TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7399~7406
e-ISSN: 2087-278X

Dominance-based Matrix algorithm for Knowledge Reductions in Incomplete Fuzzy System

Lixin Fan, Qiang Wu*
Department of Computer Science and Technology, Shaoxing University, Shaoxing, 312000, Zhejiang, People's Republic of China
*Corresponding author, e-mail: cswq@usx.edu.cn

Abstract

In this paper, definitions of knowledge granulation and rough entropy are proposed based on dominance relations in incomplete fuzzy system (fuzzy information system), and important properties are obtained. It can be found that using the definitions can measure uncertainty of an attribute set in the incomplete fuzzy information systems. A matrix algorithm for attributes reduction is acquired in the systems. An example illustrates the validity of this algorithm, and results of compared with other existing methods show that the algorithm is an efficient tool for data mining.

Keywords: *incomplete fuzzy information system, knowledge granulation, rough entropy, dominance matrix, knowledge reduction*

1. Introduction

The rough set theory proposed by Pawlak in 1982 [1] is a new mathematical tool to deal with the uncertain, vague, inaccurate information. Classic rough set theory built on equivalence relation (reflexive, symmetric, transmission) mainly solves the problem of incomplete information system. However; the vast majority of the information is incomplete, vague in reality. In order to get a broader application of the rough set theory, many scholars improved methods for incomplete information system [2-4]. For fuzzy information system, some researchers have suggested some improvement methods [5-8]. A lot of information in real problem is not only of incomplete, vague, and with preference relations. Greco, etc. [9-11] firstly introduced dominance relation into rough set model and fuzzy rough set, but they can only handled complete information. To this end, Wei et al [12] expanded the dominance relation to incomplete fuzzy information systems. Soon, they proposed rough fuzzy set in incomplete fuzzy information system based on similarity dominance relation [13].

This study aims to IFIS, the incomplete fuzzy information system. There have been unknown attribute values. We believe that the unknown attribute values are only missing, but it is also real. In other words, it is the fact that inaccurate knowledge forcing people to deal with only some of the information, incomplete information system. Each individual object has complete information on the potential. At present it is only just missed out these values. Therefore, in IFIS, the unknown attribute values can be considered as any other known attribute values are comparable. According to this explanation, using the extended dominance relations, a fuzzy rough set model is built into IFIS. Based on this fuzzy rough set model, we address knowledge granulation and rough entropy of rough set in IFIS, thus to knowledge reduction.

In next section, IFIS and its extended dominance relations are reviewed. Knowledge granulation is defined. In Section 3, after giving knowledge reduction algorithm based on dominance matrix, an example shows all reducts are enumerated by the matrices associated with IFIS. The following results of experiments were given by Section 4, Finally, we describe conclusions in Section 5.

Received June 28, 2013; Revised August 18, 2013; Accepted August 30, 2013

2. Basic Notions

A fuzzy information system is the 4-tuple *S*=<*U*,*AT*,*V*,*f*>,where *U*– is a nonempty set, called the universe, *AT* – is a finite set of fuzzy attributes, *V*– is a set of fuzzy (linguistic) values of attributes, $V=V_{AT}=\bigcup_{a\in AT} V_a$,V_a is the set of linguistic values of an attribute $a\in AT$, *f* – is an information function, $f: U\times AT\to[0,1]$, $f(x,a)\in[0,1]$ for every $x\in U$ and every $a\in AT$.

In practice, we use fuzzy decision tables, which constitute a special form of fuzzy information systems with two disjoint groups of condition and decision attributes, respectively.

When the precise values for some of the objects on some fuzzy attributes are not known, i.e. unknown values (symbol "*" is used to express unknown value), then such a fuzzy system is referred to as an incomplete fuzzy information system (IFIS). In this paper, an IFIS are still recorded as *S*=<*U*,*AT*,*V*,*f*>, at this time $V=[0,1]\cup\{*\}$.

Table 1 is an IFIS, of which $U=\{x_1,x_2,\ldots,x_{10}\}$, attribute set $AT=\{a_1,a_2,a_3,a_4\}$.

Table 1. An IFIS

U	a_1	a_2	a_3	a_4
x_1	0.9	*	0.2	0.7
x_2	0.9	0.2	0.2	0.1
x_3	0.1	0.1	0.1	0.9
x_4	0.0	0.9	*	0.8
x_5	0.1	0.1	1.0	0.8
x_6	*	0.2	0.9	0.1
x_7	0.0	0.1	0.9	0.2
x_8	0.9	0.9	0.1	1.0

In IFIS, we believe that the unknown attribute value is only missing, but they actually exist, and therefore these unknown values can be combined with any other attribute value for comparison. According to this understanding, the dominance relation can be constructed as follows.

Definition 1: Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS,$B\subseteq AT$. The dominance relation in terms of *B* is defined as:

$$R_B^{\geq}=\{(x,y)\in U^2 \mid \forall a\in B, f(x,a)\geq f(y,a)\vee f(x,a)=*\vee f(y,a)=*\} \quad (1)$$

$[x_i]_B^{\geq}$ is called a dominance class of object x_i, if:

$$[x_i]_B^{\geq}=\{x_j\in U \mid (x_j,x_i)\in R_B^{\geq}\}=\{x_j\in U\mid \forall a\in B, f(x_j,a)\geq f(x_i,a)\vee f(x_i,a)=*\vee f(x_j,a)=*\} \quad (2)$$

$$U/R_B^{\geq}=\{[x_i]_B^{\geq}\mid x_i\in U\} \quad (3)$$

$U/R_B^{\geq}$ is a classification for the object set on the attribute set *B*.

In terms of dominance classes of *B*, the pair of lower and upper approximation operators can be defined by:

$$\underline{R_B^{\geq}}(X)=\{x_i\in U \mid [x_i]_B^{\geq}\subseteq X\} \quad (4)$$

$$\overline{R_B^{\geq}}(X)=\{x_i\in U \mid [x_i]_B^{\geq}\cap X\neq\varnothing\} \quad (5)$$

An element $x\in U$ belongs to the lower approximation of *X* if all its dominance elements belong to *X*. It belongs to the upper approximation of *X* if at least one of its dominance elements belongs to *X*.

As same as Pawlak approximation space, it is also to have many similar properties. For details, please refer to literature [11].

Example 1: Table 1 gives an IFIS. As a result, by the definition of the dominance relation, we have $[x_1]_{AT}^{\geq}=\{x_1\}$, $[x_2]_{AT}^{\geq}=\{x_1,x_2,x_6\}$, $[x_3]_{AT}^{\geq}=\{x_3,x_8,x_9\}$, $[x_4]_{AT}^{\geq}=\{x_4,x_8,x_{10}\}$, $[x_5]_{AT}^{\geq}=\{x_5,x_9\}$, $[x_6]_{AT}^{\geq}=\{x_4,x_6,x_9\}$, $[x_7]_{AT}^{\geq}=\{x_4,x_5,x_7,x_9\}$, $[x_8]_{AT}^{\geq}=\{x_8\}$, $[x_9]_{AT}^{\geq}=\{x_9\}$, $[x_{10}]_{AT}^{\geq}=\{x_1,x_{10}\}$.

Furthermore, if *A*={a_1,a_2,a_3}, *B*={a_1,a_2} then $[x_1]_A^{\geq}=[x_2]_A^{\geq}=\{x_1,x_2,x_6\}$, $[x_3]_A^{\geq}=\{x_1,x_2,x_3,x_5,x_6,x_8,x_9\}$, $[x_4]_A^{\geq}=\{x_1,x_4,x_8,x_{10}\}$, $[x_5]_A^{\geq}=\{x_5,x_9\}$, $[x_6]_A^{\geq}=\{x_4,x_6,x_9\}$, $[x_7]_A^{\geq}=\{x_4,x_5,x_7,x_9\}$, $[x_8]_A^{\geq}=\{x_1,x_8\}$, $[x_9]_A^{\geq}=\{x_9\}$, $[x_{10}]_A^{\geq}=\{x_1,x_{10}\}$, as well as $[x_1]_B^{\geq}=[x_2]_B^{\geq}=\{x_1,x_2,x_6,x_8\}$, $[x_3]_B^{\geq}=\{x_1,x_2,x_3,x_5,x_6,x_8,x_9\}$, $[x_4]_B^{\geq}=\{x_1,x_4,x_8,x_{10}\}$, $[x_5]_B^{\geq}=\{x_1,x_2,x_3,x_5,x_6,x_8,x_9\}$, $[x_6]_B^{\geq}=\{x_1,x_2,x_4,x_6,x_8,x_9,x_{10}\}$, $[x_7]_B^{\geq}=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}$, $[x_8]_B^{\geq}=\{x_1,x_8\}$, $[x_9]_B^{\geq}=\{x_1,x_8,x_9\}$, $[x_{10}]_B^{\geq}=\{x_1,x_{10}\}$.

In this section, knowledge granulation and rough entropy in an IFIS are introduced. They have some very useful properties. The relationship between knowledge granulation and rough entropy in IFIS is established.

Definition 2: The granulation of knowledge *B*⊆*AT* is defined as follows, for an IFIS *S*=<*U*,*AT*,*V*,*f*>

$$GK(B)=\sum_{i=1}^{n}\frac{\left|[x_i]_B^{\geq}\right|}{|U|^2} \tag{6}$$

Example 2: For example 1, one can calculate the knowledge granulation of knowledge *A* and *B*:

GK(*A*)=1/100(3+3+7+4+2+3+4+2+1+2)=0.31,*GK*(*B*)=1/100(4+4+7+4+7+7+10+2+3+2) =0.50.

3. Algorithm

Definition 3: Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS, *B*⊆*AT, U*={x_1,x_2,…,x_n}.The dominance matrix of system *S* with respect to *B* is defined as:

$$M_B=(m_{ij})_{n\times n}=\begin{cases}1, & x_j\in[x_i]_B^{\geq}\\ 0, & others\end{cases}\quad i,j=1,2,\ldots,n. \tag{7}$$

M_B is also called *l* level dominance matrix of *S* if |*B*|=*l*.

Definition 4: The intersection of the dominance matrices M_B and M_C is defined as follows, for any *B,C*⊆*AT* on the *S*=<*U*,*AT*,*V*,*f*>,

$$M_B\cap M_C=(m_{ij})_{n\times n}\cap(m'_{ij})_{n\times n}=(\min\{m_{ij},m'_{ij}\})_{n\times n} \tag{8}$$

Property 1: Given *S*=<*U*,*AT*,*V*,*f*> and *B,C*⊆*AT*, if M_B,M_C are two dominance matrices, we have that:

(1) m_{ii} =1,*i=1,2,…,n*;

(2) if *B,C*⊆*AT,* then $M_{B\cup C}=M_B\cap M_C$.

Property 1can be obtained directly from Definition 3 and 4.

Definition 5: Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS, *B*⊆*AT,* a dominance matrix M_B with *B*.|M_B| is the dominance cardinality of *B* if it indicated that the number of non-zero elements, that is, a total number of 1 values.

Theorem 1: Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS; *B*⊆*AT*. Then |M_B|=$\sum_{i=1}^{n}\left|[x]_B^{\geq}\right|$,

$$GK(B)=\frac{1}{|U|^2}|M_B|\cdot \quad (9)$$

Proof: According to the definition of the dominance matrix, for *B*⊆*AT*, we have that $(m_{i_1},m_{i_2},\cdots,m_{i_n})$ corresponds to $(x'_{i_1},x'_{i_2},\cdots,x'_{i_n})=[x_i]_B^{\geq}$, where $x'_{i_k}=\begin{cases}x_j, & x_j\in[x_i]_B^{\geq}\\ \varnothing, & others\end{cases}$.

Therefore, |M_B|=$\sum_{i=1}^{n}\left|[x]_B^{\geq}\right|$. It holds that $GK(B)=\frac{1}{|U|^2}|M_B|$. This completes the proof.□

Definition 6: Given two *n*-dimensional *n×1* vectors $\alpha=(e_1,e_2,\cdots,e_n)^T$ and $\beta=(b_1,b_2,\cdots,b_n)^T$, *T* said the transpose, α is smaller than β if $e_i\leq b_i$ (*i*=1,…,*n*).

Definition 7: Let $M_A=(\alpha_1,\alpha_2,\cdots,\alpha_n)^T$ and $M_B=(\beta_1,\beta_2,\cdots,\beta_n)^T$ be a matrix, where α_i and β_i (*i*=1,…,*n*) are *n*-dimensional *n×1* vectors. M_A is smaller than M_B if $\alpha_i\leq\beta_i$ (*i*=1,…,*n*) , denoted $M_A\leq M_B$.

Definition 8: Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS. *B*⊆*AT* is a reduct of *AT* if *GK*(*B*)= *GK*(*AT*). If there is not *b*∈*B* makes *GK*(*B*-{*b*})= *GK*(*AT*), claimed that *B* is one of the maximum reduct about *AT*.

Let *S*=<*U*,*AT*,*V*,*f*> be an IFIS, *U*={$x_1,x_2,\ldots,x_n$}, *AT*={$a_1,a_2,\ldots,a_m$}, *B*⊆*AT*, $M_B=(\beta_1,\beta_2,\cdots,\beta_n)^T$ and $M_{AT}=(\gamma_1,\gamma_2,\cdots,\gamma_n)^T$. Based on the idea of the above subsection, a greedy algorithm for computing reduct can be constructed.

Table 2. An IFIS for Example.

U	a_1	a_2	a_3	a_4
x_1	0.1	0.2	0.1	0.1
x_2	0.1	*	0.3	0.1
x_3	0.3	0.2	0.3	*
x_4	0.1	0.2	*	0.1
x_5	*	0.2	0.1	0.3
x_6	0.3	0.1	*	0.3
x_7	0.3	0.2	*	*
x_8	0.3	0.1	0.2	0.3
x_9	0.2	0.3	*	0.2

Algorithm for calculating reduct:

Input: IFIS *S*=<*U*,*AT*,*V*,*f*>.

Output: One reduct *B* of *AT*.

Step 1. Compute the dominance matrix $M_{AT}=(\gamma_1,\gamma_2,\cdots,\gamma_n)^T$ of *AT*.

Step 2. Compute the first level matrix for every $a_l\in AT(1\leq l\leq m)$:

$M_{\{a_l\}}=M_{\{a_l\}}^{(1)}=(\tau_1^{(1)},\tau_2^{(1)},\cdots,\tau_n^{(1)})^T$.

For *i*= *1* to *n* do

If $0\neq\tau_i^{(1)}\leq\gamma_i$, then let $\tau_i^{(1)}=0$, and the new matrix is denoted by $FM_{\{a_i\}}^{(1)}$, $FM_{\{a_i\}}^{(1)}$= $(\tau_1^{(1)},\tau_2^{(1)},\cdots,\tau_n^{(1)})^T$,$a_l\in AT$,(1≤*l*≤*m*) is called the first-level reduct matrix;

Come into the next step.

Step 3. If $FM_{\{a_i\}}^{(1)}$=0 then output a first-level reduct {a_l}.Otherwise, enter the next step.

Step 4. All second-level dominance matrices are obtained by intersection of the non-0 first-level reduct matrices on Step 2: $M^{(2)}_{\{a_l a_s\}}$, $M^{(2)}_{\{a_l a_s\}} \neq M^{(1)}_{\{a_l\}}$, $M^{(2)}_{\{a_l a_s\}} \neq M^{(1)}_{\{a_s\}}$, $l \neq s, l, s = 1,2,\ldots,n$.

Find all of the second-level reducts by the method used in *Step 2*.

Step 5. Repeat *Step 4* to obtain to the third-level and more reducts, until $M^{(k)}_B = 0$ $(1 \leq k \leq m), B \subseteq AT$.

The time complexity of this algorithm is $O(|U|^2 2^{|A|})$.

Example 3: Table 2 provides an IFIS $S=<U,AT,V,f>$, where $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9\}$, $AT=\{a_1,a_2,a_3,a_4\}$. We run our matrix algorithm on Table 2 in this subsection, to demonstrate its simplicity, practicability and time efficiency.

Step 1: Construct the dominance matrices.

$$M_{AT} = \begin{pmatrix} 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 0\,1\,1\,1\,0\,1\,1\,0\,1 \\ 0\,0\,1\,0\,0\,0\,1\,0\,0 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,0\,1\,0\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0\,1 \end{pmatrix} \quad M_{\{a_1\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,1 \end{pmatrix} \quad M_{\{a_2\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,0\,0\,0\,0\,0\,0\,1 \end{pmatrix}$$

$$M_{\{a_3\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,1\,1\,0\,1\,1\,0\,1 \\ 0\,1\,1\,1\,0\,1\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,1\,1\,0\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \end{pmatrix} \quad M_{\{a_4\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,1 \end{pmatrix}$$

Steps 2 and 3: Construct the first-level reduct matrices and output reduction.

Compare their rows of $M_{\{a_1\}}$, $M_{\{a_2\}}$, $M_{\{a_3\}}$ and $M_{\{a_4\}}$ to M_{AT}'s . Find that there is no $0 \neq \tau^{(1)}_i \leq \gamma_i$ .Row 1,4 and 5 are same for $M_{\{a_2\}}$ and M_{AT}. Row 2 is same for $M_{\{a_3\}}$ and M_{AT}. Row 6 and 8 are same for $M_{\{a_4\}}$ and M_{AT}. Therefore, we can see that there is no first-level reduct.

Thus, the first-level reduction matrices are:

$$FM^{(1)}_{\{a_1\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,1 \end{pmatrix} \quad FM^{(1)}_{\{a_2\}} = \begin{pmatrix} 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,0\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,0\,0\,0\,0\,0\,0\,1 \end{pmatrix} \quad FM^{(1)}_{\{a_3\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,1\,1\,0\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,1\,1\,1\,0\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \end{pmatrix}$$

$$FM_{\{a_4\}} = \begin{pmatrix} 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1\,1\,1\,1\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,1\,1\,1\,1\,1 \end{pmatrix}$$

Step 4 and 5: Get the second-level and more dominance matrices, reduction.

$$M^{(2)}_{\{a_1,a_2\}} = FM^{(1)}_{\{a_1\}} \bigcap FM^{(1)}_{\{a_2\}} = \begin{pmatrix} 00000000\ 0 \\ 111111111 \\ 001\ 0\ 10100 \\ 000\ 000000 \\ 000\ 000000 \\ 00\ 1\ 011110 \\ 00\ 1\ 0\ 10100 \\ 00\ 1\ 0\ 11110 \\ 00\ 00\ 0000\ 1 \end{pmatrix} \quad M^{(2)}_{\{a_1,a_3\}} = FM^{(1)}_{\{a_1\}} \bigcap FM^{(1)}_{\{a_3\}} = \begin{pmatrix} 111111111 \\ 0000000\ 0\ 0 \\ 0010\ 0\ 1\ 1\ 0\ 1 \\ 111111111 \\ 111111111 \\ 001\ 0\ 1111\ 0 \\ 00\ 1\ 0\ 1111\ 0 \\ 00\ 1\ 0\ 0\ 111\ 0 \\ 00\ 1\ 0\ 11111 \end{pmatrix}$$

$$M^{(2)}_{\{a_1,a_4\}} = FM^{(1)}_{\{a_1\}} \bigcap FM^{(1)}_{\{a_4\}} = \begin{pmatrix} 111111111 \\ 111111111 \\ 001\ 0\ 11110 \\ 111111111 \\ 00000000\ 0 \\ 00000000\ 0 \\ 00\ 1011111\ 0 \\ 0000\ 0000\ 0 \\ 00\ 10\ 0\ 1111 \end{pmatrix} \quad M^{(2)}_{\{a_2,a_3\}} = FM^{(1)}_{\{a_2\}} \bigcap FM^{(1)}_{\{a_3\}} = \begin{pmatrix} 000000000 \\ 000000000 \\ 011\ 100101 \\ 00000000\ \ 0 \\ 00000000\ \ 0 \\ 111111111 \\ 11111\ 0\ 1\ 0\ 1 \\ 0\ 111\ 0\ 1111 \\ 0\ 10000\ \ 0\ 0\ 1 \end{pmatrix}$$

$$M^{(2)}_{\{a_2,a_4\}} = FM^{(1)}_{\{a_2\}} \bigcap FM^{(1)}_{\{a_4\}} = \begin{pmatrix} 000000000 \\ 111111111 \\ 11111\ 0\ 1\ 0\ 1 \\ 00000000\ \ 0 \\ 0000000\ \ 0\ 0 \\ 000\ 0000\ 0\ 0 \\ 11111\ 0\ 1\ 0\ 1 \\ 000000\ \ 0\ 0\ 0 \\ 0000000\ \ 0\ 1 \end{pmatrix} \quad M^{(2)}_{\{a_3,a_4\}} = FM^{(1)}_{\{a_3\}} \bigcap FM^{(1)}_{\{a_4\}} = \begin{pmatrix} 111111111 \\ 00000000\ \ 0 \\ 0111\ 0\ 1101 \\ 111111111 \\ 000000\ \ 00\ 0 \\ 0000000\ \ 0\ 0 \\ 111111111 \\ 0000\ \ 0000\ \ 0 \\ 0\ 0\ 1\ 0\ 11111 \end{pmatrix}$$

Compared with M_{AT}, there is no second-level reduct. By using the same method we can get $FM^{(2)}_{\{a_1,a_2,a_3\}}$. Each row of $FM^{(2)}_{\{a_1,a_3,a_4\}}$ are 0. The looping can now be stopped.

Therefore, we can see that a reduct of AT is $\{a_1,a_2,a_3\}$.

As can be seen from the discussed above, obtaining reducts from a large data table for even very complex one, is a relatively easy task.

4. Results of Experiments

In Table 3-7 results of experiments on six well-known data sets from the UCI Machine Learning Repository [14] are cited. The Matrix and Revised-quickreduct algorithm [15] have been implemented using MATLAB for the databases. Before the experiment, we done the preprocessing that the data is limited to between 0-1.

From the table, it is evident that Matrix algorithm produces minimal reduct for large data sets with more number of attributes. The performance analysis of the Matrix and the Revised-quickreduct is also depicted in Figure 1-5. (Note,at.=attributes,mv.=missing values, Mx.=Matrix)

Table 3. Comparative Analysis of Indiscernibility Relation Method

Data set	Instances	No.of at.	No.of mv.	Mx.	Revised-quickreduct
Car	8	4	5	2	2
Hepatitis	155	19	167	5	6
Heart (Switzerland)	123	13	273	5	4
Soybean (Large)	307	35	705	11	13
Water-treatment-data	527	38	591	7	8
Rchocardiogram	74	13	132	3	2

Table 4. Comparative Analysis of Mean Imputation 1

Data set	Instances	No.of at.	No.of mv.	Mx.	Revised-quickreduct
Car	8	4	5	2	2
Hepatitis	155	19	167	5	5
Heart (Switzerland)	123	13	273	6	5
Soybean (Large)	307	35	705	10	11
Water-treatment-data	527	38	591	5	6
Rchocardiogram	74	13	132	3	2

Table 5. Comparative Analysis of Mean Imputation 2

Data set	Instances	No.of at.	No.of mv.	Mx.	Revised-quickreduct
Car	8	4	5	3	3
Hepatitis	155	19	167	5	6
Heart (Switzerland)	123	13	273	6	5
Soybean (Large)	307	35	705	11	12
Water-treatment-data	527	38	591	7	6
Rchocardiogram	74	13	132	2	2

Table 6. Comparative Analysis of Median Imputation

Data set	Instances	No.of at.	No.of mv.	Mx.	Revised-quickreduct
Car	8	4	5	2	2
Hepatitis	155	19	167	5	5
Heart (Switzerland)	123	13	273	6	5
Soybean (Large)	307	35	705	11	12
Water-treatment-data	527	38	591	5	6
Rchocardiogram	74	13	132	3	2

Table 7. Comparative Analysis of Mode Imputation

Data set	Instances	No.of at.	No.of mv.	Mx.	Revised-quickreduct
Car	8	4	5	2	2
Hepatitis	155	19	167	5	6
Heart (Switzerland)	123	13	273	6	5
Soybean (Large)	307	35	705	9	10
Water-treatment-data	527	38	591	7	7
Rchocardiogram	74	13	132	2	2

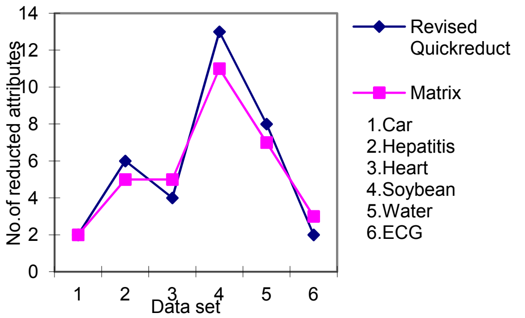

Figure 1. Performance analysis of the Matrix and the Revised Quickreduct (Indiscernibility)

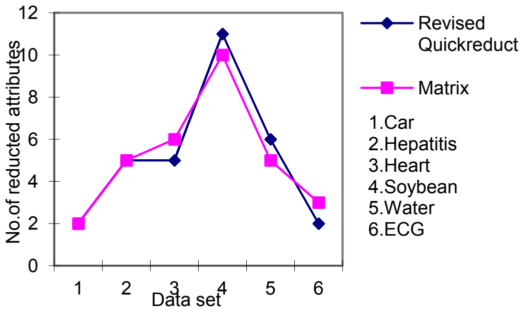

Figure 2. Performance Analysis of the Matrix and the Revised Quickreduct (Mean Imputation1)

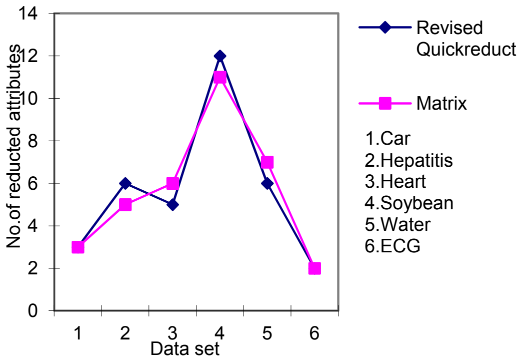

Figure 3. Performance Analysis of the Matrix and the Revised Quickreduct (Mean Imputation 2)

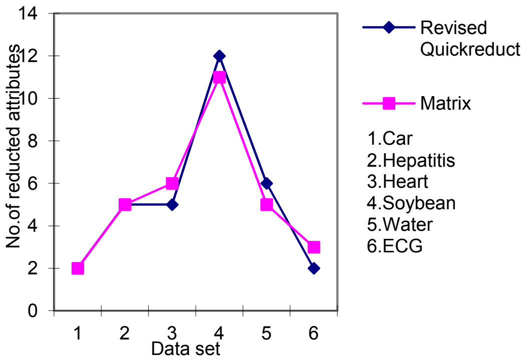

Figure 4. Performance Analysis of the Matrix and the Revised Quickreduct (Median)

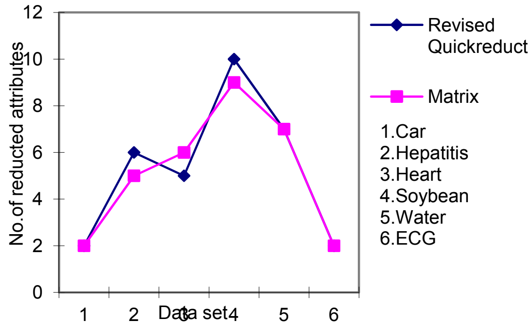

Figure 5. Performance Analysis of the Matrix and the Revised Quickreduct (Mode)

5. Conclusion

IFIS is a special information system with both fuzzy knowledge and uncertainty. In this paper, it gives a new definition of some basic concepts of rough set in IFIS by dominance relation. On this basis, we studied the IFIS knowledge granularity and advantages matrix, made some important conclusions. A IFIS attribute reduction algorithm based on the dominance matrix is built. The next step will be to use the knowledge reduction algorithm to obtain fuzzy rules, decision analysis.

References

[1] Pawlak Z. Rough sets. *International Journal of Computer and Information Sciences*.1982; 11(5): 341-356.
[2] Pawlak Z, Skowron A. Rudiments of rough sets. *Information Sciences*. 2007; 177(1): 3-27.
[3] Kryszkiewicz M. Rough set approach to incomplete information systems. *Information Sciences*. 1998; 112(1/2/3/4): 39-49.
[4] Kryszkiewicz M. Rules in incomplete information systems. *Information Sciences*. 1999; 113(3/4): 271-292.
[5] Dubols D, Prade H. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*.1990; 7(2/3): 191-209.
[6] Wu WZ, Mi JS, Zhang WX. Generalized fuzzy rough sets. *Information Sciences*. 2003; 151: 263-282.
[7] Yao YY. A comparative study of fuzzy sets and rough sets. *Information Sciences*. 1998; 109(1/2/3/4): 227-242.
[8] Chen DG, Zhao SY. Local reduction of decision system with fuzzy rough sets. *Fuzzy Sets and Systems.* 2010; 161(13): 1871-1883.
[9] Greco S, Matarazzo B, Slowinski R. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*. 2001; 129(1): 1-47.
[10] Greco S, Matarazzo B, Slowinski R. Rough approximation by dominance relations. *International Journal of Intelligent systems*. 2002; 17(2): 153-171.
[11] Greco S, Inuiguchi M, Slowinshi R. Fuzzy rough sets and multiple-premise gradual decision rules. *International Journal of Approximate Reasoning*. 2006; 41(2): 179-211.
[12] Wei LH, Tang ZM, Yang XB, Zhang LL. *Dominance-based Rough Set Approach to Incomplete Fuzzy Information System.* Proceedings of the 2008 IEEE International Conference on Granular Computing, GrC. 2008; 1: 632–637.
[13] Yang XB, Wei LH, Yu DJ, Yang JY. Rough Fuzzy Set in Incomplete Fuzzy Information System Based on Similarity Dominance Relation. *Recent Patents on Computer Science*. 2009; 2: 68-74.
[14] Blake CL, Merz CJ. UCI Repository of machine learning databases, Irvine, University of California. 1998: Availaible on:http://www.ics.uci.edu/ mlearn/.
[15] Thangavel K, Pethalakshmi A, Jaganathan P. A Novel Reduct Algorithm for Dimensionality Reduction with Missing Values Based on Rough Set Theory. *International Journal of Soft Computing.* 2006; 1(2): 111-117.
[16] Gu SM, Sun XH, Wu WZ. *Rough Set Approximations in Incomplete Multi-scale Information Systems, Rough Sets and Knowledge Technology*. Lecture Notes in Computer Science. 2012; 7414: 342-350.
[17] Zhang HY, Leung Y, Zhou L. Variable-precision-dominance-based rough set approach to interval-valued information systems. *Information Sciences.* 2013; 244(20): 75–91.