Rough Set Involving with Fuzzy Weights

Hugang Han, and Akira Ikuta
Faculty of MIS, Prefectural University of Hiroshima
1-1-71 Ujina, Minami-ku, Hiroshima-city, Hiroshima 734-8558, Japan
{hhan, ikuta}@pu-hiroshima.ac.jp

Abstract – *To deal with some defects in the traditional information system (IS), such as only implementing the fully correct or certain class labeling, importance of each attribute or tuple (each example data), some new concepts such as variable precision rough model (VP-model), uncertain information system (UIS), have been suggested by some researchers. However, how to give some proper certainty and importance for each example data, and how to classify such an IS are still open problems. In this paper, we first evaluate every attribute importance using, not a singleton value but, a fuzzy number (or, suppose the original information system is along with such fuzzy numbers), due to the fact that the fuzzy numbers are easily set in comparison with the singleton ones. Then, based on the IS with the fuzzy numbers, we give a new definition of rough set. As a result, the traditional rough set is a special case in our proposal. Example shows our model is more practical than the traditional one.*

Keywords: Rough set, information system, fuzzy weight, rule extraction.

1 Introduction

The effective use of computers in various realms of human activities strongly depends on the efficiency of algorithms implemented in these computers. So far, many theoretical foundation stones for the algorithm have been set up, in which the rough set theory [1, 2] is a powerful tool to extract classification rules from a database. In general, such a database regarding the knowledge we are interested in is given in the form of information system (IS), which actually indicates an approximation space. In the traditional IS , as partly pointed out by some researchers [3, 4, 5], the approach for rule extraction implements the fully correct or certain classification rules without considering other factors such as uncertain class labeling, importance of examples. The limitations above severely reduce the applicability of the rough set approach to problems which are more probabilistic than deterministic in nature. In order to deal with the defects above and improve the reality of IS, some new concepts, such as variable precision rough model (VP-model) [3], uncertain information system (UIS) [5], have been suggested. In an UIS, considering data's noise tolerance degrees, two classification factors, which are corresponding with the positive region, and the negative region, respectively, must firstly be set up for whole system, then, the certainty and importance for each tuple need to be given. However, when we extract rules based on the UIS, there are some tough tasks to be encountered. The following are some of them. (1) It is difficult to set up some singleton values for the classification factors and the every certainty and importance. For example, you say the importance factor for a condition contribute set is 0.85, and I may say the one is 0.86. Such a little difference 0.01 may lead to a completely different classification rule. Therefore, a nature way to avoid the problem above is adopting fuzzy numbers, say "about 0.85" or "about 0.86"; (2) When the classification and each example data's factors, where no matter the values are singleton or fuzzy, are involved in the *IS*, the traditional rough set model [1] is no longer capable of giving the Upper/Lower approximation.

In this paper, in order to cover all uncertainties and noises in a database from every aspect, we first evaluate each attribute including both of condition and decision using, not a singleton value but, a fuzzy number (or, suppose the original information system is along with such fuzzy numbers), due to the fact that the fuzzy numbers are easily set in comparison with the singleton ones. Then, based on the information system with the fuzzy weights (numbers), we give a new definition of fuzzy rough set. As a result, the traditional rough set is a special case in our proposal. Consequently, an interesting difference can be shown by applying them to a same example database.

The remainder of this paper is arranged as follows. Section 2 describes some basic definitions in an *IS*, while the way of how to extract classification rules is shown with an example. In Section 3, we first point out what are problems in the traditional *IS*, then focus our attention on full explanation of new rough set model proposed in this paper. At the same time, the relation between the traditional rough set model and our model will be made clear. Also, it gives an example to show the effectiveness of the proposed rough set model. Finally, the conclusion in this paper is given in Section 4.

2 Rough Set

The database regarding the experts' know-how is generally given in the form of the information system. The definition of the traditional information system is given by Pawlak [2].

Definition 1 (*IS*) *An information systems (IS) is an ordered quadruple*

$$IS = (U, Q, V, \rho) \tag{1}$$

where U is the universe which is a non-empty finite set of objects x; Q is a finite set of attributes q; $V = U_{q \in Q} V_q$, and V_q is the domain of attribute q; ρ is a mapping function such that $\rho(x, q) \in V_q$ for every $q \in Q$ and $x \in U$. Q is composed of two parts: a set of condition attributes (C) and a decision attribute (D), i.e., $Q = C \cup D$.

ρ also is called a decision function. If we introduce function $\rho_x : Q \to V$ such that $\rho_x(q) = \rho(x, q)$ for every $q \in Q$ and $x \in U$, ρ_x is called decision rule in IS, and x is called a label of the decision rule ρ_x.

Let $IS = (U, Q, V, \rho)$ be an information system, and let $q \in Q$, $x, y \in U$. If $\rho_x(q) = \rho_y(q)$, then we say x, y are indistinguishable, in symbols xR_qy where R_q is an equivalence relation. Also, objects $x, y \in U$ are indistinguishable with respect to $P \subset Q$ in IS, in symbols xR_Py, if xR_py for every $p \in P$. In particular, if $P = Q$, x, y are indistinguishable in IS, in symbols xRy instead of xR_Qy. Therefore each information system $IS = (U, Q, V, \rho)$ defines uniquely an approximation space $A = (U, R)$, where R is an equivalence relation generated by the information system IS. The equivalence relation R partitions U into a family of disjoint subsets which are called Q-elementary sets. Likewise, R_C leads to C-elementary sets, and R_D leads to D-elementary sets.

Given an arbitrary set $X \subseteq U$, in general it may not be possible to describe X precisely in A. One may characterize X by a pair lower and upper approximations.

Definition 2 (Rought Set) *Let R be an equivalence relation on a universe U. For any set $X \subseteq U$, the lower approximation $\underline{apr}(X)$ and the upper approximation $\overline{apr}(X)$ are defined by as follows:*

$$\underline{apr}(X) = \{x \in U \mid [x]_R \subseteq X\} \tag{2}$$
$$\overline{apr}(X) = \{x \in U \mid [x]_R \cap X \neq 0\} \tag{3}$$

where

$$[x]_R = \{y \mid xRy\} \tag{4}$$

is the equivalence class containing x.

The lower approximation $\underline{apr}(X)$ is the union of elementary sets which are subsets of X, and the upper approximation $\overline{apr}(X)$ is the union of elementary sets which have a non-empty intersection with X. The set $bnd(X) = \overline{apr}(X) - \underline{apr}(X)$ is called boundary of X in A. If $bnd(X)$ is empty, then subset X is exactly definable. Note that rough set is a set (pair) of lower and upper approximation.

An accuracy measure of set X in the approximation space $A = (U, R)$ is defined as

$$\alpha(X) = \frac{|\underline{apr}(X)|}{|\overline{apr}(X)|} \tag{5}$$

where $|\cdot|$ denotes the cardinality of a set. Clearly, it is true that $0 \leq \alpha(X) \leq 1$. Besides, X is called definable in A if $\alpha(X) = 1$, and X is called undefinable in A if $\alpha(X) < 1$.

Now, let us consider the issue of rule extraction from an information system. A natural way to extract rules, or represent experts' knowledge, is to construct a set of conditional productions, each of them having the form

IF { set of conditions} **THEN** { set of decisions}

Such a form can be easily induced by taking the advantage of rough set. In an approximation space $A = (U, R)$, regarding a subset X of U, the whole universe U is partitioned into three regions:

- Positive region $pos(X) = \underline{apr}(X)$;
- Negative region $neg(X) = U - \overline{apr}(X)$;
- Boundary region $bnd(X) = \overline{apr}(X) - \underline{apr}(X)$

which lead to the following decision rules:

- Describing $pos(X) \longrightarrow$ positive decision rules;
- Describing $neg(X) \longrightarrow$ negative decision rules;
- Describing $bnd(X) \longrightarrow$ possible decision rules.

Also, the positive decision rules, possible decision rules are referred to as certain rules, possible rules, respectively. A simple illustration example is shown as follows.

[**Example 1**] Suppose that there is an information system $IS = (U, Q, V, \rho)$, which is a database about the diagnosis of influenza (Tab.1). In the information system, $U = \{x_1, x_2, \ldots, x_8\}$ in which each object (element) expresses a patient; $Q = C \cup D = \{t, s, h, l, flu\}$ in which t denotes body temperature, s sneeze, h headache, l lumbago, and flu influenza. $V_t = \{0, 1, 2\}$ in which 0 expresses "normal", 1 "high" and 2 "very high". $V_s = V_h = V_l = V_{flu} = \{0, 1\}$ in which 0 expresses "no" and 1 "yes". Also, the mapping function ρ is given in the table.

Clearly, IS yields the following elementary sets with respects to the condition attributes $C = \{t, s, h, l\}$:

$$E_1 = \{x_1, x_5\}, \quad E_2 = \{x_2\}, \quad E_3 = \{x_3\},$$

$$E_4 = \{x_4\}, \quad E_5 = \{x_6\}, \quad E_6 = \{x_7\}, \quad E_7 = \{x_8\}$$

i.e., C-elementary sets $= \{E_1, E_2, \ldots, E_7\}$.

Now, let us consider to approximate a subset

Table 1: Influenza data

U	C				D
	t	s	h	l	flu
x_1	2	0	0	0	1
x_2	1	1	0	1	1
x_3	1	0	1	1	0
x_4	0	1	1	0	1
x_5	2	0	0	0	0
x_6	0	1	1	1	0
x_7	1	1	0	0	0
x_8	1	1	1	1	1

(Q spans C and D.)

$$X = \{x_1, x_2, x_4, x_8\}$$

which is a set of patients who have influenza. Based on the concepts defined above, we have,

$$\begin{aligned} \underline{apr}(X) &= \{x_2, x_4, x_8\} \\ \overline{apr}(X) &= \{x_1, x_5, x_2, x_4, x_8\} \\ pos(X) &= \{x_2, x_4, x_8\} \\ neg(X) &= \{x_3, x_6, x_7\} \\ bnd(X) &= \{x_1, x_5\} \end{aligned}$$

therefore, the certain rules follow from $pos(X)$ below:
(1) **IF** t=1 $\wedge$ s=1 $\wedge$ h=0 $\wedge$ l=1 **THEN** flu=1;
(2) **IF** t=0 $\wedge$ s=1 $\wedge$ h=1 $\wedge$ l=0 **THEN** flu=1;
(3) **IF** t=1 $\wedge$ s=1 $\wedge$ h=1 $\wedge$ l=1 **THEN** flu=1,
where $\wedge$ denotes "and". Furthermore, the possible rules follow from $bnd(X)$ below:
(4) **IF** t=2 $\wedge$ s=0 $\wedge$ h=0 $\wedge$ l=0 **THEN** flu=1;
(5) **IF** t=2 $\wedge$ s=0 $\wedge$ h=0 $\wedge$ l=0 **THEN** flu=0,
where we can see that in rules (3) and (4), though they have the same condition in **IF** part, the decisions are different in **THEN** part. It means in such a case (condition), you are *probably* have influenza. The negative decision rules are obtained by describing $neg(X)$ as follows:
(6) **IF** t=1 $\wedge$ s=0 $\wedge$ h=1 $\wedge$ l=1 **THEN** flu=0;
(7) **IF** t=0 $\wedge$ s=1 $\wedge$ h=1 $\wedge$ l=1 **THEN** flu=0;
(8) **IF** t=1 $\wedge$ s=1 $\wedge$ h=0 $\wedge$ l=0 **THEN** flu=0.
Form (21), the approximation accuracy $\alpha(X) = 3/5 = 0.6$. □

3 Rough Set with Fuzzy Weights

As mentioned previously, we can use the theory of rough set to extract (classification) rules from an information system which leads to an approximation space. However, such a classification is based on the facts such as noise-free, importance-identical for each example (or tuple) and each attribute. Therefore, in this section we propose an information system with fuzzy weights in order to remove the weaknesses above in the traditional information system. In this paper, the fuzzy weights are expressed by fuzzy sets with triangle fuzzy memberships, which are referred to as triangle fuzzy numbers (T.F.N.s) and widely used in the fuzzy-related fields [8]. Thus, at first we introduce the triangle fuzzy numbers.

3.1 T.F.N.s [8]

3.1.1 The four operations

Fig.1 shows the membership function for a T.F.N., which can be defined by a triplet (a_1, a_2, a_3):

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & x < a_1 \\ \frac{x-a_1}{a_2-a_1}, & a_1 \le x < a_2 \\ \frac{a_3-x}{a_3-a_2}, & a_2 \le x \le a_3 \\ 0, & x > a_3 \end{cases} \tag{6}$$

Define two T.F.N.s $\tilde{A}$ and $\tilde{B}$ by the triplets as $\tilde{A} = (a_1, a_2, a_3)$ and $\tilde{B} = (b_1, b_2, b_3)$. Their addition and subtraction operations can be calculated as follows.

$$\tilde{A} + \tilde{B} = (a_1 + b_1, a_2 + b_2, a_3 + b3) \tag{7}$$
$$\tilde{A} - \tilde{B} = (a_1 - b_3, a_2 - b_2, a_3 - b1) \tag{8}$$

Also, in R^+ their multiplication and division operations can approximately calculate by

$$\tilde{A} * \tilde{B} = (a_1 * b_1, a_2 * b_2, a_3 * b3) \tag{9}$$
$$\tilde{A}/\tilde{B} = (a_1/b_3, a_2/b_2, a_3/b1) \tag{10}$$

Note that (7) and (8) hold in any cases while (9) and (10) only approximately hold in R^+.

3.1.2 Ordering of T.F.N.s

In this paper, we adopt the criterion called *removal* in order to rank or order two fuzzy sets. Let us consider an ordinary number $k \in R$, and a fuzzy set $\tilde{A}$ as illustrated in Fig.2. The left side removal of $\tilde{A}$ with respect to k, $R_l(\tilde{A}, k)$, is defined as the area bounded by k and the left side of the fuzzy set $\tilde{A}$. Similarly, the right side removal, $R_r(\tilde{A}, k)$, is defined. The removal of the fuzzy set $\tilde{A}$ with respect to k is defined as the mean of $R_l(\tilde{A}, k)$ and $R_r(\tilde{A}, k)$:

$$R(\tilde{A}, k) = \frac{R_l(\tilde{A}, k) + R_r(\tilde{A}, k)}{2} \tag{11}$$

Fig.2 shows how the left and right removals computed from the corresponding areas. The position of k can be located anywhere including $k = 0$.
If the origin 0 is conveniently moved to the left, it is possible in this case that all of fuzzy sets will have positive removal numbers. Hence, the removal numbers become positive if k is correctly chosen. The removal numbers with respect to a given k, therefore, can be taken as a measure of distances, and can thus be used for ordering fuzzy sets. The removal number $R(\tilde{A}, k)$ defined in this criterion, relative to $k = 0$ is equivalent to an "ordinary

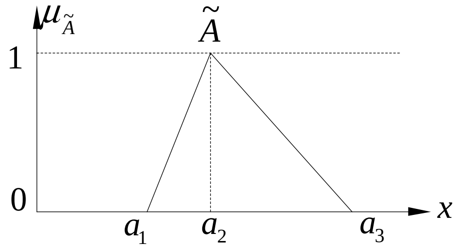

Figure 1: Triangle fuzzy number

representative" of the fuzzy set. In the case of T.F.N this ordinary representative is given by

$$\hat{A} = \frac{a_1 + 2a_2 + a_3}{4} \tag{12}$$

where $\tilde{A} = (a_1, a_2, a_3)$.

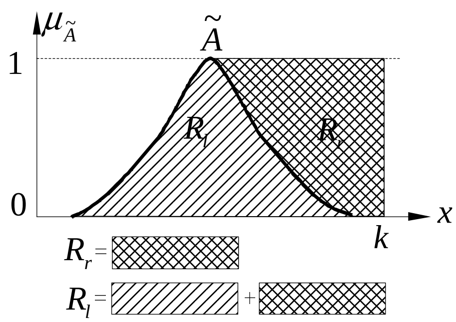

Figure 2: Left and right removals with respect to k

3.2 Fuzzy weights for attributes in C

Let us pay attention to p_4 and p_6 in Tab.1. They belonged to different elementary sets (E_4 and E_5) only because of different attribute values of l. Since each attribute in C was considered as same weight (say 1), it was a natural classification in the traditional IS based on values of condition attributes. However, in practice, each attribute does have different weight in the process of decision-making, which is corresponding to the value of attributes in D. In this case, to judge if a patient catch a cold or not (*i.e.*, $flu = 1$ or $flu = 0$), attributes (body temperature, sneeze, headache and lumbago) are adopted as some measures. Actually, the most important factor (attribute) in this judgement is body temperature, the weakest attribute is lumbago which is just as a reference to this judgement. Consequently, it is unnatural to make classification without having different weights for all attributes in C under consideration. Therefore, it is reasonable to put some fuzzy weights for the attributes like Tab.2, where $\tilde{1}$ denotes the most important attribute evaluation, and the $\tilde{0}$ denotes the weakest evaluation. Needless to say, such weights are evaluated based on the knowledge from the experts who build the database (IS). In order to take such weights into account in the process of classification, we give the following definition referred to as *absolute degree of classification.*

Definition 3 ($\tilde{D}_A$) *For two objects $x_i, x_j \in U$, their absolute degress of classfication (indistinguishableness) $\tilde{D}_A$ is defined by as follows.*

$$\tilde{D}_A(x_i, x_j) = \sum_{q \in C} |\rho(x_i, q) - \rho(x_j, q)| \tilde{w}_q \tag{13}$$

For example in Tab.2,

$$\begin{aligned} \tilde{D}_A(x_1, x_2) &= |2-1| \cdot \widetilde{1.0} + |0-1| \cdot \widetilde{0.9} \\ &\quad + |0-0| \cdot \widetilde{0.8} + |0-1| \cdot \widetilde{0.1} \\ &= \widetilde{2.0} \end{aligned}$$

It means the two tuples (x_1, x_2) are $\widetilde{2.0}$ indistinguishable, belonging to two different elementary sets. In the traditional IS, two tuples are indistinguishable with zero-tolerance if the $\tilde{D}_A$ is greater than zero. However, it is reasonable to adopt a threshold $\tilde{\beta}_A$ and consider that two tuples (x_i, x_j) belong to a same elementary set if $\tilde{D}_A(x_i, x_j) \leq \tilde{\beta}_A$. For example, setting $\tilde{\beta}_A = \widetilde{0.2}$,

$$\begin{aligned} \tilde{D}_A(x_4, x_6) &= |0-0| \cdot \tilde{1} + |1-1| \cdot \widetilde{0.9} \\ &\quad + |1-1| \cdot \widetilde{0.8} + |0-1| \cdot \widetilde{0.1} \\ &= \widetilde{0.1} \end{aligned}$$

therefore, tuples (x_4, x_6) belong to a same elementary set since $\tilde{D}_A(x_4, x_6) \leq \tilde{\beta}_A = \widetilde{0.2}$.

With the fuzzy weights $\tilde{W}_C$ for attributes in C and the concept of $\tilde{D}_A$, it is possible to consider each attribute's importance in a database, and give some flexibility in the process of classification. Clearly, once the threshold $\tilde{\beta}_A$ is given, the $\tilde{D}_A$ for every two tuples in a same elementary set should be less than or equal to $\tilde{\beta}_A$.

Table 2: Influenza data with $\tilde{W}_C$

U	Q				
	C ($\tilde{W}_C$)				D
	t ($\widetilde{1.0}$)	s ($\widetilde{0.9}$)	h ($\widetilde{0.8}$)	l ($\widetilde{0.1}$)	flu
x_1	2	0	0	0	1
x_2	1	1	0	1	1
x_3	1	0	1	1	0
x_4	0	1	1	0	1
x_5	2	0	0	0	0
x_6	0	1	1	1	0
x_7	1	1	0	0	0
x_8	1	1	1	1	1

It should be noted that for the sake of simplicity, all fuzzy numbers' operations are performed without membership functions throughout this paper. Clearly, if we take T.F.N. like Fig.1 as the membership functions of the fuzzy weights and use (12), it is easy to transfer the T.F.N.s into ordinary representatives (numbers), and order different T.F.N.s.

3.3 Fuzzy weights for attributes in D

In Tab.1, tuples x_1, x_2 have different conditions which means they have different condition attribute values, but the decisions (influenza or not) are same. In such diagnosis, which condition is easier to lead to flu? Naturally, the difference between them does occur. However, in the traditional *IS* as in Tab.1, there is indifference between the tuples which have same attribute values in D. Therefore, if there is a stronger causal relationship between the condition and decision in the case of x_1, the weight will be bigger, say $\tilde{1}$; likewise, if the causal relationship in the case of x_2 is weaker, the weight will be smaller, say $\widetilde{0.4}$. As a result, each decision attribute $q \in Q$ appears with its own fuzzy weight in the IS like Tab.3. The main purpose of putting the fuzzy weights $\widetilde{W}_D$ is to give a *evaluation* to each tuple. It means, for a same decision, there maybe are several tuples which have either same or different conditions, but considering their respective situation such as noise, confidence and so on, they do have different weights.

Table 3: Influenza data with $\tilde{W}_D$

	Q				
	C				D ($\tilde{W}_D$)
U	t	s	h	l	flu
x_1	2	0	0	0	1 ($\widetilde{0.9}$)
x_2	1	1	0	1	1 ($\widetilde{0.9}$)
x_3	1	0	1	1	0 ($\widetilde{0.9}$)
x_4	0	1	1	0	1 ($\widetilde{0.2}$)
x_5	2	0	0	0	0 ($\widetilde{0.2}$)
x_6	0	1	1	1	0 ($\widetilde{1.0}$)
x_7	1	1	0	0	0 ($\widetilde{0.5}$)
x_8	1	1	1	1	1 ($\widetilde{1.0}$)

3.4 Rough set with fuzzy weights

With the fuzzy weights $\tilde{W}_C, \tilde{W}_D$ in mind, we give a new definition.

Definition 4 **(IS_F)** *An information system with fuzzy weights (IS_F) is defined as follows:*

$$IS_F = (U, C, D, V, \rho, \tilde{W}_C, \tilde{W}_D) \tag{14}$$

where U is the universe which is a non-empty finite set of objects x; C is a finite condition set of attributes; D is a finite decision set of attributes; $V = \cup_{q \in C \cup D} V_q$, and V_q is the domain of attribute q; ρ is a mapping function such that $\rho_x(q) \in V_q$ for every $q \in C \cup D$ and $x \in U$; $\tilde{W}_C = \cup_{q \in C} \tilde{w}_q$, and $\tilde{w}_q$ is a fuzzy number defined by membership function $\mu_{\tilde{w}_q} \to [0,1]$, which is the importance evaluation of the attribute $q \in C$; $\tilde{W}_D = \cup_{x \in U} \tilde{w}_x$, and $\tilde{w}_x$ is a fuzzy number defined by membership function $\mu_{\tilde{w}_x} \to [0,1]$, which assigns each tuple an importance (weight) factor to represents how important (weighty) is for the corresponding decision.

The images of IS and IS_F are depicted in Fig.3 and Fig.4. In the IS_F (Fig.4), each object has different weight which is imaged by its size of circle, whereas each object has same size of circle in traditional *IS* (Fig.3).

At this stage, the problem left to us is how to give the

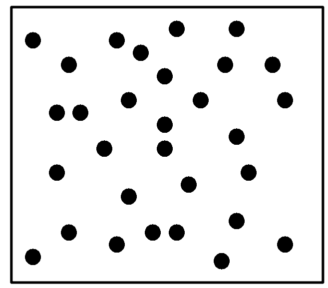

Figure 3: Image of the traditional IS

rough set in IS_F. Let E, X be a non-empty elementary set, and a non-empty subset in the approximation space, respectively. First, similarly in [3], we define a concept referred to as *relative degree of classification.*

Definition 5 **($\tilde{D}_R$)** *For an elementaty set E, its relative degree with respect to a set X is defined by as follows.*

$$\tilde{D}_R(E, X) = \frac{\sum_{x \in I} \tilde{w}_x}{\sum_{x \in E} \tilde{w}_x}; \qquad I = E \cap X \tag{15}$$

where $|E| \leq |X|$ and $|\cdot|$ denotes the cardinallity.

Fig.5 shows the image of the concept of $\tilde{D}_R$. Clearly, $0 \leq \tilde{D}_R \leq 1$, and 1 means E is completely included in X, while 0 means that E has no intersection with X.

Considering the system situation such as admissible level of misclassification, noise, and approximation precision, one can set up two thresholds $\tilde{\beta}_P, \tilde{\beta}_N$, which are called *positive threshold, negative threshold*, respectively. We say that E is included in X, if $\tilde{D}_R(E, X) \geq \tilde{\beta}_P$, and E is connected nothing with X, if $\tilde{D}_R(E, X) \leq \tilde{\beta}_N$. Based on $\tilde{D}_R$ in (15), the lower approximation, and upper approximation of a subset X with respect to thresholds $\tilde{\beta}_P$ and $\tilde{\beta}_N$, in symbols $\underline{apr}_{\tilde{\beta}_P}(X)$, $\overline{apr}_{\tilde{\beta}_N}(X)$ respectively, can be given.

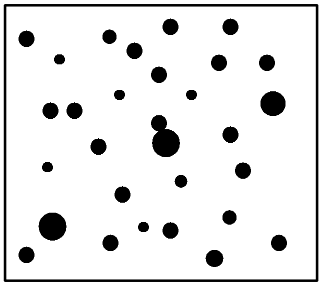

Figure 4: Image of the IS_F

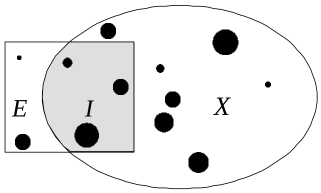

Figure 5: E, X, and their intersection I

Definition 6 (Rought Set with Fuzzy Weights) *Let R be an equivalence relation on a universe U. For any set $X \subseteq U$, the lower approximation $\underline{apr}(X)$ and the upper approximation $\overline{apr}(X)$ with two thresholds $\tilde{\beta}_P, \tilde{\beta}_N$, are defined by as follows:*

$$\underline{apr}_{\tilde{\beta}_P}(X) = POS_{\tilde{\beta}_P}(X) \tag{16}$$

$$\overline{apr}_{\tilde{\beta}_N}(X) = U - NEG_{\tilde{\beta}_N}(X) \tag{17}$$

where,

$$POS_{\tilde{\beta}_P}(X) = \bigcup\{E \in R_C^* \mid \tilde{D}_R(E,X) \geq \tilde{\beta}_P\} \tag{18}$$

$$NEG_{\tilde{\beta}_N}(X) = \bigcup\{E \in R_C^* \mid \tilde{D}_R(E,X) \leq \tilde{\beta}_N\} \tag{19}$$

and $R_C^* = \{E_1, E_2, \ldots, E_N\}$ is the C-elementary sets.

Similarly, the boundary region $BND_{\tilde{\beta}_P,\tilde{\beta}_N}(X)$ of X is composed of those elementary sets, which are neither in the positive region $POS_{\tilde{\beta}_P}(X)$, nor negative region $NEG_{\tilde{\beta}_N}(X)$ of X,

$$BND_{\tilde{\beta}_P,\tilde{\beta}_N}(X) = \bigcup\{E \in R_C^* \mid \tilde{\beta}_N < \tilde{D}_R(E,X) < \tilde{\beta}_P,\} \tag{20}$$

In this way, the accuracy measure of set X in the approximation space $A = (U, R)$ is given by

$$\tilde{\alpha}(X) = \frac{\sum_{x \in \underline{apr}_{\tilde{\beta}_P}(X)} \tilde{w}_x}{\sum_{x \in \overline{apr}_{\tilde{\beta}_N}(X)} \tilde{w}_x} \tag{21}$$

The difference between the two rough sets in IS and IS_F can be shown in Figs.6 and 7. Let us pay attention to the boundary regions. Compared the one in Fig.6, its counterpart is greatly reduced in Fig.7, where the elementary sets with arrows outward to subset X go to the negative region while the others with arrows inward to X go to the positive region.

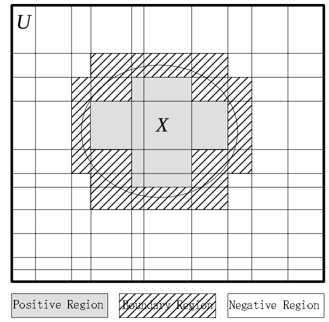

Figure 6: The three regions in IS

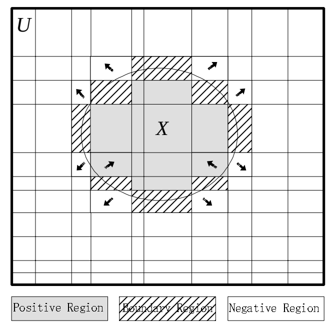

Figure 7: The three regions in IS_F

It should be noted that the IS is a special case of the IS_F, as well as rough sets. Namely,

$$IS = (U, C, D, V, \rho, W_C, W_D) \tag{22}$$

where, $W_C = \cup_{q \in C} w_q$ with $w_q = 1$, $W_D = \cup_{x \in U} w_x$ with $w_x = 1$. In other words, in the traditional IS, all tuples (objects) and all attributes have equal evaluations (weights). Besides, positive region $POS(X)$, and negative region $NEG(X)$ are the special cases of $POS_{\tilde{\beta}_P}(X)$, and $NEG_{\tilde{\beta}_N}(X)$, respectively,

$$\begin{aligned} POS(X) &= POS_1(X) \\ &= \bigcup\{E \in R_C^* \mid D_R(E,X) = 1\} \end{aligned} \tag{23}$$

$$\begin{aligned} NEG(X) &= NEG_0(X) \\ &= \bigcup\{E \in R_C^* \mid D_R(E,X) = 0\} \end{aligned} \tag{24}$$

where,

$$D_R(E,X) = \frac{\sum_{x \in I} w_x}{\sum_{x \in E} w_x}; \qquad I = E \cap X \tag{25}$$

In what follows, we apply the proposed rough set with fuzzy weights to the same date uesed in example 1 except for the fuzzy weights $\tilde{W}_C$ and $\tilde{W}_D$ (Tab.4).

[**Example 2**] Give the classification of Tab.4 with thresholds $\tilde{\beta}_c = \widetilde{0.2}, \tilde{\beta}_P = \widetilde{0.8}$, and $\tilde{\beta}_N = \widetilde{0.2}$.
Without the consideration of fuzzy weights $\tilde{W}_C$, there are 7 elementary sets as shown in example 1. However, by using the ADC with threshold $\beta_c = \widetilde{0.2}$, we can combine E_2 and E_6, E_4 and E_5 due to the facts that $\tilde{D}_R(c_2, c_7) = \tilde{D}_R(c_4, c_6) = \widetilde{0.1} \leq \tilde{\beta}_c = \widetilde{0.2}$, and finally get 5 elementary sets as follows.

$$E_1 = \{x_1, x_5\}, \quad E_2 = \{x_2, x_7\}, \quad E_3 = \{x_3\},$$

$$E_4 = \{x_4, x_6\}, \quad E_5 = \{x_8\}$$

Regarding the subset $X = \{x_1, x_2, x_4, x_8\}$ to be approximated, elementary sets in the boundary region are E_1, E_2, and E_4 from the viewpoint of the traditional rough set. However, for example, since the fact that $\tilde{D}_R(E_1, X) = \widetilde{0.9/1.1} = \widetilde{0.81} \geq \tilde{\beta}_P$, E_1 should be shifted into the positive region. In the same way, E_3 is moved out to the negative region due to the fact that $\tilde{D}_R(E_4, X) \leq \tilde{\beta}_N$. Consequently, we have

$$\underline{apr}_{\widetilde{0.8}}(X) = \{x_1, x_5, x_8\}$$

$$\overline{apr}_{\widetilde{0.2}}(X) = \{x_1, x_2, x_5, x_7, x_8\}$$

$$pos_{\widetilde{0.8}}(X) = \{x_1, x_5, x_8\} \tag{26}$$

$$neg_{\widetilde{0.2}}(X) = \{x_3, x_4, x_6\} \tag{27}$$

$$bnd_{\widetilde{0.8},\ \widetilde{0.2}}(X) = \{x_2, x_7\} \tag{28}$$

By the same manner as in example 1, we can get 3 positive rules, 2 possible rule, and 3 negative rules by describing the three regions: positive, boundary, and negative regions, respectively. Clearly, in comparison with rules (1)∼(8), the interesting thing is that although the cardinalities of positive, negative, and boundary sets are same, there contents of the three regions are quite different. For example, the contents of the boundary set are x_1 and x_5 in example 1 while they are x_2 and x_7 in example 2. Compared them, it clear that the latter reflects more the realities of the database with fuzzy weights than the former.

Table 4: Influenza data with fuzzy weights

U	C ($\tilde{W}_C$)				D ($\tilde{W}_D$)
	t ($\widetilde{1.0}$)	s ($\widetilde{0.9}$)	h ($\widetilde{0.8}$)	l ($\widetilde{0.1}$)	flu
x_1	2	0	0	0	1 ($\widetilde{0.9}$)
x_2	1	1	0	1	1 ($\widetilde{0.9}$)
x_3	1	0	1	1	0 ($\widetilde{0.9}$)
x_4	0	1	1	0	1 ($\widetilde{0.2}$)
x_5	2	0	0	0	0 ($\widetilde{0.2}$)
x_6	0	1	1	1	0 ($\widetilde{1.0}$)
x_7	1	1	0	0	0 ($\widetilde{0.5}$)
x_8	1	1	1	1	1 ($\widetilde{1.0}$)

(The header Q spans the C and D columns.)

4 Conclusion

In this paper, in order to cover the different importance of each attribute as well as some uncertainties and noises in a database we introduced an information system with fuzzy weights (IS_F). Also, a new concept of rough set is proposed in the context of IS_F. From the differences of rules gotten in IF and IS_F based on the same database in the two examples, we can see the latter is more effectively reflects the realities of a database.

References

[1] Z. Pawlak, "Rough sets," *Int. J. of Information and Computer Sciences*, vol. 11, no. 5, pp. 341-356, 1982.

[2] Z. Pawlak, "Rough classification," *Int. J. of Man-Machine Studies*, vol. 20, pp. 469-483, 1984.

[3] W. Ziarko, "Variable precision rough set model," *Journal of Computer and System Sciences*, no. 46, p. 39-59, 1993.

[4] W. Ziarko, *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Springer-Verlag, 1994

[5] J. Han, X. Hu, and N. Cercone, "Supervised learning: a generalized rough set approach," *Proceedings of the Second International Conference on Rough Sets and Current Trends in Computing (RSCTC'2000)*, pp. 284-291, Banff, Canada, October, 2000.

[6] Z. Pawlak, S. K. M. Wong, and W. Ziarko, "Rough sets: probabilistic versus deterministic approach," *Int. J. Man-Machine Studies*, vol. 29, pp. 81-95, 1988.

[7] H. Han, Y. Morioka, and K. Takano, "Rule extraction using GA-based fuzzy modeling," *Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation*, A. Grmela, and N. E. Mastorakis (eds), pp. 44-49, WSEAS, 2002.

[8] A. Kaufmann, and M. M. Gupta, *Fuzzy Mathematical Models in Engineering and Management Science*, North-Holland, 1988.