International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:2, 2017

An Improved Limited Tolerance Rough Set Model

Chen Wu, Komal Narejo, Dandan Li

Abstract**—**Some extended rough set models in incomplete information system cannot distinguish the two objects that have few known attributes and more unknown attributes; some cannot make a flexible and accurate discrimination. In order to solve this problem, this paper suggests an improved limited tolerance rough set model using two thresholds to control what two objects have a relationship between them in limited tolerance relation and to classify objects. Our practical study case shows the model can get fine and reasonable decision results.

Keywords**—**Decision rule, incomplete information system, limited tolerance relation, rough set model.

I. Introduction

IN recent years, rough set model [1]-[3], as a component of hybrid solutions in data mining and machine learning, have been successfully applied in bioinformatics, software engineering, robotics, medicine, multimedia, web and text mining, economics and finance, signal and image processing, and engineering (e.g. power systems and control engineering). People need to know more about rough set to deal with practical problems.

Gore also notices that we never have complete information yet we have to make decisions anyway [4]. This means that making decisions with partial information is ultimately inevitable. So how to process incomplete information system (IIS) with missing attribute values in the process of knowledge acquisition became a hot topic nowadays. There are three types of missing attribute values, the first of them is lost values, in this case the values were recorded but currently are unavailable; the second is attribute-concept values, which means that missing values may be substituted using any value in the same attribute; the last is the 'do not care' condition which means that the original values are irrelevant.

For the moment, two strategies data reparation and model extension have been applied to deal with IIS [5]. The former is an indirect method which transforms an IIS into a complete information system. The latter is a direct method which expands basic concepts in the classical rough set theory to IIS. Because the first approach may change the original information, most experts now pay their attention to study IIS by using the second approach.

A tolerance relation rough set model is put forward and researched in [5], as a pioneer work in the IIS processing area.

Chen Wu is with the the School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang City, Jiangsu Province, China (corresponding author; phone: +8613861390697; e-mail: wuchenzj@ sina.com).

Komal Narejo and Dandan Li are with the the School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang City, Jiangsu Province, China (e-mail: 3533202963@qq.com, simpler@sina.com).

Similarity rough set model is proposed and studied in [6], [7]. Many variable precision rough set models are also suggested and inspected. A limited tolerance rough set model is explored in [8]. In [9], [10], an improved limited rough set model which distinguishes two objects by setting a threshold is put forward. But it is a pity that it only works as an irreflexive relation. A symmetry and improved limited rough set model is suggested in [11], which works well for it is a both reflexive and symmetric relation. The last two models make much progress in dealing with incomplete information system but also have limited ability to distinguish two objects which have few known attributes and more unknown attributes. Another limited rough set model, proposed in [12], overcomes some shortcomings in several models. Even now some experts are studying multi-granular rough set model by generalizing rough set theory [13]-[15]. Reductions in incomplete information system are also hot topics [16].

In the present paper, we are going to discuss about extending limited tolerance rough set model. We find that in [12], it is hard for the model to distinguish two objects with a big precision. For example, the probability of the object *A* and object *B* is 1/10; the probability of the object *A* and object *C* is 1/1000; the model may classify object *B* and object *C* into the same indiscernibility class of object *A*. In order to solve this problem, this paper presents an improved limited tolerance set model with two thresholds, which can distinguish two objects with an accurate precision.

II. Basic Concepts

An IIS can be denoted as a four-tuple *IIS*=<*U*,*AT*,*V*,*f*>,where *U* is called the universe of discourse, being a non-empty set of finite objects. *AT* is a non-empty set of finite attributes. V_a is domain of the attribute a (usually with "*" to represent unknown attribute), set $V=\cup_{a\in AT} V_a$, f is an information function, that is, for $\forall a\in AT$, $\forall x\in U$, then $f(x,a)\in V_a$. If the domain V_a of the information table *IIS* contains at least one attribute a whose value is * for some object, then we say that the information system IIS is incomplete, otherwise *IIS* is complete.

Definition 1. Let *IIS* be an incomplete information system. $\forall A\subseteq AT$. The variable precision classification relation in terms of *A* is denoted by $V^{\alpha}(A)$ [6]:

$$V^{\alpha}(A)=\{(x,y)\in U^2 \mid \forall a\in P_A(x) \cap P_A(y),\, f(x,a)=f(y,a) \wedge \frac{|P_A(x)\cap P_A(y)|}{|P_A(x)|}\geq \alpha\}\cup I_U$$

where $\alpha \in [0,1]$, $P_A(x)=\{a : a \in A, f(x,a) \neq *\}$, $|\cdot|$ represents the cardinality of a set, and $I_U = \{(x,x) | x \in U\}$.

It is easy to find that $V^{\alpha}(A)$ meets only reflexive, but is not necessarily symmetric and transitive. In the Definition 2, we thought x={*,1,*,2,3,*,1,*} and y={1,*,0,*,*,*,1,*} belong to the same class. But x and y have only one attribute with the same value. Therefore, it is impossible that both of them belong to the same class and putting them into a class becomes very far-fetched.

Definition 2. Let *IIS* be an incomplete information system. $\forall A \subseteq AT$, $\forall X \subseteq U$. The upper approximation and lower approximation of X in terms of $V^{\alpha}(A)$ are denoted by $\overline{V_A^{\alpha}}(X)$ and $\underline{V_A^{\alpha}}(X)$:

$$\overline{V_A^{\alpha}}(X) = \{x \in U \mid V_A^{\alpha}(x) \cap X \neq \varnothing\}$$
$$\underline{V_A^{\alpha}}(X) = \{x \in U \mid V_A^{\alpha}(x) \subseteq X\}$$

where $V_A^{\alpha}(x) = \{y \in U \mid (x,y) \in V^{\alpha}(A)\}$.

An extended limited tolerance rough set model with symmetry is given in Definition 3.

Definition 3. Let *IIS* be an incomplete information system and *IIS*=<*U*,*AT*=*C* ∪ {*d*},*V*,*f*>, *x*,*y* ∈ *U*, *B* ⊆ *C*, $0 \le \alpha \le 1$, $\mu(x,y) = \frac{|P_A(x) \cap P_A(y)|}{|P_A(x) \cup P_A(y)|}$. The binary relation L_B^{α} is called the limited and variable precision tolerance relation, where

$$L_B^{\alpha} = \{(x,y) \in U^2 \mid (x,y) \in L_B, \; f(x,a) = f(y,a) \wedge \mu(x,y) \ge \alpha\}$$

We can see that the limited and variable precision tolerance relation is reflexive and symmetric, which overcame the shortcomings in the variable precision rough set model in [10]. But for any two objects x={1,*,*,*,*,*,*,*,*,*} and y={1,*,*,*,*,*,*,*,*,*}, in Definition 1 and Definition 3, no matter what value α takes, they are regarded as the same class. So another limited rough set model is proposed in [12] as following definitions.

Definition 4. Let *IIS* be an incomplete information system in which $\forall A \subseteq AT$, the limited relation in terms of *A* is denoted by $SV^{\alpha}(A)$:

$$SV^{\alpha}(A) = \{(x,y) \in U^2 \mid \forall a \in P_A(x) \cap P_A(y), f(x,a) = f(y,a) \wedge \frac{|P_A(x) \cap P_A(y)|}{|A|} \ge \alpha\} \cup I_U$$

where |·| is the function to calculate the cardinality of a set, $\alpha \in [0,1]$ is a threshold; $I_U = \{(x,x) | x \in U\}$ is the identity relation; $P_A(x) = \{a : a \in A, f(x,a) \neq *\}$ is the number of attribute in *A* for *x* being not * at the attribute.

Definition 5. Let *IIS* be an incomplete information system in which $\forall A \subseteq AT$, $\forall X \subseteq U$, then the upper approximation and lower approximation of *X* in terms of *A* are denoted by $\overline{SV_A^{\alpha}}(X)$ and $\underline{SV_A^{\alpha}}(X)$:

$$\overline{SV_A^{\alpha}}(X) = \{x \in U \mid SV_A^{\alpha}(x) \cap X \neq \varnothing\}$$
$$\underline{SV_A^{\alpha}}(X) = \{x \in U \mid SV_A^{\alpha}(x) \subseteq X\}$$

where $SV_A^{\alpha}(x) = \{y \in U \mid (x,y) \in SV^{\alpha}(A)\}$.

In Definition 4 and Definition 5, we can see the limited and variable precision rough set model as well as its upper approximation and lower approximation. It is obvious that $SV^{\alpha}(A)$ is reflexive and symmetric, but not transitive. And if $a = 1 \subseteq A$, the limited and variable precision rough set model can distinguish object x={1,*,*,*,*,*,*,*,*,*} and object y={1,*,*,*,*,*,*,*,*,*} by setting a higher precision.

III. An Improved Limited RSM

According to the model in [9], [10], for three objects x={1,2,*,*}, y={*,2,*,*} and z={*,2,1,3}, to a certain degree, objects *y* and *z* may be classified into the indiscernibility class of *x*. It means that the probability of that they are the same to *x* is identical. Actually, both the object *y* and the object *z* have different known attributes as well as unknown attributes, so we need a more accurate classification to distinguish them. In order to solve this problem, this paper proposes a new improved limited tolerance rough set model with two thresholds related to the filling factors of the system table to form limited tolerance relation as in Definition 6.

Definition 6. Let IIS be an incomplete information system. $\forall A \subseteq AT$. The limited tolerance relation in terms of *A* is denoted by $IV^{\alpha,\beta}(A)$:

$$IV^{\alpha,\beta}(A) = \{(x,y) \in U^2 \mid \forall a \in P_A(x) \cap P_A(y), f(x,a) = f(y,a) \wedge \frac{|P_A(x) \cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}} \ge \alpha \wedge \frac{\min\{|P_A(x)|,|P_A(y)|\}}{|A|} \ge \beta\} \cup I_U$$

where α, $\beta \in [0,1]$, $P_A(x) = \{a : a \in A, f(x,a) \neq *\}$, $|\cdot|$ represents the cardinality of the set, and $I_U = \{(x,x)|x \in U\}$.

Definition 7. Let *IIS* be an incomplete information system. $\forall A \subseteq AT$, $\forall X \subseteq U$. The upper approximation and lower approximation of *X* in terms of $IV^{\alpha,\beta}(A)$ are denoted by $\overline{IV_A^{\alpha,\beta}}(X)$ and $\underline{IV_A^{\alpha,\beta}}(X)$:

$$\overline{IV_A^{\alpha,\beta}}(X)=\{x\in U \mid IV_A^{\alpha,\beta}(x)\cap X\neq\varnothing\}$$
$$\underline{IV_A^{\alpha,\beta}}(X)=\{x\in U \mid IV_A^{\alpha,\beta}(x)\subseteq X\}$$

where $IV_A^{\alpha,\beta}(x)=\{y\in U \mid (x,y)\in IV^{\alpha,\beta}(A)\}$.

Theorem 1. Let *IIS* be an incomplete information system. $\forall A\subseteq AT$. We can get:

$$IV^{0,0}(A)=\{(x,y)\in U^2 \mid \forall a\in P_A(x)\cap P_A(y), f(x,a)=f(y,a)\wedge\frac{|P_A(x)\cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}}\geq 0\wedge\frac{\min\{|P_A(x)|,|P_A(y)|\}}{|A|}\geq 0\}\cup I_U$$
$$\Leftrightarrow \{(x,y)\in U^2 \mid \forall a\in P_A(x)\cap P_A(y), f(x,a)=f(y,a)\}\cup I_U$$

$$IV^{1,1}(A)=\{(x,y)\in U^2 \mid \forall a\in P_A(x)\cap P_A(y), f(x,a)=f(y,a)\wedge\frac{|P_A(x)\cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}}\geq 1\wedge\frac{\min\{|P_A(x)|,|P_A(y)|\}}{|A|}\geq 1\}\cup I_U$$
$$\Leftrightarrow \{(x,y)\in U^2 \mid \forall a\in P_A(x)\cap P_A(y), f(x,a)=f(y,a)\wedge|P_A(x)|=|P_A(y)|=|A|\}$$

Actually, $IV^{0,0}(A)$ is equivalent to the tolerance relation in [6]. $IV^{1,1}(A)$ is an absolute equivalence relation. $\forall x,y\in IV^{1,1}(A)$, they cannot have null attribute values and each attribute value of them in attribute set A is identical which illustrates that $IV^{1,1}(A)$ has stricter constraint. From the Theorem 1, it is easy to conclude that the new improved limited tolerance rough set model with two parameters has a wider change scope.

Theorem 2. Let *IIS* be an incomplete information system. $\forall A\subseteq AT$. If $0\leq\alpha_1\leq\alpha_2\leq 1$ and $0\leq\beta_1\leq\beta_2\leq 1$, for $\forall X\subseteq U$, then

$$IV_A^{\alpha_2,\beta_2}(x)\subseteq IV_A^{\alpha_1,\beta_1}(x)$$
$$\overline{IV_A^{\alpha_2,\beta_2}}(X)\subseteq\overline{IV_A^{\alpha_1,\beta_1}}(X)$$
$$\underline{IV_A^{\alpha_1,\beta_1}}(X)\subseteq\underline{IV_A^{\alpha_2,\beta_2}}(X)$$

Proof.

(i) $\forall y\in IV_A^{\alpha_2,\beta_2}(x)$, we have

$$\frac{|P_A(x)\cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}}\geq\alpha_2,\ \frac{\min\{|P_A(x)|,|P_A(y)|\}}{|A|}\geq\beta_2,$$

and $\alpha_1\leq\alpha_2$, $\beta_1\leq\beta_2$, that is $y\in IV_A^{\alpha_1,\beta_1}(x)$.

(ii) $\forall y\in\underline{IV_A^{\alpha_1,\beta_1}}(X)$, we have $IV_A^{\alpha_1,\beta_1}(y)\subseteq X$, besides we also have $IV_A^{\alpha_2,\beta_2}(x)\subseteq IV_A^{\alpha_1,\beta_1}(x)$, so $IV_A^{\alpha_2,\beta_2}(x)\subseteq X$, that is $y\in IV_A^{\alpha_2,\beta_2}(x)$.

(iii) $\forall y\in\overline{IV_A^{\alpha_2,\beta_2}}(x)$, we have $IV_A^{\alpha_2,\beta_2}(x)\bigcap X\neq\varnothing$, because of $IV_A^{\alpha_2,\beta_2}(x)\subseteq IV_A^{\alpha_1,\beta_1}(x)$, we can get that $IV_A^{\alpha_1,\beta_1}(x)\bigcap X\neq\varnothing$, that is $y\in\overline{IV_A^{\alpha_1,\beta_1}}(x)$.

Theorem 3. Let IIS be an incomplete information system. $\forall A\subseteq AT$, $\forall X,Y\subseteq U$. Then

$$\underline{IV_A^{\alpha,\beta}}(X\cap Y)=\underline{IV_A^{\alpha,\beta}}(X)\cap\underline{IV_A^{\alpha,\beta}}(Y)$$
$$\underline{IV_A^{\alpha,\beta}}(X\cup Y)\supseteq\underline{IV_A^{\alpha,\beta}}(X)\cup\underline{IV_A^{\alpha,\beta}}(Y)$$
$$\overline{IV_A^{\alpha,\beta}}(X\cap Y)\subseteq\overline{IV_A^{\alpha,\beta}}(X)\cap\overline{IV_A^{\alpha,\beta}}(Y)$$
$$\overline{IV_A^{\alpha,\beta}}(X\cup Y)=\overline{IV_A^{\alpha,\beta}}(X)\cup\overline{IV_A^{\alpha,\beta}}(Y)$$

Proof.

(i) $z\in\underline{IV_A^{\alpha,\beta}}(X\bigcap Y)\quad\Leftrightarrow IV_A^{\alpha,\beta}(z)\subseteq X\bigcap Y$
$\Leftrightarrow IV_A^{\alpha,\beta}(z)\subseteq X\wedge IV_A^{\alpha,\beta}(z)\subseteq Y$
$\Leftrightarrow z\in\underline{IV_A^{\alpha,\beta}}(X)\wedge z\in\underline{IV_A^{\alpha,\beta}}(Y)$
$\Leftrightarrow z\in\underline{IV_A^{\alpha,\beta}}(X)\cap\underline{IV_A^{\alpha,\beta}}(Y)$

(ii) $z\in\underline{IV_A^{\alpha,\beta}}(X)\cup\underline{IV_A^{\alpha,\beta}}(Y)\quad\Leftrightarrow z\in\underline{IV_A^{\alpha,\beta}}(X)\vee z\in\underline{IV_A^{\alpha,\beta}}(Y)$
$\Leftrightarrow IV_A^{\alpha,\beta}(z)\subseteq X\vee IV_A^{\alpha,\beta}(z)\subseteq Y$
$\Leftrightarrow IV_A^{\alpha,\beta}(z)\subseteq X\vee IV_A^{\alpha,\beta}(z)\subseteq Y$
$\Rightarrow IV_A^{\alpha,\beta}(z)\subseteq X\cup Y\Leftrightarrow z\in\underline{IV_A^{\alpha,\beta}}(X\cup Y)$

(iii) $z\in\overline{IV_A^{\alpha,\beta}}(X\bigcap Y)\Leftrightarrow IV_A^{\alpha,\beta}(z)\cap(X\cap Y)\neq\varnothing$
$\Rightarrow IV_A^{\alpha,\beta}(z)\cap X\neq\varnothing\wedge IV_A^{\alpha,\beta}(z)\cap Y\neq\varnothing$
$\Leftrightarrow z\in\overline{IV_A^{\alpha,\beta}}(X)\wedge z\in\overline{IV_A^{\alpha,\beta}}(Y)$
$\Leftrightarrow z\in\overline{IV_A^{\alpha,\beta}}(X)\cap\overline{IV_A^{\alpha,\beta}}(Y)$

(iv) $z\in\overline{IV_A^{\alpha,\beta}}(X\bigcup Y)\quad\Leftrightarrow IV_A^{\alpha,\beta}(z)\cap(X\cup Y)\neq\varnothing$
$\Leftrightarrow IV_A^{\alpha,\beta}(z)\cap X\neq\varnothing\vee IV_A^{\alpha,\beta}(z)\cap Y\neq\varnothing$
$\Leftrightarrow z\in\overline{IV_A^{\alpha,\beta}}(X)\vee z\in\overline{IV_A^{\alpha,\beta}}(Y)$
$\Leftrightarrow z\in\overline{IV_A^{\alpha,\beta}}(X)\cup\overline{IV_A^{\alpha,\beta}}(Y)$

Theorem 4. Let *IIS* be an incomplete information system, $\forall\ A \subseteq AT$, $\forall\ X \subseteq U$, $\forall x \in U$, $\alpha, \beta \in [0,1]$, $\max\{|P_A(x)|,|P_A(y)|\} = |A|$. Then

(i) $IV_A^{\alpha,\beta}(x) \subseteq SV_A^{\alpha}(x) \subseteq V_A^{\alpha}(x)$

(ii) $\underline{IV_A^{\alpha,\beta}}(X) \subseteq X \subseteq \overline{IV_A^{\alpha,\beta}}(X)$

(iii) $\underline{V_A^{\alpha}}(X) \subseteq \underline{SV_A^{\alpha}}(X) \subseteq \underline{IV_A^{\alpha,\beta}}(X)$

(iv) $\overline{IV_A^{\alpha,\beta}}(X) \subseteq \overline{SV_A^{\alpha}}(X) \subseteq \overline{V_A^{\alpha}}(X)$

Proof. Since $SV_A^{\alpha}(x) \subseteq V_A^{\alpha}(x)$, $\underline{V_A^{\alpha}}(X) \subseteq \underline{SV_A^{\alpha}}(X)$, $\overline{SV_A^{\alpha}}(X) \subseteq \overline{V_A^{\alpha}}(X)$ are correct, we have $\underline{IV_A^{\alpha,\beta}}(X) \subseteq X \subseteq \overline{IV_A^{\alpha,\beta}}(X)$ according to the Definition 7. So only $IV_A^{\alpha,\beta}(x) \subseteq SV_A^{\alpha}(x)$, $\underline{SV_A^{\alpha}}(X) \subseteq \underline{IV_A^{\alpha,\beta}}(X)$, $\overline{IV_A^{\alpha,\beta}}(X) \subseteq \overline{SV_A^{\alpha}}(X)$ need to be proved.

(1) Firstly, when $\beta \in [\alpha,1]$, $\max\{|P_A(x)|,|P_A(y)|\} = |A|$, $\frac{|P_A(x) \cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}} \geq \alpha \Leftrightarrow \frac{|P_A(x) \cap P_A(y)|}{|A|} \geq \alpha$. Then

$$y \in IV_A^{\alpha,\beta}(x) \Leftrightarrow \forall a \in P_A(x) \cap P_A(y), f(x,a) = f(y,a) \wedge \frac{|P_A(x) \cap P_A(y)|}{\max\{|P_A(x)|,|P_A(y)|\}} \geq \alpha \wedge \frac{\min\{|P_A(x)|,|P_A(y)|\}}{|A|} \geq \alpha$$
$$\Rightarrow \forall a \in P_A(x) \cap P_A(y), f(x,a) = f(y,a) \wedge \frac{|P_A(x) \cap P_A(y)|}{|A|} \geq \alpha$$
$$\Leftrightarrow y \in SV_A^{\alpha}(x).$$

(2) By the Definition 5, we have that $y \in \underline{SV_A^{\alpha}}(X) \Leftrightarrow SV_A^{\alpha}(y) \subseteq X$. At the same time, we have $IV_A^{\alpha,\beta}(x) \subseteq SV_A^{\alpha}(x)$ by the proof of (1), so $IV_A^{\alpha,\beta}(y) \subseteq SV_A^{\alpha}(y)$ and $IV_A^{\alpha,\beta}(y) \subseteq X$, that is $y \in \underline{IV_A^{\alpha,\beta}}(X)$ by the Definition 7.

By the Definition 7, we have

$$y \in \overline{IV_A^{\alpha,\beta}}(X) \Leftrightarrow IV_A^{\alpha,\beta}(y) \cap X \neq \varnothing.$$

At the same time, we have $IV_A^{\alpha,\beta}(x) \subseteq SV_A^{\alpha}(x)$ by the proof of (1), so $IV_A^{\alpha,\beta}(y) \subseteq SV_A^{\alpha}(y)$ and $SV_A^{\alpha}(y) \cap X \neq \varnothing$. That is $y \in \overline{SV_A^{\alpha}}(X)$ by the Definition 5.

In summary, the proposed model in this paper meets all the theorems met by relations defined in Definitions 1, 2 and 3 under certain conditions. But the relationship between the model proposed in this paper and similarity relation [6] proposed by Stefanowski which is caused by the strict constraint.

TABLE I
AN INCOMPLETE INFORMATION SYSTEM

A	c_1	c_2	c_3	c_4	d
a_1	3	2	1	0	Φ
a_2	2	3	2	0	Φ
a_3	2	3	2	0	Ψ
a_4	*	2	*	1	Φ
a_5	*	2	*	1	Ψ
a_6	2	3	2	1	Ψ
a_7	3	*	*	3	Φ
a_8	*	0	0	*	Ψ
a_9	3	2	1	3	Ψ
a_{10}	1	*	*	*	Φ
a_{11}	*	2	*	*	Ψ
a_{12}	3	2	1	*	Φ

IV. Case Analysis and Rule Generation

The core of rough set theory is knowledge reduction and knowledge reasoning. After knowledge reduction, we can get simplify rules so as to guide us to make better decision. But for an information system, there are many different knowledge reductions which can generate different decision rules that are difficult to choose. So, this paper generates all the relatively determinative and probable rules without attributes reduction and uses the decision matrix in Definition 8 [10].

Definition 8. Let *IIS=<U,AT=C* $\cup$ *D,V,f>* be an incomplete information system. Here, AT is a non-empty set of finite attributes, D is a non-empty set of decision attributes, and $AT \cap D = \varnothing$, $* \notin V_D$. Let $A \subseteq AT$, $U/IND(D) = \{D_i \mid i = 1,2,...,q\}$ denoted the classification of domain. Then the decision matrix in terms of D_l is denoted by $M_{x,y}^{k}$, where

$$M_{x,y}^{k} = \begin{cases} (a, f(x,a)),\ f(x,a) \neq * \wedge f(y,a) \neq * \\ \qquad\qquad \wedge\ f(x,a) \neq f(y,a) \\ \varnothing, otherwise \end{cases}$$

where $x \in \underline{IV_A^{\alpha,\beta}}(D_k)$, $y \in U - D_k (1 \leq k \leq q)$, $a \in P_A(x)$.

Here we still use a real IIS to analyze and compare the performances of these extended rough set models. Table I is an incomplete information table which is used by Stefanowski to analyze tolerance relation, similarity relations and the similar tolerance relation in [10]. For easiness to compare, we still use this information table [10]. As the table shows, $a_1,a_2,\ldots,a_{12}$ represent all the objects; c_1,c_2,c_3,c_4 represent all the attributes and the domain is {0,1,2,3}. * represents an unknown attribute; d represents decision attribute. The objects are divided into decision classes Φ and Ψ.

(1) First we use the improved limited rough set model proposed in [10] to analyze the information table, let $\alpha = 0.5$, we can get the results in [10]. According to the

decision attribute d to classify, we can get that

$$d_\Phi=\{a_1,a_2,a_4,a_7,a_{10},a_{12}\},\ d_\Psi=\{a_3,a_5,a_6,a_8,a_9,a_{11}\}$$

and $\underline{A^{0.5}}(\Phi)=\{a_{10}\}$,

$$\overline{A^{0.5}}(\Phi)=\{a_1,a_2,a_3,a_4,a_5,a_7,a_9,a_{10},a_{11},a_{12}\},$$

$$\underline{A^{0.5}}(\Psi)=\{a_6,a_8\},$$

$$\overline{A^{0.5}}(\Psi)=\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{11},a_{12}\}.$$

The results are not optimistic. For example, objects a_4 and a_{11} have only one attribute c_2 with the same value. Two other attribute values of a_4 are unknown and all the others attributes of a_{11} are unknown, so the probability of that the two objects are the same is only $(1/4)^5=1/1024$. It is unconvincing to take them into the same class.

The symmetry and improved limited rough set model proposed in [11] also can not differentiate between a_4 and a_{11} if we set $\alpha=0.5$.

(2) We use the limited rough set model proposed in [12] to analyze the information. Let $\alpha=0.5$. We can get a result that is more reasonable than those in [9], [10]. And in [12], object a_4 and the object a_{11} are classified into different classes. The result is more accurate than the previous model. But when α is too big or small, the result is different. Let $\alpha=$ 0.25, then $SV_c^{0.25}(a_4)=\{a_4,a_5,a_{11},a_{12}\}$. It is easy to calculate that the probability for objects a_4 and a_{11} being the same is $(1/4)^5=1/1024$ and the probability for objects a_4 and a_{12} being the same is $(1/4)^3=1/64$. But, in the model in [12] objects a_{11} and a_{12} are the same as a_4.

We use our model in definition 6 to analyze the information. Let $\alpha=0.5,\beta=0.5$. We can obtain

$$IV_c^{0.5,0.5}(a_1)=\{a_1,a_{12}\},\ IV_c^{0.5,0.5}(a_2)=\{a_2,a_3\},$$

$$IV_c^{0.5,0.5}(a_3)=\{a_2,a_3\},\ IV_c^{0.5,0.5}(a_4)=\{a_4,a_5\},$$

$$IV_c^{0.5,0.5}(a_5)=\{a_4,a_5\},\ IV_c^{0.5,0.5}(a_6)=\{a_6\},$$

$$IV_c^{0.5,0.5}(a_7)=\{a_7,a_9\},IV_c^{0.5,0.5}(a_8)=\{a_8\},$$

$$IV_c^{0.5,0.5}(a_9)=\{a_7,a_9,a_{12}\},\ IV_c^{0.5,0.5}(a_{10})=\{a_{10}\},$$

$$IV_c^{0.5,0.5}(a_{11})=\{a_{11}\},\ IV_c^{0.5,0.5}(a_{12})=\{a_1,a_9,a_{12}\}.$$

$$\underline{IV_c^{0.5,0.5}}(\Phi)=\{a_1,a_{10}\},$$

$$\overline{IV_c^{0.5,0.5}}(\Phi)=\{a_1,a_2,a_3,a_4,a_5,a_7,a_9,a_{10},a_{12}\},$$

$$\underline{IV_c^{0.5,0.5}}(\Psi)=\{a_6,a_8,a_{11}\},$$

$$\overline{IV_c^{0.5,0.5}}(\Psi)=\{a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{11},a_{12}\}.$$

As you have seen, when $\alpha=\beta=0.5$, our results are the same as [12]. We can also classify the object a_4 and the object a_{11} into different class. Next, we let $\alpha=0.25,\beta=0.5$, then $IV_c^{0.25,0.5}(a_4)=\{a_4,a_5,a_{12}\}$. Here we classified the object a_{12} into the class of the object a_4. At the same time, the object a_{11} is not in the class of a_4. But in the model in [12], whatever α takes under $\alpha<0.5$, it cannot distinguish a_{11} and a_{12}, due to that it does not consider the number of known and unknown attributes. From this point of view, it is obvious that our model can classify finer and more accurate. So Φ's discernibility matrix for relatively determinative rule generation by using $x\in\underline{IV_c^{0.5,0.5}}(d_\Phi)$, $y\in U-d_\Phi$ is as in Table II. Blank grid means empty. Thus, relatively determinative rules generated from Table II are:

$$(c_2,2)\wedge(c_4,0)\rightarrow(e,\Phi);(c_3,1)\wedge(c_4,0)\rightarrow(e,\Phi);$$
$$(c_1,1)\rightarrow(e,\Phi).$$

TABLE II
DISCERNIBILITY MATRIX FOR RELATIVELY DETERMINATIVE RULE GENERATION TO Φ

	a_1	a_{10}
a_3	$(c_1,3)(c_2,2)(c_3,1)$	$(c_1,1)$
a_5	$(c_4,0)$	
a_6	$(c_1,3)(c_2,2)(c_3,1)(c_4,0)$	$(c_1,1)$
a_8	$(c_2,2)(c_3,1)$	
a_9	$(c_4,0)$	$(c_1,1)$
a_{11}	$(c_1,3)(c_2,2)(c_3,1)$	$(c_1,1)$

TABLE III
DISCERNIBILITY MATRIX FOR RELATIVELY PROBABLE RULE GENERATION TO Φ

	a_6	a_8	a_{11}
a_1	$(c_1,3)(c_2,2)(c_3,1)(c_4,0)$	$(c_2,2)(c_3,1)$	
a_2	$(c_4,0)$	$(c_2,3)(c_3,2)$	$(c_2,3)$
a_4	$(c_2,2)$	$(c_2,2)$	
a_7	$(c_1,3)(c_4,3)$		
a_{10}	$(c_1,1)$		
a_{11}	$(c_1,3)(c_2,2)(c_3,1)$	$(c_2,2)(c_3,1)$	

The results have another two extra determinative rules more than those in the model in [10] which has only rule: $(c_1,1)\rightarrow(e,\Phi)$.

Φ's discernibility matrix for relatively probable rule generation by $x\in\overline{IV_c^{0.5,0.5}}(d_\Phi)$, $y\in U-\overline{IV_c^{0.5,0.5}}(d_\Phi)$ is as in Table III. Thus, relatively probable rules generated from Table III are:

$$(c_2,2)\rightarrow(e,\Phi);(c_3,1)\rightarrow(e,\Phi);(c_2,3)\wedge(c_4,0)\rightarrow(e,\Phi);$$
$$(c_1,3)\rightarrow(e,\Phi);(c_4,3)\rightarrow(e,\Phi);(c_1,1)\rightarrow(e,\Phi).$$

Here we let $x\in\overline{IV_c^{0.5,0.5}}(d_\Phi)$, then $x\neq a_3,a_5,a_9$. In this way we can get relatively reasonable result. In fact, the relatively probable rules generated from Table III are more meaningful than the model in [10]. In the same way, we can also construct $\Psi's$ discernibility matrix for relatively determinative rule generation and get relatively probable rules for decision class Ψ . The relatively determinative rules are:

$$(c_2,3) \wedge (c_4,1) \to (e,\Psi)\,;(c_2,0) \to (e,\Psi)\,;$$
$$(c_2,2) \to (e,\Psi)\,.$$

The relatively probable rules for decision class Ψ are:

$$(c_1,2) \to (e,\Psi)\,;(c_2,3) \to (e,\Psi)\,;(c_3,2) \to (e,\Psi)\,;$$
$$(c_4,1) \to (e,\Psi)\,;(c_2,0) \to (e,\Psi)\,;(c_3,0) \to (e,\Psi)\,;$$
$$(c_4,3) \to (e,\Psi)\,;(c_1,3) \to (e,\Psi)\,.$$

V. Conclusions

In this paper, we do not mention some classical rough set theory in detail, but all the new rough set models are undoubtedly based on them. And they are widely used in the IIS in the real word. Our model in this paper has necessary connection with them as presented in Theorem 1. Owing to a large amount of incomplete information existing in the real world, different extended rough set models will be studied in now and in the future. Comparing with the rough set models proposed, our model is accurate and reasonable.

Acknowledgment

This work was partially and financially supported by a Chinese National Science Foundation (61100116) and a Foundation from Educational Bureau of Jiangsu provincial government and a Foundation of Graduate Department of Jiangsu University of Science and Technology.

References

[1] Pawlak. S. “Rough set theory and its applications to data analysis”. Journal of Cybernetics and Systems. 1998, (29):661-688.
[2] Pawlak. Z. “Rough sets and intelligent data analysis”. Journal of Information Sciences. 2002, (147): 1-12.
[3] Pawlak Z. “Rough Sets. Theoretical Aspects of Reasoning about Data”. Dordecht: Kluwer Academic Publishers. 1991.
[4] Gore, A.: “Earth in the balance”. New York: Plume Books (1992).
[5] M. Kryszkiewicz, Rough set approach to incomplete information systems. Information Science, 112(1), 1998, 39-49.
[6] J. Stefanowski, S. A. Tsouki, “Incomplete information tables and rough classification”. Computational Intelligence, 17(3), 2001, 545-566.
[7] Stefanowski J., Tsoukias. A. “On the extension of rough sets under incomplete information”. In: N Zhong, A. Tsoukias, S. Ohsugaeds. Proc of the 7th Int’l Workshop on New Directions in Rough Sets. Data Mining, and Granular-Soft Computing. Berlin: Springer-Verlag, 1999, 73-81.
[8] G. Y. Wang, “Extension of rough set under incomplete information systems”. Journal of Computer Research and Development, 39(10), 2002, 1238-1243.
[9] X. B. Yang, J. Y. Yang. “Incomplete Information System and Rough Set Theory”. Science Press, 2011, 9.
[10] X. B. Yang, “Rough set model based on variable parameter classification in incomplete information systems”. Systems Engineering-Theory & Practice, 28(5), 2008, 116-121.
[11] L. J. Wang, C. Wu, “A limited and variable precision rough model with symmetry”. Journal of Jiangnan University (Natural Science Edition), 6(6), 2007, 825-829.
[12] M. X. Li. “Limited and Variable Precision Rough Set Model”. Journal of Information & Computational Science 11:10 (2014) 3493-3501tion,” in Conf. Rec. 1995 World Academy of Science, Engineering and Technology Int. Conf. Communications, pp. 3–8.
[13] Y.H. Qian, J.Y. Liang, Y.Y. Yao, “MGRS: A multi-granulation rough set”, Information Sciences, vol.180, no.6, pp.949–970, 2010.
[14] Y.H. Qian, J.Y. Liang, C.Y. Dang, “Incomplete multi-granulation rough set”, IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 40, no.2, pp.420-431,2010.
[15] Y.H Qian, J.Y. Liang, W. Wei, “Pessimistic rough decision”, in: Second International Workshop on Rough Sets Theory, Zhoushan, China, pp. 440–449, 2010.
[16] W. X. Zhang, J. S. Mi, W. Z. Wu, “Knowledge Reductions in Inconsistent Information Systems”, Chinese Journal of Computers (in Chinese), Vol.1, No.1, pp.12-18, 2003.