

Incremental algorithms based on metric for finding reduct in dynamic decision tables

Nguyen Thi Lan Huong, Nguyen Long Giang

Email: bmtointin@hvcnsd.edu.vn, nlgiang@ioit.ac.vn

Abstract: Feature selection is a crucial problem need to be effectively solved in knowledge discovery in databases because of two basic reasons: to minimize cost and to accurately classify data. Feature selection using rough set theory also called attribute reduction have attracted much attention from researchers and many results are gained. However, attribute reduction in dynamic databases is still in the first stage. This paper focus on develop incremental methods and algorithms to derive reducts hiring a distance measure when adding, deleting or updating objects. Since not re-implement the algorithms on the varied universal set, our algorithms significantly reduce the complexity of implementation time.

Keyword: Rough set, decision tables, attribute reduction, reduct, metric.

1. INTRODUCTION

Feature selection is a common technique in the data mining process of knowledge discovery in databases (KDD). Rough set theory by professor Zdzislaw Pawlak in 1982 right after its start-up has been recognized as a mathematical tool for data analysis. In the field of KDD, rough sets have numbers of applications such as discretization, feature selection, data reduction. Feature selection or attribute reduction means to determine an appropriate subset of attributes that can preserve the ability to classify the set of objects. The subset is called a reduct.

In recent years, there are rapid growth of interests to attribute reduction based on rough set. Many advanced techniques have been proposed and the common points of those methods are to give a definition of reduct and a heuristic algorithm for finding the best reduct with criteria is classification quality of features (also referred to as feature significance). In the works [4, 11], the authors

summarized and categorized feature reduction methods in decision tables into three groups: (1) Positive region method (including attribute reduction methods based on positive region); (2) Shannon entropy method (including the method using Shannon's entropy and method using relational algebra); (3) Liang entropy method (including the methods using Liang entropy, methods using information entropy and methods using discernibility matrix). In the context of static database, those techniques are quite rational. However, the number of researches on dynamic databases is still modest.

In the real world, data usually is updated and changed with time. It can vary from adding or deleting objects or features to updating existing objects. As a natural solution, users can carry repeatedly the exist algorithms to finding new reducts. As a result, the time spent for re-computation is quite huge, especially when the number of objects being added, deleted or updated is large.

Recent researches on incremental methods to find reducts on dynamic decision tables have employed different measures. In [2, 3, 12], authors used positive region and discernibility matrix when adding new objects. In [6, 9], authors constructed formulas for three entropies (Shannon's entropy, Liang entropy and combination entropy) when adding or deleting objects. In [10], using entropy method, an attribute reduction algorithm is proposed for decision tables in case an object is updated. However, proposed formulas are quite complex and the methods mentioned above have not yet completely dealt with dynamic decision tables. To deal with the variation of the decision tables and complex formulas, we use distance measure – a major classic approach to understand the different

between two objects in many fields, such as algebra, geometry, set theory, ...

Some distance measures were hiring to construct algorithms for finding reduct such as in [7], author used Jaccard distance. In [16], authors used a variant of Jaccard distance to work out the reduct and proved that their methods belong to the group of Liang entropy methods. However, these reasearches are not completely different from entropy methods - in [16] authors proved that the metric used in their algorithms is equal tothe Liang conditional entropy [5] – and for the case of updating objects, these methods are not yet to deal with.

Here in this paper, we employ a different distance measure and gain advantages of better complexity as well as simplicity and understandability of the formulas in the algorithms. Even better, we prove that our method is in Shanon’s entropy group and as in [4] the authors pointed out that reducts finding in this group are better in terms of the number of attributes than the reducts finding by Liang entropy methods.

Using the metricsuggested by authors in [7], we first work out the incremental formulas to derive the metric from the original one. Then we propose two incremental algorithms for finding reducts of decision tables in cases of adding, deleting and updating objects. Our algorithms are theoretically proved better in terms of computing complexity in comparison to the case of re-implement the algorithms for the varied object set. More over, the formulas using in algorithms are not too complex as they are in other techniques.

This paper is organized as follows. Basic concepts of rough set theory and attribute reduction are recalled in section 2. Section 3 presents the attribute reduction method based on distance measure in [7]. In section 4, we work out the incremental formulas for the new metric after one object is added, deleted from the object set. Also in the section, base on the metric we propose algorithms for finding reduct. Similarly, in section 5 we proposed the mechanism for update the metric in case one objetc is updated and the algorithm for finding the reduct. The conclusion of the paper and further reasearches are presented in section 6.

2. PRELIMINARY KNOWLEDGE ON ROUGH SETS AND ATTRIBUTE REDUCTION

In this section we summary some basic concepts in rough set theory [8] and an overview of rough set based methods for attribute reduction in decision tables.

An information system is a couple $IS = (U, A)$ where U is a finite nonempty set of objects and A is a finite nonempty set of features. Each $a \in A$ determines a map $a : U \rightarrow V_a$ where V_a is the value set of a .

Given an information system $IS = (U, A)$, each $P \subseteq A$ determines an equivalence relation:

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

$IND(P)$ defines a partition of U . Let $U/IND(P)$ or briefly U/P denotes a family of all equivalence classes of the relation $IND(P)$ and $[u]_P = \{v \in U \mid (u, v) \in IND(P)\}$. Then, $U/P = \{[u]_P \mid u \in U\}$.

Given an information system $IS = (U, A)$, $B \subseteq A$ and $X \subseteq U$. $\underline{B}X = \{u \in U \mid [u]_B \subseteq X\}$ and

$\overline{B}X = \{u \in U \mid [u]_B \cap X \neq \emptyset\}$ are respectively called lower approximation and upper approximation of X respect to B .

A decision table is an information system (U, A) , where A includes two separate subsets: condition attribute subset C and decision attribute subset D . So that, a decision table $DS = (U, C \cup D)$ where $C \cap D = \emptyset$.

Let $DS = (U, C \cup D)$ be a decision table, denote D_i as an equivalence class of partition U/D , then $U/D = \{D_i\}$. $POS_C(D) = \bigcup_{D_i \in U/D} (\underline{C}D_i)$ is called C-Positive

region of D . One can easily obtain that $POS_C(D)$ is a set of objects belonging to U which are partitioned by $IND(C)$ right into decision classes of U/D . A decision table DS is consistent if and only if $POS_C(D) = U$; other wise, it is inconsistent.

In an information system, not all the features are necessary for recognition or classification. The concept of attribute reduction was first introduced by Z.Pawlak in [8], [15] which target to eliminate irrelevant or redundant features in such that the discernible ability of the attribute set is preserved. The remain feature set after elimination is called a reduct.

Given an information system $IS = (U, A)$, $B \subseteq A$ is a reduct of IS if it satisfies two following conditions:

- 1) $U/B = U/A$
- 2) $\forall a \in B, U/(B - \{a\}) \neq U/B$

There are usually more than one reduct in an information system, the intersection of all the reducts is called the core.

In the next section, we briefly present the attribute reduction method based on the metric in [7]. The method is proved that it belongs to Shannon entropy method group.

3. METRIC BASED ATTRIBUTE REDUCTON

Given a decision table $DS = (U, C \cup D)$, for each $P \subseteq C$, $K(P) = \{[u_i]_P | u_i \in U\}$ is called a knowledge of P on U [8]. $K(P)$ includes $|U|$ elements where each one is a class in U/P , also referred as a knowledge granule. Let denote family of all knowledges on U by $\mathbf{K}(U)$ and $U/C = \{C_1, C_2, \dots, C_m\}$, $U/D = \{D_1, D_2, \dots, D_n\}$. For any $K(P), K(Q) \in \mathbf{K}(U)$, the metric between $K(P), K(Q) \in \mathbf{K}(U)$ is defined as folow

Definition 3.1

$$d(K(C), K(C \cup D)) = 1 - \sum_{i=1}^n \sum_{j=1}^m \frac{|D_i \cap C_j|^2}{|U||C_j|}$$

Based on the metric, the author in [7] proposed a metric based attribute reduction method. The method consists of steps: Define a reduct based on the distance; Definethesignificance of an attribute based on the distance and constructthe attribute reduction algorithm to find the best reduct.

Definition 3.2. Given a decision table $DS = (U, C \cup D)$, $R \subseteq C$. If:

- 1) $d(K(R), K(R \cup D)) = d(K(C), K(C \cup D))$;
- 2) $\forall r \in R, d(K(R - \{r\}), K(R - \{r\} \cup D)) \neq d(K(C), K(C \cup D))$;

then R is a reduct of C based on metric. This reduct is proved equivalence to the reduct based on Shannon entropy.

Definition 3.3. Given a decision table $DS = (U, C \cup D)$, $B \subseteq C$ and $b \in C - B$. The significance of b is defined by

$$SIG_B(b) = d(K(B), K(B \cup D)) - d(K(B \cup \{b\}), K(B \cup \{b\} \cup D))$$

It is easy to see that $SIG_B(b) \geq 0$ and $SIG_B(b)$ is caculated by the changed amount of the distance between B and $B \cup \{b\}$ and the greater $SIG_B(b)$ is, the greater changed amount is or the more significant b is and via versa. This significance is the attribute selection criteria of the heuristic algorithm for finding reducts of decision tables.

Algorithm 3.1. Heuristic algorithm for finding a reduct based on metric

Input: Decision table $DS = (U, C \cup D)$.

Output: The best reduct R .

//Finding the core set $CORE(C)$;

1. $CORE(C) = \emptyset$;
 2. For $c \in C$
 3. If $d(K(C - \{c\}), K(C - \{c\} \cup D)) \neq d(K(C), K(C \cup D))$
- then $CORE(C) := CORE(C) \cup \{c\}$;

//Finding the reduct R based on metric

4. $R = CORE(C)$;
5. While $d(K(R), K(R \cup D)) \neq d(K(C), K(C \cup D))$ do
6. Begin

7. For $a \in C - R$ calculate $SIG_R(a)$;

8. Select $a_m \in C - R$ such that

$$SIG_R(a_m) = \text{Max}_{a \in C-R} \{SIG_R(a)\};$$

9. $R = R \cup \{a_m\}$;

10. End;

11. Return R ;

Given a decision table $DS = (U, C \cup \{d\})$ (suppose that decision set D includes only one element $D = \{d\}$), according to [13], the time complexity (hereinafter referred to as complexity) for getting condition partition U / C is $O(|U||C|)$, so that the complexity for computing the metric $d(K(C), K(C \cup \{d\}))$ is

$$O\left(|U||C| + |U| + \sum_{i=1}^n D_i \sum_{j=1}^m C_j\right) = O(|U||C| + |U|^2)$$
 and the

complexity for computing the core set $CORE(C)$ from instructions 1 to 3 of algorithm 3.1 is $O(|C|(|U||C| + |U|^2)) = O(|C|^2|U| + |C||U|^2)$. The

complexity for computing the reduct from instructions 4 to 9 is $O(|C|^2|U| + |C||U|^2)$. Hence, the

complexity of algorithm 3.1 is $O(|C|^2|U| + |C||U|^2)$.

In case of dynamic decision tables, algorithm 3.1 could be re-implemented to get reducts. However, here in section 4 and section 5 we propose incremental algorithms which use calculated metric before the decision table be changed to work out the new reducts. We prove in theory that our algorithms are better than re-implement algorithm 3.1 in complexity for computing.

4. PROPOSED ALGORITHMS FOR FINDING THE REDUCT BASED ON METRIC WHEN ADDING OR DETETING ONE OBJECT

The main points of our algorithms are to derive the metric after object set varied from the original one without re-calculating it. At first, we consider the case that only one object is added or deleted. If

multi objects are varied we can repeatedly implement the algorithms.

4.1 Incremental formula for metric when adding one object

Given a decision table $DS = (U, C \cup D)$, $B \subseteq C$. Denote $U / B = \{X_1, X_2, \dots, X_m\}$ and $U / D = \{Y_1, Y_2, \dots, Y_n\}$. The metric between two knowledges $K(B)$ and $K(B \cup D)$ on U is $d(K(B), K(B \cup D))$.

Proposition 4.1. *Suppose that object x is added to U , the metric between $K(B)$ and $K(B \cup D)$ on U is $d_U(K(B), K(B \cup D))$. Then one can obtain:*

1) *If $x \notin X_i$ for any $i = 1..m$ and $x \notin Y_j$ for any $j = 1..n$ then*

$$d_{U \cup \{x\}}(K(B), K(B \cup D)) = \frac{1}{|U| + 1} \left(1 - |U| \cdot d_U(K(B), K(B \cup D))\right)$$

2) *If $x \notin X_i$ for any $i = 1..m$ and $x \in Y_q$ for $q \leq n$ then*

$$d_{U \cup \{x\}}(K(B), K(B \cup D)) = \frac{1}{|U| + 1} \left(|U| \cdot d_U(K(B), K(B \cup D))\right)$$

3) *If $x \in X_p$ for $p \leq m$ and $x \notin Y_j$ for any $j = 1..n$ then*

$$d_{U \cup \{x\}}(K(B), K(B \cup D)) = \frac{1}{|U| + 1} \left(1 - \frac{1}{|X_p|} + |U| \cdot d_U(K(B), K(B \cup D))\right)$$

4) *If $x \in X_p$ for $p \leq m$ and $x \in Y_q$ for $q \leq n$ then*

$$\begin{aligned}
 d_{U \cup \{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U| + 1} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &= \frac{1}{|U| + 1} \left(|U| \cdot d_U(K(B), K(B \cup D)) - \frac{2|X_p \cap Y_q| + 1}{|X_p| + 1} \right. \\
 &\quad \left. + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{(|X_p| + 1)|X_p|} + 1 \right) \quad \left. + \sum_{j=1}^n \frac{|X_{m+1} \cap Y_j|^2}{|X_{m+1}|} \right)
 \end{aligned}$$

Proof.

1) Suppose that $X_{m+1} = \{x\}$ and $Y_{n+1} = \{x\}$.
We have

$$\begin{aligned}
 d_{U \cup \{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U| + 1} \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} \frac{|X_i \cap Y_j|^2}{|X_i|}
 \end{aligned}$$

$$\begin{aligned}
 &= 1 - \frac{1}{|U| + 1} \sum_{i=1}^{m+1} \left(\sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &\quad \left. + \frac{|X_i \cap Y_{n+1}|^2}{|X_i|} \right)
 \end{aligned}$$

$$\begin{aligned}
 &= 1 - \frac{1}{|U| + 1} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &\quad \left. + \sum_{j=1}^n \frac{|X_{m+1} \cap Y_j|^2}{|X_{m+1}|} \right)
 \end{aligned}$$

$$= \frac{1}{|U| + 1} (1 - |U| \cdot d_U(K(B), K(B \cup D)))$$

2) Suppose that $X_{m+1} = \{x\}$ and $x \in Y_q$ for $q \leq n$.
We have:

$$\begin{aligned}
 d_{U \cup \{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U| + 1} \sum_{i=1}^{m+1} \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|}
 \end{aligned}$$

$$= \frac{1}{|U| + 1} (|U| \cdot d_U(K(B), K(B \cup D)))$$

3) Suppose that $x \in X_p$ for $p \leq m$ and $Y_{n+1} = \{x\}$. We have:

$$\begin{aligned}
 d_{U \cup \{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U| + 1} \sum_{i=1}^m \sum_{j=1}^{n+1} \frac{|X_i \cap Y_j|^2}{|X_i|}
 \end{aligned}$$

$$\begin{aligned}
 &= 1 - \frac{1}{|U| + 1} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &\quad \left. + \sum_{i=1}^m \frac{|X_i \cap Y_{n+1}|^2}{|X_i|} \right)
 \end{aligned}$$

$$= 1 - \frac{1}{|U| + 1} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \frac{1}{|X_p|} \right)$$

$$= \frac{1}{|U| + 1} \left(1 - \frac{1}{|X_p|} + |U| \cdot d_U(K(B), K(B \cup D)) \right)$$

4) Suppose that $x \in X_p$ for $p \leq m$ and $x \in Y_q$ for $q \leq n$. We have:

$$\begin{aligned}
 d_{U \cup \{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U| + 1} \sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \\
 &= 1 -
 \end{aligned}$$

$$\frac{1}{|U|+1} \left(\sum_{\substack{i=1 \\ i \neq p}}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|(X_p \cup \{x\}) \cap Y_j|^2}{|X_p|+1} + \frac{|(X_p \cup \{x\}) \cap (Y_q \cup \{x\})|^2}{|X_p|+1} \right)$$

$$= 1 - \frac{1}{|U|+1} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \frac{2|X_p \cap Y_q|+1}{|X_p|+1} - \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p|+1)} \right)$$

$$= \frac{1}{|U|+1} \left(|U| \cdot d_U(K(B), K(B \cup D)) - \frac{2|X_p \cap Y_q|+1}{|X_p|+1} + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{(|X_p|+1)|X_p|} + 1 \right)$$

4.2. Incremental formula for metric when deleting one object

Proposition 4.2. Let $x \in U$ is the object to be deleted from U . Suppose that $x \in X_p$ for $p \leq m$ and $x \in Y_q$ for $q \leq n$. We denote $d_U(K(B), K(B \cup D))$ as the metric between knowledges $K(B)$ and $K(B \cup D)$ on U . One can obtain:

$$\begin{aligned} d_{U-\{x\}}(K(B), K(B \cup D)) &= \frac{1}{|U|-1} (|U|d_U(K(B), K(B \cup D)) - \Delta) \end{aligned}$$

$$\text{Where } \Delta = 1 + \frac{1-2|X_p \cap Y_q|}{|X_p|-1} + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p|-1)}$$

Proof. We have

$$\begin{aligned} d_{U-\{x\}}(K(B), K(B \cup D)) &= 1 - \frac{1}{|U|-1} \left(\sum_{\substack{i=1 \\ i \neq p}}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} + \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p|-1} + \frac{|(X_p - \{x\}) \cap (Y_q - \{x\})|^2}{|X_p|-1} \right) \end{aligned}$$

$$= 1 - \frac{1}{|U|-1} \left(\sum_{\substack{i=1 \\ i \neq p}}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap Y_q|^2}{|X_i|} + \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_p \cap Y_j|^2}{|X_p|-1} + \frac{|X_p \cap Y_q|^2}{|X_p|-1} + \frac{(|X_p \cap Y_q|-1)^2}{|X_p|-1} \right)$$

$$= \frac{1}{|U|-1} \left(|U|d_U(K(B), K(B \cup D)) - 1 - \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p|-1)} - \frac{1-2|X_p \cap Y_q|}{|X_p|-1} \right)$$

$$= \frac{1}{|U| - 1} (|U| d_U(K(B), K(B \cup D)) - \Delta)$$

$$\text{Where } \Delta = 1 + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)}$$

4.3. Algorithms for finding reduct when adding or deleting one object

At first, we construct an incremental algorithm for finding the reduct after adding a new object. Here proposition 4.3 is the fundamentals for constructing the algorithm.

Proposition 4.3. Given a decision table $DS = (U, C \cup D)$, $B \subseteq C$ is a reduct of DS based on metric and x is the new one added to U . Let $U/B = \{X_1, X_2, \dots, X_m\}$. Then one can obtain:

- 1) If $x \notin X_j$ for any $j = 1..m$ then $d_{U \cup \{x\}}(K(B), K(B \cup D)) = d_{U \cup \{x\}}(K(C), K(C \cup D))$
- 2) If $x \in X_p$ for $p \leq m$ then

$$d_{U \cup \{x\}}(K(B), K(B \cup D)) \neq d_{U \cup \{x\}}(K(C), K(C \cup D))$$

Proof. Proposition 4.3 can be directly drawn from proposition 4.1 and definition 3.1 of a reduct based on metric.

Proposition 4.3 shows that if x does not belong to any equivalence class in U/B then the metric $d_{U \cup \{x\}}(K(B), K(B \cup D)) = d_{U \cup \{x\}}(K(C), K(C \cup D))$, i.e the reduct is unchanged. From that, we construct the incremental algorithm for finding reduct as follow:

Algorithm 4.1. The incremental algorithm for finding reduct based on metric when adding a new object.

Input: Decision table $DS = (U, C \cup D)$, reduct R_U on U and new object x .

Output: Reduct $R_{U \cup \{x\}}$ on $U \cup \{x\}$.

1. Assign $R = R_U$, calculate

$$U/R = \{X_1, X_2, \dots, X_m\};$$

2. If $x \in X_p$, $X_p \in U/R$ then
3. Begin
4. While $d_{U \cup \{x\}}(K(R), K(R \cup D)) \neq d_{U \cup \{x\}}(K(C), K(C \cup D))$ do
5. Begin
6. For $a \in C - R$ calculate $SIG_R(a)$ according to the incremental formular in Proposition 4.1
7. Select $a_m \in C - R$ such that $SIG_R(a_m) = \text{Max}_{a \in C - R} \{SIG_R(a)\};$
8. $R = R \cup \{a_m\};$
9. End;
10. End;
11. For $a \in R$ do
12. Begin
13. Calculate $d_{U \cup \{x\}}(K(R - \{a\}), K((R - \{a\}) \cup D));$
14. If $d_{U \cup \{x\}}(K(R - \{a\}), K((R - \{a\}) \cup D)) = d_{U \cup \{x\}}(K(R), K(R \cup D))$ then $R = R - \{a\};$
15. End
16. Return R .

Similar to the case of adding one new object, the fundamentals for constructing an algorithm for finding reducts when deleting one object is proposition 4.4.

Proposition 4.4. Given a decision table $DS = (U, C \cup D)$, $B \subseteq C$ is a reduct of DS based on metric and $x \in U$. Let $U/B = \{X_1, X_2, \dots, X_m\}$, then we have:

$$d_{U - \{x\}}(K(B), K(B \cup D)) \neq d_{U - \{x\}}(K(C), K(C \cup D))$$

It is easy to see that the algorithm for

finding reducts in this case is worked out as the same way to algorithm 4.1.

4.4. Advantages of proposed algorithms

According to [13], the complexity for calculating partition U/C is $O(|C||U|)$, so that the complexity of the incremental formula for calculating metric in proposition 4.1 is $O(|C||U| + m|C| + |U| + |X_p||Y_q|) = O(|C||U| + |X_p||Y_q|)$. The complexity of the loop While from 7 to 12 is $O(|C|(|C||U| + |X_p||Y_q|))$. The complexity of the loop For from 15 to 19 is $O(|C|(|C||U| + |X_p||Y_q|))$. So that, the complexity of algorithm 4.1 is $O(|C|^2|U| + |C||X_p||Y_q|)$. Obviously, $|X_p||Y_q|$ is much less than $|U|^2$ so we can say the complexity of algorithm 4.1 is much less than algorithm 3.1 (the original one).

5. ALGORITHM FOR FINDING REDUCT BASED ON METRIC WHEN UPDATING ONE OBJECT

5.1 Incremental formula for metric when updating one object.

Given a decision table $DS = (U, C \cup D)$, $B \subseteq C$, $U/B = \{X_1, X_2, \dots, X_m\}$ and $U/D = \{Y_1, Y_2, \dots, Y_n\}$. The metric between two knowledges $K(B)$ and $K(B \cup D)$ on U is $d_U(K(B), K(B \cup D))$.

Proposition 5.1. Suppose that object $x \in U$ is updated to x' and $x \in X_p$, $x \in Y_q$ ($p \leq m$; $q \leq n$). Name the metric between knowledges $K(B)$ and $K(B \cup D)$ on $\{U - \{x\}\} \cup \{x'\}$ is $d_N(K(B), K(B \cup D))$. Then the new metric can be obtained from $d_U(K(B), K(B \cup D))$ as following:

1) If $x' \notin X_i$, $\forall i = 1, \dots, m$ and $x' \notin Y_j$, $\forall j = 1, \dots, n$ then,

$$\begin{aligned} d_N(K(B), K(B \cup D)) &= d_U(K(B), K(B \cup D)) \\ &+ \frac{1}{|U|} \left(\frac{2|X_p \cap Y_q| + 1}{|X_p| - 1} \right. \\ &\left. + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(1 - |X_p|)} - 1 \right) \end{aligned}$$

2) If $x' \in X_r$, $r \leq m$ and $x' \notin Y_j$, $\forall j = 1, \dots, n$ then,

$$\begin{aligned} d_N(K(B), K(B \cup D)) &= d_U(K(B), K(B \cup D)) \\ &+ \frac{1}{|U|} \left(\frac{1}{|X_r| + 1} \right. \\ &+ \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \left. \right) \\ &+ \frac{1}{|U|} \sum_{j=1}^n \left(\frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \right. \\ &\left. - \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} \right) \end{aligned}$$

3) If $x' \notin X_i$, $\forall i = 1, \dots, m$ and $x' \in Y_s$, $s \leq n$ then,

$$\begin{aligned} d_N(K(B), K(B \cup D)) &= d_U(K(B), K(B \cup D)) \\ &+ \frac{1}{|U|} \left(1 + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \right. \\ &\left. + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \right) \end{aligned}$$

4) If $x' \in X_r$, $r \leq m$ and $x' \in Y_s$, $s \leq n$ then,

$$\begin{aligned}
 d_N(K(B), K(B \cup D)) &= d_U(K(B), K(B \cup D)) \\
 &+ \frac{1}{|U|} \left(\frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \right. \\
 &+ \frac{1 + 2|X_r \cap Y_s|}{|X_r| + 1} \\
 &+ \sum_{j=1}^n \left(\frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \right. \\
 &\left. \left. - \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} \right) \right)
 \end{aligned}$$

$$\begin{aligned}
 1 - \frac{1}{|U|} &\left(\sum_{\substack{i=1 \\ i \neq p}}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &+ \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} \\
 &+ \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap \{x'\}|^2}{|X_i|} \\
 &+ \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p| - 1} \\
 &+ \frac{|(X_p - \{x\}) \cap (Y_q - \{x\})|^2}{|X_p| - 1} \\
 &+ \frac{|(X_p - \{x\}) \cap \{x'\}|^2}{|X_p| - 1} \\
 &+ \frac{|x' \cap (Y_q - \{x\})|^2}{|\{x'\}|} \\
 &\left. + \frac{|x' \cap \{x'\}|^2}{|\{x'\}|} \right)
 \end{aligned}$$

Proof.

1) Since $x' \notin X_i, \forall i = 1, \dots, m$ and $x' \notin Y_j, \forall j = 1, \dots, n$ partitions of updated U on B and D can be presented as follows:

$$(\{U - \{x\}\} \cup \{x'\})/B = \{X_1, X_2, \dots, X_p - \{x\}, X_{p+1}, \dots, X_m, X_{m+1} = \{x'\}\}$$

$$(\{U - \{x\}\} \cup \{x'\})/D = \{Y_1, Y_2, \dots, Y_q - \{x\}, Y_{q+1}, \dots, Y_n, Y_{n+1} = \{x'\}\}$$

Then we have

$$\begin{aligned}
 d_N(K(B), K(B \cup D)) &= 1 \\
 &- \frac{1}{|U|} \left(\sum_{i=1}^{m+1} \sum_{j=1}^{n+1} \frac{|X_i \cap Y_j|^2}{|X_i|} \right) =
 \end{aligned}$$

Because for any $i = 1 \dots, m$ and $i \neq p$ we have $x \notin X_i$, so $\sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} = \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap Y_q|^2}{|X_i|}$. For any $i = 1, \dots, m$, since $x' \notin X_i$ then $\sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap \{x'\}|^2}{|X_i|} = 0$ and $\frac{|(X_p - \{x\}) \cap \{x'\}|^2}{|X_p| - 1} = 0$. For any $j = 1, \dots, n$, since $x' \notin Y_j$ then $\frac{|x' \cap (Y_q - \{x\})|^2}{|\{x'\}|} = 0$

Note that for any $j = 1, \dots, n$ and $j \neq q$ since $x \notin Y_j$ so $\{x\} \cap Y_j = \emptyset$.

Consequently, we have

$$\begin{aligned}
 d_N(K(B), K(B \cup D)) &= 1 \\
 &- \frac{1}{|U|} \left(\sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &+ \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap Y_q|^2}{|X_i|} \\
 &+ \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_p \cap Y_j|^2}{|X_p| - 1} \\
 &\left. + \frac{(|X_p \cap Y_q| - 1)^2}{|X_p| - 1} + 1 \right) \\
 &= d_U(K(B), K(B \cup D)) \\
 &+ \frac{1}{|U|} \left(\frac{2|X_p \cap Y_q| + 1}{|X_p| - 1} \right. \\
 &+ \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(1 - |X_p|)} - 1 \left. \right) \\
 &= 1 - \frac{1}{|U|} \left(\sum_{\substack{i=1 \\ i \neq p; i \neq r}}^m \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\
 &+ \sum_{\substack{i=1 \\ i \neq p; i \neq r}}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} \\
 &+ \sum_{\substack{i=1 \\ i \neq p; i \neq r}}^m \frac{|X_i \cap \{x'\}|^2}{|X_i|} \\
 &+ \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p| - 1} \\
 &+ \frac{|(X_p - \{x\}) \cap (Y_q - \{x\})|^2}{|X_p| - 1} \\
 &+ \frac{|(X_p - \{x\}) \cap \{x'\}|^2}{|X_p| - 1} \\
 &+ \sum_{\substack{j=1 \\ j \neq q}}^n \frac{|(X_r \cup \{x'\}) \cap Y_j|^2}{|X_r \cup \{x'\}|} \\
 &+ \frac{|(X_r \cup \{x'\}) \cap (Y_q - \{x\})|^2}{|X_r \cup \{x'\}|} \\
 &\left. + \frac{|(X_r \cup \{x'\}) \cap \{x'\}|^2}{|X_r \cup \{x'\}|} \right)
 \end{aligned}$$

2) If $x' \in X_r$, $r \leq m$ and $x' \notin Y_j$, $\forall j = 1, \dots, n$.
Partitions of updated U on B and D can be presented as follows:

$$(\{U - \{x\}\} \cup \{x'\})/B = \{X_1, X_2, \dots, X_p \setminus \{x\}, X_{p+1}, \dots, X_r \cup \{x'\}, \dots, X_m\}$$

$$(\{U \setminus \{x\}\} \cup \{x'\})/D = \{Y_1, Y_2, \dots, Y_q \setminus \{x\}, Y_{q+1}, \dots, Y_n, Y_{n+1} = \{x'\}\}$$

Then we have

$$\begin{aligned}
 d_N(K(B), K(B \cup D)) &= 1 \\
 &- \frac{1}{|U|} \left(\sum_{i=1}^m \sum_{j=1}^{n+1} \frac{|X_i \cap Y_j|^2}{|X_i|} \right)
 \end{aligned}$$

$$= 1 - \frac{1}{|U|} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{(|X_p| - 1)|X_p|} - \sum_{j=1}^n \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} + \frac{1}{|X_r| + 1} \right)$$

$$= d_U(K(B), K(B \cup D)) + \frac{1}{|U|} \left(\frac{1}{|X_r| + 1} + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \right) + \frac{1}{|U|} \sum_{j=1}^n \left(\frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} - \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} \right)$$

3) If $x' \notin X_i, \forall i = 1, \dots, m$ and $x' \in Y_s, s \leq n$. We present the new partitions as follows:

$$(\{U - \{x\}\} \cup \{x'\})/B = \{X_1, X_2, \dots, X_p \setminus \{x\}, X_{p+1}, \dots, X_m, \{x'\}\}$$

$$(\{U - \{x\}\} \cup \{x'\})/D = \{Y_1, Y_2, \dots, Y_q \setminus \{x\}, Y_{q+1}, \dots, Y_s \cup \{x'\}, \dots, Y_n\}$$

Then we have

$$d_N(K(B), K(B \cup D)) = 1 - \frac{1}{|U|} \left(\sum_{i=1}^{m+1} \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right)$$

$$= 1 - \frac{1}{|U|} \left(\sum_{\substack{i=1 \\ i \neq p}}^m \sum_{\substack{j=1 \\ j \neq q; j \neq s}}^n \frac{|X_i \cap Y_j|^2}{|X_i|} + \sum_{j=1}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p| - 1} + \sum_{\substack{j \neq q; j \neq s}}^n \frac{|(X_p - \{x\}) \cap (Y_q - \{x\})|^2}{|X_p| - 1} + \frac{|(X_p - \{x\}) \cap (Y_s \cup \{x'\})|^2}{|X_p| - 1} + \sum_{\substack{j=1 \\ j \neq q; j \neq s}}^n \frac{|x' \cap Y_j|^2}{|x'|} + \frac{|x' \cap (Y_q - \{x\})|^2}{|x'|} + \frac{|x' \cap (Y_s \cup \{x'\})|^2}{|x'|} + \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} + \sum_{\substack{i=1 \\ i \neq p}}^m \frac{|X_i \cap (Y_s \cup \{x'\})|^2}{|X_i|} \right)$$

$$= d_j(K(B), K(B \cup D)) + \frac{1}{|U|} \left(1 + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \right)$$

4) If $x' \in X_r, r \leq m$ and $x' \in Y_s, s \leq n$. We present the new partitions as follows:

$$(\{U - \{x\}\} \cup \{x'\})/B = \{X_1, X_2, \dots, X_p \setminus \{x\}, X_{p+1}, \dots, X_r \cup \{x'\}, \dots, X_m\}$$

$$\frac{(\{U - \{x\}\} \cup \{x'\})/D = \{Y_1, Y_2, \dots, Y_q \setminus \{x\}, Y_{q+1}, \dots, Y_s \cup \{x'\}, \dots, Y_n\}}$$

Then we have

$$\begin{aligned} d_N(K(B), K(B \cup D)) &= 1 - \frac{1}{|U|} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right) \\ &= 1 - \frac{1}{|U|} \left(\sum_{i \neq p; i \neq r}^m \sum_{j \neq q; j \neq s}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\ &\quad + \sum_{j=1}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p| - 1} \\ &\quad + \frac{|(X_p - \{x\}) \cap (Y_q - \{x\})|^2}{|X_p| - 1} \\ &\quad + \frac{|(X_p - \{x\}) \cap (Y_s \cup \{x'\})|^2}{|X_p| - 1} \\ &\quad + \sum_{j \neq q; j \neq s}^n \frac{|(X_r \cup \{x'\}) \cap Y_j|^2}{|X_r| + 1} \\ &\quad + \frac{|(X_r \cup \{x'\}) \cap (Y_q - \{x\})|^2}{|X_r| + 1} \\ &\quad + \frac{|(X_r \cup \{x'\}) \cap (Y_s \cup \{x'\})|^2}{|X_r| + 1} \\ &\quad + \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} \\ &\quad \left. + \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap (Y_s \cup \{x'\})|^2}{|X_i|} \right) \end{aligned}$$

Because $x \notin Y_j$ for any $j \neq q, j = 1, \dots, n$ we have

$$\sum_{j \neq q; j \neq s}^n \frac{|(X_p - \{x\}) \cap Y_j|^2}{|X_p| - 1} = \sum_{j \neq q; j \neq s}^n \frac{|X_p \cap Y_j|^2}{|X_p| - 1},$$

$$\frac{|(X_p - \{x\}) \cap (Y_s \cup \{x'\})|^2}{|X_p| - 1} = \frac{|X_p \cap Y_s|^2}{|X_p| - 1}$$

$$\text{And } \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap (Y_q - \{x\})|^2}{|X_i|} = \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap Y_q|^2}{|X_i|}$$

Similarly, $x' \notin Y_j$ for any $j \neq s, j = 1, \dots, n$ and $x \notin X_r$ so we have

$$\begin{aligned} \sum_{j \neq q; j \neq s}^n \frac{|(X_r \cup \{x'\}) \cap Y_j|^2}{|X_r| + 1} &= \sum_{j \neq q; j \neq s}^n \frac{|X_r \cap Y_j|^2}{|X_r| + 1} + \\ &\quad \frac{|(X_r \cup \{x'\}) \cap (Y_q - \{x\})|^2}{|X_r| + 1} = \frac{|X_r \cap Y_q|^2}{|X_r| + 1}, \end{aligned}$$

$$\frac{|(X_r \cup \{x'\}) \cap (Y_q - \{x\})|^2}{|X_r| + 1} = \frac{|X_r \cap Y_q|^2}{|X_r| + 1},$$

$$\text{And } \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap (Y_s \cup \{x'\})|^2}{|X_i|} = \sum_{i \neq p; i \neq r}^m \frac{|X_i \cap Y_s|^2}{|X_i|}$$

Consequently,

$$\begin{aligned} d_N(K(B), K(B \cup D)) &= 1 \\ &\quad - \frac{1}{|U|} \left(\sum_{i=1}^m \sum_{j=1}^n \frac{|X_i \cap Y_j|^2}{|X_i|} \right. \\ &\quad + \sum_{j=1}^n \frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \\ &\quad + \frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \\ &\quad - \sum_{j=1}^n \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} \\ &\quad \left. + \frac{2|X_r \cap Y_s| + 1}{|X_r| + 1} \right) \end{aligned}$$

$$\begin{aligned}
 &= d_j(K(B), K(B \cup D)) \\
 &\quad + \frac{1}{|U|} \left(\frac{1 - 2|X_p \cap Y_q|}{|X_p| - 1} \right. \\
 &\quad \left. + \frac{1 + 2|X_r \cap Y_s|}{|X_r| + 1} \right) \\
 &\quad + \frac{1}{|U|} \sum_{j=1}^n \left(\frac{|X_p \cap Y_j|^2}{|X_p|(|X_p| - 1)} \right. \\
 &\quad \left. - \frac{|X_r \cap Y_j|^2}{|X_r|(|X_r| + 1)} \right)
 \end{aligned}$$

The proof is completed.

5.2. Algorithm for finding reduct when update one object.

Algorithm 5.1. The incremental algorithm for finding reducts based on metric when updating a object.

Input: Decision table $DS = (U, C \cup D)$, reduct R_U on U and the object x is updated to x'

Output: Reduct $R_{\{U-\{x\}\} \cup \{x'\}}$ on $\{U - \{x\}\} \cup \{x'\}$.

1. Assign $R = R_U$, calculate $U/R = \{X_1, X_2, \dots, X_m\}$;
2. While $d_{\{U-\{x\}\} \cup \{x'\}}(K(R), K(R \cup D)) \neq d_{\{U-\{x\}\} \cup \{x'\}}(K(C), K(C \cup D))$ do
3. Begin
4. For $a \in C - R$ calculate $SIG_R(a)$ according to the incremental formulas in Proposition 5.1
5. Select $a_m \in C - R$ such that $SIG_R(a_m) = \text{Max}_{a \in C - R} \{SIG_R(a)\}$;
6. $R = R \cup \{a_m\}$;
7. End;
8. For $a \in R$ do
9. Begin
10. Calculate

$$d_{\{U-\{x\}\} \cup \{x'\}}(K(R - \{a\}), K((R - \{a\}) \cup D));$$

11. If $d_{\{U-\{x\}\} \cup \{x'\}}(K(R - \{a\}), K((R - \{a\}) \cup D)) = d_{\{U-\{x\}\} \cup \{x'\}}(K(R), K(R \cup D))$ then $R = R - \{a\}$;
12. End
13. Return R .

Similar to algorithm 4.1, one can prove that the complexity of algorithm 5.1 is $O(|C|^2|U| + \max(|C||X_p||Y_q|, |C||X_r||Y_s|))$. Obviously, $\max(|C||X_p||Y_q|, |C||X_r||Y_s|) < |C||U|^2$ and this means that the complexity of our algorithm is much better than that of the original one (algorithm 3.1).

6. CONCLUSIONS

In the real world of today, databases have being gradually got increased, changed and updated. To propose effective methods in order to optimize consuming time for carrying algorithms finding reducts is a goal of researchers. In this paper, we construct mechanisms for updating the metric used in finding reduct algorithms and propose algorithms to work out the reducts in the case of adding, deleting and updating an object from the universal set of a decision table. Since our algorithms do not reimplement to the varied database so that we increase calculating time for finding the reduct. Also we point out that our method are better than technique used in [16] in terms of the number of attributes in reducts. Besides, employing distance in difference measuring is known for its simplicity and effectiveness and we gained that in this work.

At the first stage, we limit the number of the varied objects for one. To develop our research to construct algorithms finding reducts when multi objects varied is completely capable. Our next research could be using metric to constructing algorithms for finding reducts in case of multi objects or attribute set varied.

6. REFERENCES

- [1] Deza M. M. and Deza E., “*Encyclopedia of Distances*”, Springer, 2009.
- [2] Guan L. H., “*An incremental updating algorithm of attribute reduction set in decision tables*”, FSKD'09 Proceedings of the 6th international conference on Fuzzy systems and knowledge discovery, Vol 2, 2009, pp. 421-425
- [3] Hu F., Wang G.Y., Huang H., Wu Y., “*Incremental attribute reduction based on elementary sets*”, Proceedings of the 10th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, Regina, Canada, 2005, pp. 185-193.
- [4] Nguyễn Long Giang, Phạm Hoàng Tuyền, “*Nghiên cứu sự thay đổi giá trị các độ đo đánh giá hiệu năng tập luật quyết định trên các tập rút gọn*”, Kỷ yếu Hội thảo Quốc gia lần thứ XV “Một số vấn đề chọn lọc của Công nghệ thông tin và truyền thông”, Hà Nội 11/2012, 2013, Tr. 295-301
- [5] Liang J.Y, Chin K.S., Dang C.Y. and Richard C.M.YAM, “*New method for measuring uncertainty and fuzziness in rough set theory*”, International Journal of General Systems **31**, 2002, pp. 331-342.
- [6] Liang J.Y, Wang F., Dang C.Y., Qian Y.H., “*A group incremental approach to feature selection applying rough set technique*”, IEEE Transactions on Knowledge and Data Engineering, 2014, **26(2)**:294 - 308.
- [7] Long Giang Nguyen, “*Metric Based Attribute Reduction in Decision Tables*”, The 2012 International Workshop on Rough Sets Applications (RSA'2012), FedCSIS Proceedings, IEEE, 2012, pp. 333-338.
- [8] Pawlak Z., “*Rough sets: Theoretical Aspects of Reasoning About Data*”, Kluwer Academic Publishers, 1991.
- [9] Wang F., Liang J. Y, Qian Y. H., “*Attribute reduction: A dimension incremental strategy*”, Knowledge-Based Systems, **Volume 39**, 2013, pp. 95–108
- [10] Feng Wang, Jiye Liang, Chuangyin Dang, “*Attribute reduction for dynamic data sets*”, Applied Soft Computing, 2013, **13(1)**: 676-689.
- [11] Wei W, J.Y. Liang, Y.H. Qian, F. Wang, C.Y. Dang, “*Comparative study of decision performance of decision tables induced by attribute reductions*”, International Journal of General Systems, Vol. **39**, No. **8**, 2010, 813–838.
- [12] Zhang C. S, Jing Ruan J.,Tan Y. H., “*An Improved Incremental Updating Algorithm for Core Based on Positive Region*”, Journal of Computational Information Systems 7: **9**, 2011, pp. 3127-3133.
- [13] Z. Y. Xu, Z. P. Liu, B. R. Yang, W. Song., “*A quick attribute reduction algorithm with complexity of $\max(O(|C||U|), O(|C|^2|U/C|))$* ”, Journal of Computer, Vol. 29, no. **3**, pp. 391-398, 2006.
- [14] Qian Y.H., Liang J.Y., Li D.Y., Zhang H.Y. and Dang C.Y. “*Measures of Evaluating The Decision Performace of a Decision Table in Rough Set Theory*”, Information Sciences, **Vol.178**, 2008, pp.181-202.
- [15] Z.Pawlak, A.Skowron, “*Rudiments of rough sets*”, Information Sciences 177(1), 2007, pp. 3-27.
- [16] J.Demetrovics, Vu Duc Thi, Nguyen Long Giang, “*Metric based attribute reduction in dynamic decision tables*”, Annales Univ, Sci Budapest, Sect. Comp. Vol.42, 2014, pp. 157-172.

AUTHORS' BIOGRAPHIES



Nguyen Thi Lan Huong graduated from Ha Noi Pedagogical University 2 in 1994, major in Mathmatics and Informatics. Received the Master degree in computer science from Military Technical Academy in 2007.

Research interests include data mining, database and machine learning.



Nguyen Long Giang graduated from Ha Noi University of Science and Technology in 1997. Received Master degree from VNU University of Engineering and Technology in 2003. Received the Ph.D degree in 2012 from Institute of Information

Technology - Vietnamese Academy of Science and Technology (VAST). Research interests include database, data mining and machine learning.