# Attribute Reduction Algorithm Based on Structure Discernibility Matrix in Composite Information Systems

Mei-Jun GE[1], Nian-Bai FAN[1,a] and Tao SUN[2]

[1]*College of Information Science and Engineering, Hunan University, Changsha 410082, China*
[2]*College of Mathematics and Econometric, Hunan University, Changsha 410082, China*

**Abstract.** Attribute reduction, as an important preprocessing step for knowledge acquiring in data mining, is one of the key issues in rough set theory. It can only deal with attributes of a specific type in the information system by using a specific binary relation. However, there may be attributes of multiple different types in information systems in real-life applications. A composite relation is proposed to process attributes of multiple different types simultaneously in composite information systems. In order to solve the time-consuming problem of traditional heuristic attribute reduction algorithms, a novel attribute reduction algorithm based on structure discernibility matrix was proposed in this paper. The proposed algorithms can choose the same attribute reduction as its previous version, but it can be used to accelerate a heuristic process of attribute reduction by avoiding the process of intersection and adopting the forward greedy attribute reduction approach. The theoretical analysis and experimental results with UCI data sets show that the proposed algorithm can accelerate the heuristic process of attribute reduction.

## 1 Introduction

Pawlak proposed the Rough set theory in 1980s[1], this theory has become a powerful mathematical tool for analyzing one of various types of data[2,3]. It can be used in an attribute value representation model to describe the dependencies among attributes, evaluate the significance of attributes and derive reduction[4,5].

The classical rough set model can only be used to deal with categorical attributes, However, there may be attributes of multiple different types in real-life applications. many extended rough set models have been developed for attributes of multiple different types. A neighborhood relation was proposed by Hu to deal with numerical attributes[6]. Guan defined a tolerance relation and used the maximal tolerance classes to derive optimal decision rules from set-valued information systems[7]. Qian used a binary dominance relation to process set-valued data in set-valued ordered information systems[8]. Leung defined $\alpha$-tolerance relations and employed the $\alpha$-misclassification rate for rule acquisition from interval-valued information systems[9]. To deal with missing data, the toleration and similarity relations as well as the limited tolerance relation were proposed[10]. Grzymała-Busse combined the toleration and similarity relations and presented characteristic relations for missing data in information systems[11].

[a] Corresponding author: 398000862@qq.com

Most of the classical rough set methods fail to deal with more than attributes of two different types. Many scholars introduced the composite rough set model and proposed the basic idea to deal with attributes of multiple different types[12-14]. we introduced a structure discernibility matrix[15] to solve the time-consuming problem of traditional heuristic attribute reduction algorithms in this paper. The proposed algorithms can choose the same attribute reduction as its original version, but it can be used to accelerate a heuristic process of attribute reduction by avoiding the process of intersection and adopting the forward greedy attribute reduction approach. Extensive experiments on different data sets from UCI show that the proposed structure discernibility matrix-based method can process large data sets efficiently.

## 2 Composite rough set model

In many practical issues, there are attributes of multiple different types in the information system, we call it a composite information system. A composite information system can be written as $CIS=(U,A,V,f)$, where $U$ is a non-empty finite set of objects; $A$ is a non-empty finite set of attributes; $V=\bigcup_{a\in A}V_a$ and $V_a$ is a domain of attribute $a$; $f:U\times A\to V$ is an information function such that $f(x,a)\in V_a$ for every $x\in U,a\in A$.

More specifically, a composite information system is also called a composite decision table if there are condition and decision attributes in the information system, which is denoted by $CDT=(U,A\cup D,V,f)$.

**Definition 1**[16]. Given $x,y\in U$ and $B\subseteq A$, the composite relation $CR_B$ is defined as

$$CR_B=\{(x,y)\mid(x,y)\in\bigcap R_a,\forall a\in B\} \tag{1}$$

Where $R_a\subseteq U\times U$ is an indiscernibility relation defined by an attribute $a$ on $U$.

When $(x,y)\in CR_B$, we call $x$ and $y$ are indiscernible on $B$. Let $CR_B(x)=\{y\mid y\in U,\forall a\in B,yR_ax\}$, we call $CR_B(x)$ the composite class for $x$ on $CR_B$.

**Definition 2**[12]. Given a composite information system $CIS=(U,A,V,f)$, $\forall X\subseteq U,B\subseteq A$, the lower and upper approximations of $X$ in terms of the composite relation $CR_B$ are defined as

$$CR_B(X)=\{x\in U\mid CR_B(x)\subseteq X\} \tag{2}$$

$$CR^B(X)=\{x\in U\mid CR_B(x)\cap X\neq\varnothing\} \tag{3}$$

Here, the positive region $POS_{CR_B}(X)=CR_B(X)$ (4)

**Definition 3**[12]. Given a composite decision table $CDT=(U,A\cup D,V,f)$, $B\subseteq A$. Let $U/D=\{D_1,D_2,...,D_r\}$ be a partition over the decision $D$. Then the lower and upper approximations of the decision $D$ with respect to attributes $B$ are defined as

$$CR_B(D)=\bigcup_{j=1}^{r}CR_B(D_j) \tag{5}$$

$$CR^B(D)=\bigcup_{j=1}^{r}CR^B(D_j) \tag{6}$$

The positive region $POS_{CR_B}(D)=CR_B(D)$ (7)

**Definition 4**[17]. Let $S=(U,A\cup D,V,f)$ be a decision table, $B\subseteq A$ and $\forall a\in B$, if $POS_B(D)=POS_{B-\{a\}}(D)$, $a$ is unnecessary in $B$ relative to $D$, else $a$ is necessary in $B$ relative to $D$. If $\forall a\in B$ is necessary, $B$ is independent relative to $D$.

**Definition 5**[18]. Let $S=(U,A\cup D,V,f)$ be a decision table, $B\subseteq A$. If $POS_B(D)=POS_A(D)$, and $\forall a\in B$ is necessary, $B$ is the reduce of $A$ relative to $D$.

**Definition 6**[18]. Let $S=(U,A\cup D,V,f)$ be a decision table, $B\subseteq A$ and $\forall a\in B$. The significance measure of a in B is defined as

$$Sig^{inner}(a,B,D)=(|POS_B(D)-POS_{B-\{a\}}(D)|)/U \tag{8}$$

If $Sig^{inner}(a,B,D)>0$, $a$ is core attribute of $B$ related to $D$.

**Definition 7**[18]. Let $S=(U,A\cup D,V,f)$ be a decision table, $B\subseteq A$ and $\forall a\in A-B$. The significance measure of $a$ in $B$ is defined as

$$Sig^{outer}(a,B,D)=(|POS_{B\cup\{a\}}(D)-POS_B(D)|)/U \tag{9}$$

Example 1. A composite decision table $CDT=(U,A\cup D,V,f)$ is presented in Table 1. Let $A=\bigcup_{k=1}^{5}a_k$. We set the neighborhood parameter $\delta=0.15$ and adopt Manhattan distance.

**Table 1.** A composite decision table

| $U$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $D$ |
|---|---|---|---|---|---|---|
| $x_1$ | Y | {1,2} | 0.2 | 0.1 | * | Yes |
| $x_2$ | Y | {1} | 0.2 | 0.3 | ? | No |
| $x_3$ | Y | {0} | 0.1 | 0.1 | Small | Yes |
| $x_4$ | Y | {0,1,2} | 0.1 | 0.2 | Small | Yes |
| $x_5$ | N | {1} | 0.1 | 0.3 | Large | Yes |
| $x_6$ | N | {0,2} | 0.2 | 0.2 | Large | No |

**Table 2.** Results of composite classes

| $U$ | $R_{a_1}(x)$ | $R_{a_2}(x)$ | $R_{a_3}(x)$ | $R_{a_4}(x)$ | $R_{a_5}(x)$ | $CR_B(x)$ |
|---|---|---|---|---|---|---|
| $x_1$ | $\{x_1,x_2,x_3,x_4\}$ | $\{x_1,x_2,x_4,x_5,x_6\}$ | $U$ | $\{x_1,x_3,x_4,x_6\}$ | $U$ | $\{x_1,x_4\}$ |
| $x_2$ | $\{x_1,x_2,x_3,x_4\}$ | $\{x_1,x_2,x_4,x_5\}$ | $U$ | $\{x_2,x_4,x_5,x_6\}$ | $U$ | $\{x_2,x_4\}$ |
| $x_3$ | $\{x_1,x_2,x_3,x_4\}$ | $\{x_3,x_4,x_6\}$ | $U$ | $\{x_1,x_3,x_4,x_6\}$ | $\{x_1,x_3,x_4\}$ | $\{x_3,x_4\}$ |
| $x_4$ | $\{x_1,x_2,x_3,x_4\}$ | $U$ | $U$ | $U$ | $\{x_1,x_3,x_4\}$ | $\{x_1,x_3,x_4\}$ |
| $x_5$ | $\{x_5,x_6\}$ | $\{x_1,x_2,x_4,x_5\}$ | $U$ | $\{x_2,x_4,x_5,x_6\}$ | $\{x_1,x_5,x_6\}$ | $\{x_5\}$ |
| $x_6$ | $\{x_5,x_6\}$ | $\{x_1,x_3,x_4,x_6\}$ | $U$ | $U$ | $\{x_1,x_5,x_6\}$ | $\{x_6\}$ |

According to the introduction in Section 2, it is easy to know that $R_{a_1},R_{a_2},R_{a_3},R_{a_4},R_{a_5}$ are equivalence relation, tolerance relation, neighborhood relation, neighborhood relation and characteristic relation respectively. By Definition 4, $CR_B(x)=\{y\mid y\in U,\forall B_k\in B,yR_{B_k}x\}$, the results are listed in Table 2.

It is easy to obtain that $U/D=\{D_1,D_2\}$, where $D_1=\{x_1,x_3,x_4,x_5\}$, $D_2=\{x_2,x_6\}$.

Since $CR_A(x_1)=\{x_1,x_4\}$, $CR_A(x_2)=\{x_2,x_4\}$, $CR_A(x_3)=\{x_3,x_4\}$,

$CR_A(x_4)=\{x_1,x_3,x_4\}$, $CR_A(x_5)=\{x_5\}$, $CR_A(x_6)=\{x_6\}$.

$POS_A(D) = \{x_1, x_3, x_4, x_5, x_6\}$.

$CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_1) = \{x_1, x_4, x_6\}$, $CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_2) = \{x_2, x_4, x_5\}$,

$CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_3) = \{x_3, x_4\}$, $CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_4) = \{x_1, x_3, x_4\}$,

$CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_5) = \{x_5\}$, $CR_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(x_6) = \{x_6\}$.

$POS_{(a_2 \cup a_3 \cup a_4 \cup a_5)}(D) = \{x_3, x_4, x_5, x_6\} \neq POS_A(D)$.

So $a_1$ is necessary in $A$ relative to $D$. Use the same method to calculate every attribute in $A$, for simplicity, just think $a_1$, finally the core is $\{a_1, a_2, a_4, a_5\}$.Then use the same method we obtain that

$CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_1) = \{x_1, x_4\}$, $CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_2) = \{x_2, x_4\}$,

$CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_3) = \{x_3, x_4\}$, $CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_4) = \{x_1, x_3, x_4\}$,

$CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_5) = \{x_5\}$, $CR_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(x_6) = \{x_6\}$.

$POS_{(a_1 \cup a_2 \cup a_4 \cup a_5)}(D) = \{x_1, x_3, x_4, x_5, x_6\} = POS_A(D)$.

According Definition 5, $\{a_1, a_2, a_4, a_5\}$is the reduction of attribute sets $A$ relative to $D$.

The traditional heuristic attribute reduction algorithms is time-consuming, section 3 gives structure discernibility matrix-based heuristic attribute reduction algorithms, which can accelerate a heuristic process of attribute reduction by avoiding the process of intersection and adopting the forward greedy attribute reduction approach.

## 3 Structure discernibility matrix-based attribute reduction algorithm

In this section, the attribute reduction algorithm based on structure discernibility matrix in complex information system is presented and the sample is listed.

### 3.1 The sample analysis of attribute reduction algorithm based on structure discernibility matrix

**Definition 8**[12]. Given an composite information system $CIS = (U, A, V, f)$. Let $a \in A$ and $R_a$ be an discernibility relation on $U$, $M^{R_a}_{n \times n} = (\xi_{ij})_{n \times n}$ be an $n \times n$ matrix representing $R_a$, called the relation matrix on $a$. Then

$$\xi^a_{ij} = \begin{cases} 1, (x_i, x_j) \in R_a \\ 0, (x_i, x_j) \notin R_a \end{cases} \tag{10}$$

**Definition 9**[15]. Given an composite information system $CIS = (U, A, V, f)$. Let $B \subseteq A$ and $R_B$ be an discernibility relation on $U$, $M^{R_B}_{n \times n} = (\xi_{ij})_{n \times n}$ be an $n \times n$ matrix representing $R_B$, called the relation matrix on $B$. Then

$$\xi^B_{ij} = \begin{cases} |B|, (x_i, x_j) \in R_B \\ \sum_{k=1}^{|B|} \xi^k_{ij}, (x_i, x_j) \notin R_B \end{cases} \tag{11}$$

$\xi^k_{ij}$ is the value of the kth attribute in $B$.

**Proposition 1** A composite decision table $CDT = (U, A \cup D, V, f)$, Let $B \subseteq A$ and $R_B$ be an discernibility relation on $U$, $M^{R_B}_{n \times n} = (\xi_{ij})_{n \times n}$ is a relation matrix on $B$, $M^{R_D}_{n \times n} = (\xi_{ij})_{n \times n}$ is a relation matrix on

$D$, If the value is $|B|$ in the ith row of $M_{n\times n}^{R_B}$, corresponding location is 1 in $M_{n\times n}^{R_D}$, then $x_i$ is in $POS_{CR_B}(D)$.

**Proof:** $\exists X \subseteq U/D, R_B(X) = \{x_i \in U \mid CR_B(x_i) \subseteq X\}$,

$CR_B(x_i) \subseteq R_B(X)$, so $x_i \in POS_{CR_B}(D)$.

$POS_{CR_A}(D)$ can be calculated by proposition 1.

Let $B \subseteq A, a \in B$ and $R_a$ be an discernibility relation on $U$, let $M'_{sum} = M_{sum} - M_{n\times n}^{R_a} = (\xi'_{ij})_{n\times n}$, $B' = B - \{a\}$, then $POS_{B'}(D)$ can be calculated by the same method.

**Propositon 2** A composite decision table $CDT = (U, A \cup D, V, f)$, $B \subseteq A$. If $POS_{B'}(D) \neq POS_B(D)$, then $a \in B$ is a core attribute of $B$ related $D$. The proof of proposition 2 is easy by definition 6.

Example 2. A composite decision table $CDT = (U, A \cup D, V, f)$ is presented in Table 1. Let $A = \bigcup_{k=1,2,3,4,5} a_k$. We set the neighborhood parameter $\delta = 0.15$ and adopt Manhattan distance. The process of get one attribute reduction usually includes three steps.

**Step1** The construction of the relation matrix

According to Definition 8, we have

$$M_{6\times6}^{R_{a_1}} = \begin{bmatrix} 1&1&1&1&0&0\\ 1&1&1&1&0&0\\ 1&1&1&1&0&0\\ 1&1&1&1&0&0\\ 0&0&0&0&1&1\\ 0&0&0&0&1&1 \end{bmatrix} M_{6\times6}^{R_{a_2}} = \begin{bmatrix} 1&1&0&1&1&1\\ 1&1&0&1&1&0\\ 0&0&1&1&0&1\\ 1&1&1&1&1&1\\ 1&1&0&1&1&0\\ 1&0&1&1&0&1 \end{bmatrix} M_{6\times6}^{R_{a_3}} = \begin{bmatrix} 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1 \end{bmatrix}$$

$$M_{6\times6}^{R_{a_4}} = \begin{bmatrix} 1&0&1&1&0&1\\ 0&1&0&1&1&1\\ 1&0&1&1&0&1\\ 1&1&1&1&1&1\\ 0&1&0&1&1&1\\ 1&1&1&1&1&1 \end{bmatrix} M_{6\times6}^{R_{a_5}} = \begin{bmatrix} 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&0&1&1&0&0\\ 1&0&1&1&0&0\\ 1&0&0&0&1&1\\ 1&0&0&0&1&1 \end{bmatrix} M_{6\times6}^{R_D} = \begin{bmatrix} 1&0&1&1&1&0\\ 0&1&0&0&0&1\\ 1&0&1&1&1&0\\ 1&0&1&1&1&0\\ 1&0&1&1&1&0\\ 0&1&0&0&0&1 \end{bmatrix}$$

**Step2** The calculation of core attributes

First we add the structure discernibility matrix of all the condition attributes, namely

$$M_{sum} = \sum_{k=1}^{5} M_{6\times6}^{R_{a_k}} = \begin{bmatrix} 5&4&4&5&3&4\\ 4&5&3&5&4&3\\ 4&2&5&5&1&3\\ 5&4&5&5&3&3\\ 3&3&1&3&5&4\\ 4&3&3&3&4&5 \end{bmatrix}$$

If $\xi_{ij} = k$, it illustrated that $i$ and $j$ are in the same class on $k$ attributes. in this example, there are 5 conditional attributes, so if $\xi_{ij} = 5$, $i$ and $j$ are in the same class on the conditional attributes. If all of the location is 5 in the $i$ row, corresponding location in the $M_{6\times6}^{R_D}$ is 1, $x_i$ is in $POS_{CR_A}(D)$, so we have $POS_{CR_A}(D) = \{x_1, x_3, x_4, x_5, x_6\}$.

According to $M_{sum}$, we have

$CR_A(x_1) = \{x_1, x_4\}, CR_A(x_2) = \{x_2, x_4\}, CR_A(x_3) = \{x_3, x_4\}$,

$CR_A(x_4) = \{x_1, x_3, x_4\}, CR_A(x_5) = \{x_5\}, CR_A(x_6) = \{x_6\}$.

$CR_D(x_1) = \{x_1, x_3, x_4, x_5\}, CR_D(x_2) = \{x_2, x_6\}, CR_D(x_3) = \{x_1, x_3, x_4, x_5\}$,

$CR_D(x_4) = \{x_1, x_3, x_4, x_5\}, CR_D(x_5) = \{x_1, x_3, x_4, x_5\}, CR_D(x_6) = \{x_2, x_6\}$.

According to Definition 3, we have $POS_{CR_A}(D) = \{x_1, x_3, x_4, x_5, x_6\}$.

$$M_{(A-a_1)} = M_{sum} - M_{6\times6}^{R_{a_1}} = \begin{bmatrix} 4 & 3 & 3 & 4 & 3 & 4 \\ 3 & 4 & 2 & 4 & 4 & 3 \\ 3 & 1 & 4 & 4 & 1 & 3 \\ 4 & 3 & 4 & 4 & 3 & 3 \\ 3 & 3 & 1 & 3 & 4 & 3 \\ 4 & 3 & 3 & 3 & 3 & 4 \end{bmatrix}$$

In order to save the storage space, we use char to Store the matrix. If $\xi_{ij}^{(A-a_1)} = 4$, $i$ and $j$ are in the same class, we use $1$ to represent it ,else we use $0$ to represent it.

If all of the location is $1$ in the $i$ row, corresponding location in the $M_{6\times6}^{R_D}$ is $1$, $x_i$ is in $POS_{CR_{(A-a_1)}}(D)$, so we have $POS_{CR_{(A-a_1)}}(D) = \{x_3, x_4, x_5\}$.

According to $M_{(A-a_1)}$, we have

$CR_{(A-a_1)}(x_1) = \{x_1, x_4, x_6\}$, $CR_{(A-a_1)}(x_2) = \{x_2, x_4, x_5\}$, $CR_{(A-a_1)}(x_3) = \{x_3, x_4\}$,

$CR_{(A-a_1)}(x_4) = \{x_1, x_3, x_4\}$, $CR_{(A-a_1)}(x_5) = \{x_5\}$, $CR_{(A-a_1)}(x_6) = \{x_1, x_6\}$.

$POS_{CR_{(A-a_1)}}(D) = \{x_3, x_4, x_5\} \neq POS_{CR_A}(D)$.

According to Definition 4, we have $a_1$ is necessary of $A$ relative to $D$.

The method of determining whether $a_2, a_3, a_4, a_5$ is core is same as the method of determining $a_1$, no longer list here. After calculation, we know that $\{a_1, a_2, a_4, a_5\}$ is the core attributes of $A$ relative to $D$.

**Step3** The calculation of attribute significance

After calculated core attributes, we with the core attributes as the opening of attribute reduction. First we add the structure discernibility matrix of the core attributes, namely

$$M_{core} = M_{6\times6}^{R_{a_1}} + M_{6\times6}^{R_{a_2}} + M_{6\times6}^{R_{a_4}} + M_{6\times6}^{R_{a_5}} = \begin{bmatrix} 4 & 3 & 3 & 4 & 2 & 3 \\ 3 & 4 & 2 & 4 & 3 & 2 \\ 3 & 1 & 4 & 4 & 0 & 2 \\ 4 & 3 & 4 & 4 & 2 & 2 \\ 2 & 2 & 0 & 2 & 4 & 3 \\ 3 & 2 & 2 & 2 & 3 & 4 \end{bmatrix}$$

If all of the location is $4$ in the ith row, corresponding location in $M_{6\times6}^{R_D}$ is $1$, $x_i$ is in $POS_{CR_{core}}(D)$, so we have $POS_{CR_{core}}(D) = \{x_1, x_3, x_4, x_5, x_6\} = POS_{CR_A}(D)$. According to definition 5, we have $\{a_1, a_2, a_4, a_5\}$ is the reduce of $A$ relative to $D$. If the core attributes set is not the reduce of $A$ relative to $D$, we will calculate the significance of the attribute in $\{A - core\}$ according definition 7. Then add the attribute whose significance is the biggest in order until find the reduce of $A$ relative to $D$.

### 3.2 Attribute reduction algorithm based on structure discernibility matrix

In this section, we design attribute reduction algorithm based on structure discernibility matrix in the composite decision table. MRPR algorithm is a structure discernibility matrix reduction algorithm based on positive region.

Step 1 is to construct the relation matrix and its key steps are to compute the equivalence relation, the neighborhood relation, the tolerance relation, the characteristic relation. Suppose $A = \bigcup_{k=1,2,3,4} a_k$ and $a_1, a_2, a_3, a_4$ generate the equivalence relation, the neighborhood relation, the tolerance relation, the characteristic relation on the universe. The time complexity of computing the equivalence relation is $O(|U|)$ [19]; The time complexity of computing the neighborhood relation is $O(|U|\log|U|)$[6]; The time complexity of computing the tolerance relation is about $O(|U|)$[20]; The time complexity of computing the characteristic relation is about $O(|U|)$[18].

Step 2 is calculate the sum of the structure discernibility matrix of all the condition attributes and its time complexity is $O(|A||U|^2)$.

Step 3 is calculation core attributes of the condition attributes $A$ relative to $D$ and its time consume is $2|A||U|^2$. This step is the key step outperform previous heuristic attribute reduction algorithms by avoid the process of intersection.

Step 4 is calculation significance of attribute in $A$ but not in core, its time complexity is $O(|A||U|^2)$.

Step 5 is add the attribute whose significance is the biggest in order until find the reduce of $A$ relative to $D$. we adopt greedy and forward search algorithms. These search algorithms start with a nonempty set, and keep adding one attribute of highest significance into a pool each time until the dependence has not been increased. its time complexity is $O(|A||U|)$ [18].

**Algorithm.** A structure discernibility matrix reduction algorithm based on positive region (MRPR)

Input: A composite decision table $CDT = (U, A \cup D, V, f)$.

Output: A reduce of the composite decision table.

begin

1 Construct the relation matrix: $M^{R_A}_{|U|\times|U|} = (\xi^a_{ij})_{|U|\times|U|}$. // According to Definition 8

2 Calculate the sum of the structure discernibility matrix of all the condition attributes. // According to Definition 9

3 Calculation core attributes of the condition attributes $A$ relative to $D$. // According to Definition 6

4 Calculation significance of attribute in $A$ but not in core. // According to Definition 7

5 Add the attribute whose significance is the biggest in order until find the reduce of $A$ relative to $D$. // According to Definition 5

end

## 4 Experimental analysis

The experiments are carried out on a PC with the operation system win7 (64-bit), which has 4 GB main memory and uses Inter Core (TM) i3-3240 CPU with a clock frequency of 3.40 GHz. All the algorithms are coded in C++ and complied with Dev-C++.

To test and compare the performances of the MRPR algorithms and traditional heuristic attribute reduction algorithms based on positive region (RPR)[17] and A general improved feature selection algorithm based on the positive region (FSPR)[18], we download six data sets from UCI[21]. All these data sets are outlined in Table 3.

**Table 3.** A description of data sets.

| Data sets | Sam-ples | Attri-butes | Cla-sses |
|---|---|---|---|
| Lenses | 24 | 4 | 3 |
| Lung | 32 | 56 | 3 |
| Zoo | 101 | 16 | 7 |
| Vehicle | 846 | 18 | 4 |
| Car | 1728 | 6 | 4 |
| Chess | 3196 | 36 | 2 |

**Table 4.** The time-consuming of different attribute reduction algorithms.

| Data sets | Core attribute | Reduction result | MRPR time(s) | FSPR time(s) | RPR time(s) |
|---|---|---|---|---|---|
| Lenses | 4 | 4 | 0.002 | 0.003 | 0.003 |
| Lung | 0 | 20 | 0.012 | 0.025 | 0.035 |
| Zoo | 2 | 6 | 0.011 | 0.020 | 0.026 |
| Vehicle | 0 | 4 | 0.747 | 1.665 | 2.419 |
| Car | 6 | 6 | 0.670 | 0.735 | 0.739 |
| Chess | 27 | 30 | 19.590 | 96.542 | 119.479 |

Table 4 shows the experimental result, the time is average value of 10 times reduction time-consuming. Figure 1-6 can express the results more clearly.

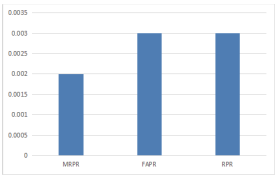


**Figure 1.** Lenses database

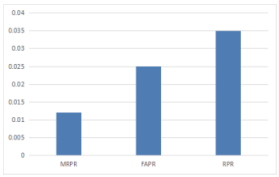


**Figure 2.** Lung database

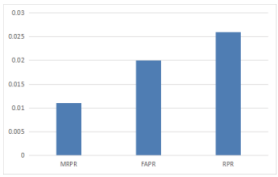


**Figure 3.** Zoo database

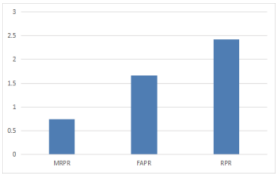


**Figure 4.** Vehicle database

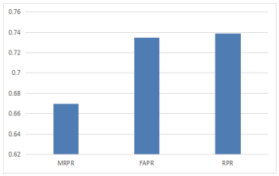


**Figure 5.** Car database

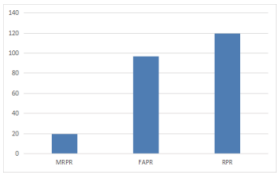


**Figure 6.** Chess database

We can see that the modified algorithms are faster than their original counterparts on these six data sets, which shows that the proposed structure discernibility matrix-based method can process data sets more efficiently. Sometimes, the effect of this reduction can reduce over half the computational time and even more. For example, the MRPR algorithm reduced time achieves 0.012 seconds on the data set Lung, while the reduced time is 0.025 seconds of FSPR algorithm and the reduced time is 0.035 seconds of RPR algorithm. The result on large data sets is more outstanding, For example, the MRPR algorithm reduced time achieves 19.590 seconds on the data set Chess, while the reduced time is 96.542 seconds of FSPR algorithm and the reduced time is 119.479 seconds of RPR algorithm. So the proposed structure discernibility matrix-based method can accelerate the heuristic process of attribute reduction and process large data sets more efficiently.

## Conclusions

To overcome the time limitations of the existing heuristic attribute reduction schemes, in this paper, a theoretic framework based on rough set theory have been proposed, called attribute reduction algorithm based on structure discernibility matrix, which can be used to accelerate algorithms of heuristic attribute reduction. Based on this framework, a structure discernibility matrix reduction algorithm based on positive region (MRPR) has been presented. Note that the MRPR algorithm can choose the same feature subset as the previous attribute reduction algorithm. Experiment on six UCI data sets show that the modified algorithms can significantly reduce computing time of attribute reduction while producing the same attribute reductions and classification accuracy as those coming from the previous methods. The results show that the attribute reduction algorithm based on structure

discernibility matrix is an effective accelerator and can efficiently obtain an attribute reduction. We will develop a parallel method to process attribute reduction in future work.

## Acknowledgment

This paper was supported by the key program of Hunan Provincial Department of Science and Technology support (Project Number: 2016JC2014).

## References

1. Z. Pawlak,Rough Sets[J].International Journal of Information Computer Science,1982,11(5):341-356.
2. Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, System Theory, Knowledge Engineering and Problem Solving, vol. 9, Kluwer Academic Publishers, Dordrecht, 1991.
3. Z. Pawlak, A. Skowron, Rudiments of rough sets, Information Sciences 177 (2007) 3–27.
4. D. Liu, T.R. Li, D. Ruan, J.B. Zhang, Incremental learning optimization on knowledge discovery in dynamic business intelligent systems, Journal of Global Optimization 51 (2011) 325–344, http://dx.doi.org/10.1007/s10898-010-9607-8.
5. J. Qian, D.Q. Miao, Z.H. Zhang, W. Li, Hybrid approaches to attribute reduction based on indiscernibility and discernibility relation, International Journal of Approximate Reasoning 52 (2011) 212–230.
6. Q.H. Hu, D.R. Yu, J.F. Liu, C.X. Wu, Neighborhood rough set based heterogeneous feature subset selection, Information Sciences 178 (2008) 3577–3594.
7. Y. Guan, H. Wang, Set-valued information systems, Information Sciences 176 (2006) 2507–2525.
8. Y.H. Qian, C.Y. Dang, J.Y. Liang, D.W. Tang, Set-valued ordered information systems, Information Sciences 179 (2009) 2809–2832.
9. Y. Leung, M.M. Fischer, W.Z. Wu, J.S. Mi, A rough set approach for the discovery of classification rules in interval-valued information systems,International Journal of Approximate Reasoning 47 (2008) 233–246.
10. G.Y. Wang, Extension of rough set under incomplete information systems, in: Proceedings of the 2002 IEEE International Conference on Fuzzy Systems,vol. 2, pp. 1098–1103.
11. J. Grzymala-Busse, Characteristic relations for incomplete data: a generalization of the indiscernibility relation, in: S. Tsumoto, R. Slowinski, J.Komorowski, J. Grzymala-Busse (Eds.), Rough Sets and Current Trends in Computing, Lecture Notes in Computer Science, vol. 3066, Springer, Berlin/Heidelberg, 2004, pp. 244–253.
12. J.B. Zhang, T.R. Li, H.M. Chen, Composite rough sets for dynamic data mining, Information Sciences, 2014, 257(2):81-100.
13. J.B. Zhang, T.R. Li, H.M. Chen, Composite rough sets, in: J. Lei, F. Wang, H. Deng, D. Miao (Eds.), Artificial Intelligence and Computational Intelligence, Lecture Notes in Computer Science, vol. 7530, Springer, Berlin/ Heidelberg, 2012, pp. 150–159.
14. H.M. Abu-Donia, Multi knowledge based rough approximations and applications, Knowledge-Based Systems 26 (2012) 20–29.
15. GE Hao, LI Longshu, YANG Chuanjian. Discernibility matrix-based reduct representation and quick algorithms [J]. Control and Decision, 2016, 31(1):12-20.
16. Z. Pawlak, A. Skowron, Rough sets: some extensions, Information Sciences 177 (2007) 28–40.
17. Zhang Wenxiu,Wu Weizhi,liang Jiye, The theory and method of rough set.Science Publishing Company.2003.1.
18. Y.H. Qian, J.Y. Liang, W. Pedrycz, C.Y. Dang, Positive approximation: an accelerator for attribute reduction in rough set theory, Artificial Intelligence 174 (2010) 597–618.
19. Z.Y. Xu, Z.P. Liu, B.R. Yang, W. Song, A quick attribute reduction algorithm with complexity of max(O(jCjjUj), O(jCj2jU/Cj)), Chinese Journal of Computers 29 (2006) 391–398.

20. Y.H. Qian, J.Y. Liang, F. Wang, A positive-approximation based accelerated algorithm to feature selection from incomplete decision tables, Chinese Journal of Computers 34 (2011) 435–442.
21. J. Stefanowski, A. Tsoukias, Incomplete information tables and rough classification, Computational Intelligence 17 (2001) 545–566.