
13.    For j = 1 to t  calculate the classification accuracy on $T_j$ by a classifier;
14.    $R_1 := T_{jo}$ where $T_{jo}$ has the highest classification accuracy;
15.    Return $R_1$;
```

Next, we evaluate the complexity of the algorithm IFW_FDAR_DA**.** Denoted by $|C|, |U|, |B|$ are number of conditional attributes, the number of objects, and the number of conditional attributes deleted from C, respectively. Let us consider the While loop of statements 4 to 10, the complexity of While loop is $O\left(|R-B|^2 * |U|^2\right)$ and the complexity of the filter phase is $O\left(|R-B|^2 * |U|^2\right)$. Assume that the classifier complexity is $O(T)$, then the complexity of the wrapper phase is $O(|R-B| * T)$. So, the complexity of algorithm IFW_FDAR_DA is $O\left(|R-B|^2 * |U|^2\right) + O(|R-B| * T)$. If the non-incremental FW_FDAR filter-wrapper algorithm is implemented [15] directly on the decision system with the attributes $C - B$, the complexity is $O\left(|C-B|^2 * |U|^2\right) + O(|C-B| * T)$. Therefore, the algorithm IFW_FDAR_DA is effective. If R is smaller, algorithm IFW_FDAR_DA will be more efficient.

5. EXPERIMENTAL RESULTS

In this section, we present the test results to evaluate the effectiveness of the proposed filter-wrapper incremental algorithm IFW_FDAR_AA with the filter incremental algorithm FRSA-IFS-HIS (AA) in [18] on the number of attribute reduction and the accuracy of the classification model. FRSA-IFS-HIS (AA) is the state-of-the-art filter incremental algorithm that finds reduct using fuzzy dependence in fuzzy rough set in case of adding attribute set. The testing was carried out on 04 sets of sample datasets from UCI data warehouse [20] as described in Table 1. On each data set, for attributes with real-value domains, we normalize the data domain to the fragment [0, 1] using the formula [14]

$$a'(x_i) = \frac{a(x_i) - \min(a)}{\max(a) - \min(a)} \tag{2}$$

with max(a), min(a) is the maximum and minimum value in the attribute value domain a. We use the fuzzy equivalence relation R_a in [14] on attribute a as follows

$$R_a(x_i, x_j) = 1 - |a(x_i) - a(x_j)| \text{ with } x_i, x_j \in U \tag{3}$$

Each attribute set is randomly divided into two parts: the original attribute set (column 5 Table 1) denoted by C_0, and the incremental attribute set (column 6 Table 1). The incremental attribute set is randomly divided into 5 equal parts, denoted by C_1, C_2, C_3, C_4, $C_{5,}$ respectively.

Table 1. Experimental data sets.

ID	**Data set**	**Number of objects**	**Number of conditional attributes**	**Original attribute set**	**Incremental attribute set**	**Number of decision class**
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Wisconsin diagnostic breast cancer (WDBC)	569	30	15	15	2
2	Credit approval (Credit)	690	15	5	10	2
3	German credit data (German)	1000	20	10	10	2
4	Waveform (Wave)	5000	21	11	10	3

To test IFW_FDAR_AA and FRSA-IFS-HIS (AA) algorithms, we firstly implement two algorithms on the data set with the original attribute set (consider the initial attribute set as increment). Next, perform two algorithms when adding from the first to the fifth part of the incremental attribute set, respectively. With the proposed algorithm of the IFW_FDAR_AA filter-wrapper hybrid approach, we use the CART classifier (Classification & Regression tree) to calculate the classification accuracy in the best accurate reduction set. We use the 10-fold cross-check method, which means that the data set is divided into 10 roughly equal parts, randomly taking 1 part as test data set, and the remaining 9 as training data. The process is repeated 10 times. Test implementation tool is Matlab R2016a. Test environment is PC with Intel (R) Core (TM) i7-3770CPU @ 3.40 GHz configuration, using Windows 7, 32 bit operating system.

Table 2 presents the comparison results of the number of attribute reductions (denoted by $|R|$) and the classification accuracy of the IFW_FDAR_AA and FRSA-IFS-HIS(AA) algorithms. Table 2 shows that, for each iteration step with incremental addition of the attribute set and across the entire attribute, IFW_FDAR_AA's classification accuracy is slightly higher than FRSA-IFS-HIS (AA) on all data sets. Moreover, the number of attribute reductions of IFW_FDAR_AA is much smaller than FRSA-IFS-HIS (AA), especially on the reduct with a large number of attributes like Libra. Therefore, implementation time and generalization of the classification rule set on the reduct of IFW_FDAR_AA are more efficient than FRSA-IFS-HIS (AA).

Table 2. The number of attributes in reduct and the classification accuracy of IFW_FDAR_AA and FRSA-IFS-HIS(AA).

ID	**Data set**	**Attribute set**	**The number of attributes**	**The total number of attribute**	**IFW_FDAR_AA**		**FRSA-IFS-HIS(AA)**	
					$\vert R\vert$	**Accuracy**	$\vert R\vert$	**Accuracy**
1	WDBC	C_0	15	15	3	76.14	5	75.96
		C_1	3	18	4	79.02	8	78.25

		C_2	3	21	4	79.02	9	79.82
		C_3	3	24	5	85.98	12	84.85
		C_4	3	27	6	93.18	15	89.36
		C_5	3	30	**6**	**93.18**	16	92.86
2	Credit	C_0	5	5	3	78.64	4	77.92
		C_1	2	7	4	81.92	5	80.15
		C_2	2	9	5	84.26	6	82.39
		C_3	2	11	5	84.26	6	82.39
		C_4	2	13	6	86.05	7	84.72
		C_5	2	15	**6**	**86.05**	8	85.96
3	German	C_0	10	10	5	72.16	6	70.46
		C_1	2	12	5	72.16	7	72.02
		C_2	2	14	6	73.08	8	73.08
		C_3	2	16	6	73.08	8	73.08
		C_4	2	18	7	74.28	10	73.92
		C_5	2	20	**7**	**74.28**	11	74.16
4	Wave	C_0	11	11	4	65.96	9	65.02
		C_1	2	13	5	68.72	11	67.78
		C_2	2	15	6	69.08	13	68.25
		C_3	2	17	6	69.08	14	68.97
		C_4	2	19	7	70.88	16	70.02
		C_5	2	21	**8**	**71.49**	17	70.85

Table 3 presents the comparison results of execution time of two algorithms IFW_FDAR_AA and FRSA-IFS-HIS (AA) (in seconds - s). Table 3 shows that the time of IFW_FDAR_AA is higher than FRSA-IFS-HIS (AA) on all datasets, the reason is that IFW_FDAR_AA takes extra time to implement the classifier in the wrapper phase. This is also a common drawback of algorithms that follow the filter-wrapper approach. However, with the aim of minimizing the complexity and increasing the accuracy of the classification rule set, the cost of time to find the reduct of the proposed algorithm is acceptable.

Table 3. The execution time of IFW_FDAR_AA and FRSA-IFS-HIS(AA) (s).

ID	**Data set**	**Attribute set**	**The number of attributes**	**The total number of attribute**	**IFW_FDAR_AA**		**FRSA-IFS-HIS(AA)**	
					Time	**Total time**	**Time**	**Total time**
1	WDBC	C_0	15	15	2.92	2.92	2.16	2.16

		C_1	3	18	0.33	3.25	0.28	2.44
		C_2	3	21	0.34	3.59	0.32	2.76
		C_3	3	24	0.22	3.81	0.20	2.96
		C_4	3	27	0.21	4.02	0.18	3.14
		C_5	3	30	0.24	4.26	0.16	**3.30**
2	Credit	C_0	5	5	2.05	2.05	1.74	1.74
		C_1	2	7	0.24	2.29	0.18	1.92
		C_2	2	9	0.29	2.58	0.22	2.14
		C_3	2	11	0.26	2.84	0.21	2.35
		C_4	2	13	0.28	3.12	0.20	2.55
		C_5	2	15	0.22	3.34	0.18	**2.73**
3	German	C_0	10	10	3.08	3.08	2.64	2.64
		C_1	2	12	0.21	3.29	0.17	2.81
		C_2	2	14	0.30	3.59	0.17	2.98
		C_3	2	16	0.32	3.91	0.21	3.19
		C_4	2	18	0.38	4.29	0.24	3.43
		C_5	2	20	0.35	4.64	0.26	**3.69**
4	Wave	C_0	11	11	64.56	64.56	56.02	56.02
		C_1	2	13	8.00	72.56	6.8	62.82
		C_2	2	15	6.52	79.08	5.62	68.44
		C_3	2	17	7.17	86.25	6.08	74.52
		C_4	2	19	5.79	92.04	4.94	79.46
		C_5	2	21	6.68	98.72	5.18	**84.64**

6. CONCLUSIONS

In the current trend of big data, decision systems are increasingly large in size and always changing, updating typically in the case of changing attribute sets. On such decision systems, the construction of algorithms to improve the efficiency of finding reduct to improve the efficiency of the classification rule set has high practical significance. Up to now, the incremental algorithms for finding reduct in case of adding and deleting attribute set are filter algorithms have followed the traditional filter approach. Therefore, the reduct of these algorithms is not optimal in terms of the number of attributes and classification accuracy. In this paper, we build up the distance update formulas in the paper [15] with the addition and deletion case of the attribute set. For using in-built formulas, we hereby propose two filter-wrapper incremental algorithms that find the reduct of the decision system: algorithm IFW_FDAR_AA in case of adding an attribute set and algorithm IFW_FDAR_DA in case of deleting attribute set. Test results on data sets from UCI data warehouse [20] showed that the number of attributes in reduct of algorithm IFW_FDAR_AA is much smaller than algorithm FRSA-IFS-HIS (AA) [18]. Furthermore, the classification accuracy of IFW_FDAR_AA is higher than that of FRSA-IFS-HIS (AA). Therefore, the classification rule set on the reduct IFW_FDAR_AA is more efficient than FRSA-IFS-HIS (AA). However, the execution time of the IFW_FDAR_AA algorithm is higher than that of FRSA-IFS-HIS (AA), the reason is that IFW_FDAR_AA takes extra time to implement the classifier in the wrapper phase. The next research direction is to continue to improve the filter-wrapper incremental algorithms to shorter the execution time with incremental computation solution when running the classifier.

Acknowledgements. This research is funded by the project NVCC02.04/21-21 “The task of providing funding for scientific research activities to senior researchers in 2021, NVCC02.04/21-21”.

CRediT authorship contribution statement. Nguyen Long Giang: Concept, Methodology, Validation, Writing - review and editing. Ho Thi Phuong: Software, Data curation, Writing - original draft preparation. All authors have read and agreed to the published version of the manuscript

Declaration of competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. Pham Dinh Phong - An Application of Feature Selction for the Fuzzy Rule Based Classifier Design Based on an Enlarged Hedge Algebras for High-Dimention Datasets, Vietnam J. Sci. Technol. **53** (5) (2015) 583-597.
2. Pham Thi Thiet - Applying the Attributed Prefix Tree for Mining Closed Sequential Patterns, Vietnam J. Sci. Technol. **54** (3A) (2016) 106-114.
3. Vu Duc Thi - Some Problems Related to Database and Data Mining, Vietnam J. Sci. Technol. **50** (6) (2012) 679-704 (in Vietnamese).
4. Dübois D., Prade H. - Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems **17** (1990) 191-209.
5. Anoop Kumar Tiwari, Shivam Shreevastava, Tanmoy Som, K. K. Shukla. - Tolerance-based intuitionistic fuzzy-rough set approach for attribute reduction, Expert Systems With Applications **101** (2018) 205-212.
6. Zhang C. L., Mei D. G., Chen Y. Y., Yang Y. - A fuzzy rough set-based feature selection method using representative instances, Knowledge-Based Systems **151** (2018) 216-229.
7. Sheeja T. K., Sunny Kuriakose A. - A novel feature selection method using fuzzy rough sets, Computers in Industry **97** (2018) 111- 116.
8. Lin Y., Li Y., Wang C., Chen J. - Attribute reduction for multi-label learning with fuzzy rough set, Knowl.-Based Syst. **152** (2018) 51-61.
9. Dai J. H., Yan Y. J., Li Z. W., Liao B. S. - Dominance-based fuzzy rough set approach for incomplete interval-valued data, J. of Intelligent & Fuzzy Systems **34** (2018) 423-436.
10. Wang C. Z., Huang Y., Shao M. W., Fan X. D. - Fuzzy rough setbased attribute reduction using distance measures, Knowledge-Based Systems **164** (2019) 205-212.
11. Cao Chinh Nghia, Demetrovics Janos, Nguyen Long Giang, Vu Duc Thi - About a fuzzy distance between two fuzzy partitions and attribute reduction problem, Cybernetics and Information Technologies **16** (4) (2016) 13-28.
12. Liu Y. M., Zhao S. Y., Chen H., Li C.P., Lu Y. M. - Fuzzy Rough Incremental Attribute Reduction Applying Dependency Measures, APWeb-WAIM 2017: Web and Big Data (2017) 484-492.

13. Yang Y. Y., Chen D. G., Wang H., Eric C. C. Tsang, Zhang D. L. - Fuzzy rough set based incremental attribute reduction from dynamic data with sample arriving, Fuzzy Sets and Systems **312** (2017) 66-86.
14. Yang Y. Y., Chen D. G., Wang H., Wang X. H. - Incremental perspective for feature selection based on fuzzy rough sets, IEEE Transactions on Fuzzy Systems **26** (3) (2017) 1257-1273.
15. Giang N. L., Ngan T. T., Tuan T. M., Phuong H. T., Abdel-Basset M., Macêdo A. R. L., Albuquerque V. - Novel Incremental Algorithms for Attribute Reduction from Dynamic Decision systems using Hybrid Filter-Wrapper with Fuzzy Partition Distance, IEEE Transactions on Fuzzy Systems **28** (5) (2020) 858-873.
16. Zhang X., Mei C. L., Chen D. G., Yang Y. Y., Li J. H. - Active Incremental Feature Selection Using a Fuzzy-Rough-Set-Based Information Entropy, IEEE Transactions on Fuzzy Systems **28** (5) (2020) 901-915.
17. Ni P., Zhao S. Y., Wang X. H., Chen H., Li C. P., Tsang E. C. C. - Incremental Feature Selection Based on Fuzzy Rough Sets, Information Sciences **536** (2020) 185-204.
18. Zeng A. P., Li T. R., Liu D., Zhang J. B., Chen H. M. - A fuzzy rough set approach for incremental feature selection on hybrid information systems, Fuzzy Sets and Systems **258** (2015) 39-60.
19. Nguyen Ba Quang, Nguyen Long Giang, Dang Thi Oanh - A distance based Incremental Filter-wrapper Algorithm for Finding Reduct in Incomplete Decision Tables, Vietnam J. Sci. Technol. **57** (4) (2019) 499-512.
20. The UCI machine learning repository, http://archive.ics.uci.edu/ml/datasets.html.