A Fuzzy Rough Set Approach for Incrementally Updating Approximations in Hybrid Information Systems

Anping Zeng[1,2], Tianrui Li[1,⋆], Chuan Luo[1], Junbo Zhang[1], and Yan Yang[1]

[1] School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China
[2] School of Computer and Information Engineering, Yibin University, Yibin 644007, China
zengap@126.com, trli@swjtu.edu.cn, luochuan@my.swjtu.edu.cn, junbozhang86@163.com, yyang@swjtu.edu.cn

Abstract. In real-applications, there may exist missing data and many kinds of data (e.g., categorical, real-valued and set-valued data) in an information system which is called as a Hybrid Information System (HIS). A new Hybrid Distance (HD) between two objects in HIS is developed based on the value difference metric. Then, a novel fuzzy rough set is constructed by using the HD distance and the Gaussian kernel. In addition, the information systems often vary with time. How to use the previous knowledge to update approximations in fuzzy rough sets is a key step for its applications on hybrid data. The fuzzy information granulation methods based on the HD distance are proposed. Furthermore, the principles of updating approximations in HIS under the variation of the attribute set are discussed. A fuzzy rough set approach for incrementally updating approximations is then presented. Some examples are employed to illustrate the proposed methods.

Keywords: Fuzzy Rough Set, Incrementally Learning, Hybrid Information Systems.

1 Introduction

Rough Set Theory(RST) is a powerful mathematical tool proposed by Pawlak [1] for processing inexact, uncertain, or vague information, and it has been widely used in several research areas including knowledge discovery, pattern recognition, artificial intelligence, and data mining [2–5].

In fact, categorical, real-valued and set-valued features usually coexist in real-world databases. A disadvantage of the Pawlak's rough set is that this model is concerned with categorical features assuming some discrete values. Some discretization algorithms can be used to divide the domain of the corresponding numerical feature into several intervals, but the discretization usually causes

⋆ Corresponding author.

D. Ciucci et al. (Eds.): RSFDGrC 2013, LNAI 8170, pp. 157–168, 2013.

information loss. Therefore, an extended model of RST, fuzzy rough set, was proposed to deal with these cases [6, 7].

When developing a fuzzy rough set model, one of important issues is generating fuzzy relations between the samples and inducing a set of fuzzy granules with the fuzzy relations. Combined with the Euclidean distance, Gaussian kernels are first introduced to acquire fuzzy relations between samples described by fuzzy or numeric attributes in order to generate fuzzy information granules in the approximation space [11]. But the Euclidean distance has difficult to deal with the categorial and set-valued data, this paper will introduces a new hybrid distance.

In real-life applications, information systems may be big data [12,13] and vary with time. In fuzzy rough sets, the generating of fuzzy relations between samples inevitably elapses a lot of time, and frequently computing the fuzzy relations will reduce efficiency of the algorithms. Incremental updating approximations is a feasible solution. In fact, in RST and its extensions, more and more serious problems are arising due to the big data and dynamic property. Some researchers have paid attention to the problem of updating approximations of RST and its extensions incrementally in dynamic information systems [14–26]. Under the variation of attribute set, Li et al. proposed some approaches for incremental updating approximations and extracting rules in RST [14–17]. However, the incremental approach for updating approximations based on fuzzy rough sets under the variation of attribute set has not been taken into account until now.

The rest of this paper is organized as follows. In Section 2, some preliminaries are introduced. In Section 3, the generating methods of fuzzy information granules in hybrid information systems are presented. In Section 4, the updating principles for lower and upper approximations are analyzed under the variation of attribute set. Some illustrative examples are conducted. In Section 5, we conclude the paper.

2 Preliminaries

The rough set theory describes a crisp subset of a universe by two definable subsets called lower and upper approximations [1]. By using the lower and upper approximations, the knowledge hidden in information systems can be discovered and expressed in the form of decision rules.

Definition 1. *Let (U, R) be a Pawlak approximation space. The universe $U \neq \emptyset$. $R \subseteq U \times U$ is an equivalence relation on U. U/R denotes the family of all equivalence classes R, and $[x]_R$ denotes an equivalence class of R containing an element $x \in U$. For any $X \subseteq U$, the lower approximation and upper approximation of X are defined respectively as follows:*

$$\begin{aligned} \underline{R}X &= \{x \in U | [x]_R \subseteq X\}; \\ \overline{R}X &= \{x \in U | [x]_R \cap X \neq \emptyset\}. \end{aligned} \tag{1}$$

The concept of fuzzy rough sets was first proposed by Dubois and Prade [6].

Definition 2. *Let R be a fuzzy equivalence relation on U and X be a fuzzy subset of U. The fuzzy lower and upper approximations of X were defined as*

$$\begin{aligned} \underline{R}X(x) &= \inf_{y\in U}\{max(1-R(x,y),X(y))\}; \\ \overline{R}X(x) &= \sup_{y\in U}\{min(R(x,y),X(y))\}. \end{aligned} \tag{2}$$

More generally, Yeung et al. proposed a model of fuzzy rough sets with a pair of T-norm and S-norm in [10].

$$\begin{aligned} \underline{R}X(x) &= \inf_{y\in U}\{S(N(R(x,y)),X(y))\}; \\ \overline{R}X(x) &= \sup_{y\in U}\{T(R(x,y),X(y))\}. \end{aligned} \tag{3}$$

In [11], based on the Gaussian kernel function, Hu et al. proposed a Gaussian kernelized fuzzy rough set model with a pair of T_{cos}-norm and S_{cos}-norm.

Definition 3. *Let R_G be a Gaussian kernelized T_{cos}-fuzzy equivalence relation on U and X be a fuzzy subset of U. The fuzzy lower and upper approximations of X are defined as*

$$\begin{aligned} \underline{R_G}X(x) &= \inf_{y\in U} S_{cos}(N(R_G(x,y)),X(y)); \\ \overline{R_G}X(x) &= \sup_{y\in U} T_{cos}(R_G(x,y),X(y)). \end{aligned} \tag{4}$$

Where $\forall x,y\in U, R_G(x,y)=k(x,y), T_{cos}(a,b)=max\{ab-\sqrt{1-a^2}\sqrt{1-b^2},0\}$ is a T-norm, and its dual $S_{cos}(a,b)=min\{a+b-ab+\sqrt{2a-a^2}\sqrt{2b-b^2},1\}$. In [11], the Gaussian kernel function $k(x,y)$ is definiton as follow.

Let U be a finite universe, and $U\neq\emptyset$. The samples are m-dimension vectors. $\forall x_i,x_k\in U,\ x_i=<x_{i1},x_{i2},...,x_{im}>,\ x_k=<x_{k1},x_{k2},...,x_{km}>$. The gaussian kernel function

$$k(x_i,x_k)=\exp\left(-\frac{||x_i-x_k||^2}{2\delta^2}\right) \tag{5}$$

can be used to compute the similarity between samples x_i and x_k. $||x_i-x_k||$ is the Euclidean distance between x_i and x_k.

A disadvantage of the Euclidean distance is that it is concerned with real values. In fact, categorical, real-valued and set-valued attributes usually coexist in real-world databases. In next section, a new hybrid distance will be introduced.

3 Gaussian Kernelized Fuzzy Rough Set in Hybrid Information Systems

Definition 4. *A Hybrid Information System (HIS) can be written as $(U,C\cup D,V,f)$, where U is the set of objects, $C=C^r\cup C^s\cup C^c$, C^r is the real-valued* *attribute set, C^s is the set-valued attribute set, C^c is the categorical attribute set, D denotes the set of decision attributes, $C^r \cap C^s = \emptyset, C^r \cap C^c = \emptyset, C^s \cap C^c = \emptyset, C \cap D = \emptyset$.*

Example 1. Table 1 is a HIS with two categorical attributes "Headache", "Muscle Pain", a real-valued attribute "Temperature", a set-valued attribute "Syndrome" (denoted as $a_1,\ a_2,\ a_3,\ a_4$, respectively), and a decision attribute d. "?" denotes the unknown value.

Table 1. A hybrid information systems

U	**Headache(a_1)**	**Muscle Pain(a_2)**	**Temperature(a_3)**	**Syndrome(a_4)**	d
x_1	Sick	Yes	40	{C, R, A}	Flu
x_2	Sick	Yes	39.5	{C, R, A}	Flu
x_3	Middle	?	39	{C}	Flu
x_4	Middle	Yes	36.8	{R}	Rhinitis
x_5	Middle	No	?	{R}	Rhinitis
x_6	No	No	36.6	{R, A}	Health
x_7	No	?	?	{A}	Health
x_8	No	Yes	38	{C, R, A}	Flu
x_9	?	Yes	37	{R}	Health

3.1 Hybrid Distance

In HIS, there are different type of attributes, to construct the distance among objects efficiently, a novel distance function should be presented. Firstly, value difference under different type of attribute should be defined.

In order to deal with the value difference under the categorical attributes, Stanfill and Waltz [27] introduced a Value Difference Metric (VDM). Based it, the normalized value difference under the categorical attributes is defined as:

Definition 5. *Let HIS=$< U, C \cup D, V, f >, \forall x, y \in U, \forall a \in C$ and a is a categorical attribute,*

$$vdm(a(x), a(y)) = \sqrt{\frac{1}{|U/D|} \sum_{d_i \in U/D} (\frac{|a(x) \cap d_i|}{|a(x)|} - \frac{|a(y) \cap d_i|}{|a(y)|})^2}. \tag{6}$$

Where $|.|$ denotes support degree, and it is clear that $vdm(a(x), a(y)) \in [0, 1]$. In [27], Wilson et al. also defined value difference under real-valued attributes.

Definition 6. *Let HIS=$< U, C \cup D, V, f >, \forall x, y \in U, \forall a \in C$ and a is a real-valued attribute,*

$$vdr(a(x), a(y)) = \frac{|a(x) - a(y)|}{4\delta_a} \tag{7}$$

where δ_a is the standard deviation under the attribute a.

In order to deal with the unknown values (denoted by "?"), Wilson et al. also defined value difference as [27]:

Definition 7. *Let HIS=$< U, C \cup D, V, f >, \forall x, y \in U, \forall a \in C$, $a(x) =?$ or $a(y) =$? and $x \neq y$,*

$$vdi(a(x), a(y)) = 1. \tag{8}$$

According to Definition 7, the value difference will be set as 1 between an unknown value and another one.

To set-valued attribute, it can be seen as a set of multiple categorical attributes. For example, to the set-valued attribute d in Table 1, the subset which has maximum cardinal number in the domain V_d is {C, R, A}. Therefore, attribute d can be divided to three categorical attributes (C, R, A, respectively). Therefore, set-value {C, R, A}={C=Yes, R=Yes,A=Yes}, {C, R}={C=Yes, R=Yes,A=?}. Because the value difference between "?" and other values is equal to 1, the value difference between {C, R, A} and {C, R} is 1/3. Therefore, the value difference of set-valued attributes is defined as follow:

Definition 8. *Let HIS=$< U, C \cup D, V, f >, \forall x, y \in U, \forall a \in C$ and a is a set-valued attribute. Let V_a be the domain of a.*

$$vds(a(x), a(y)) = 1 - \frac{|a(x) \cap a(y)|}{s} \tag{9}$$

where s is the maximum cardinal number (cardinality) in the subset of V_a.

In order to deal with the hybrid and incomplete attributes, according to Definitions 5, 6, 7 and 8, a novel Hybrid Distance (HD) can be defined as follows:

Definition 9. *Given a HIS, the Hybrid Distance (HD) is defined as:*

$$HD(x, y) = \sqrt{\sum_{a=1}^{m} d^2(a(x), a(y))} \tag{10}$$

where m is the number of attributes, and

$$d(a(x), a(y)) = \begin{cases} 1, & a(x) =? \text{ or } a(y) =? \text{ and } x \neq y \\ vdm(a(x), a(y)), & a \text{ is a categorical attribute} \\ vds(a(x), a(y)), & a \text{ is a set} - \text{valued attribute} \\ vdr(a(x), a(y)), & a \text{ is a real} - \text{valued attribute} \end{cases} \tag{11}$$

Example 2. Based on Example 1, we can compute the HD distance matrix. According to formula (10), the following results hold:

(1) Because attribute a_1 is categorical, $d(a_1(x_1), a_1(x_3)) = vdm(a_1(x_1), a_1(x_3))$ $= \sqrt{\frac{1}{3}((\frac{2}{2} - \frac{1}{3})^2 + (\frac{0}{2} - \frac{2}{3})^2 + (\frac{0}{2} - \frac{0}{3})^2)} = 0.54.$

(2) Because $a_2(x_3) =?$, $d(a_2(x_1), a_2(x_3)) = 1$.

(3) Because attribute a_3 is real-valued, $d(a_3(x_1), a_3(x_3)) = \frac{a_3(x_1)-a_3(x_3)}{4\delta_{a_3}} = (40 - 39)/(4 \times 1.28) = 0.19$.

(4) Because attribute a_4 is set-valued, $d(a_4(x_1), a_4(x_3)) = vds(a_4(x_1), a_4(x_3)) = \frac{3-1}{3} = 0.67$.

$HD(x_1, x_3) = (\sum_{a=1}^{4} d^2(a(x), a(y)))^{1/2} = \sqrt{0.54^2 + 1^2 + 0.19^2 + 0.67^2} = 1.33.$

3.2 Generating Fuzzy Relations under the Hybrid Attributes

Based on the gaussian kernel function in Formula (5), the Euclidean distance is replaced by HD distance, the new gaussian kernel function

$$k_H(x_i, x_k) = \exp\left(-\frac{||x_i - x_k||^2}{2\delta^2}\right) \tag{12}$$

$||x_i - x_k||$ is the HD distance between x_i and x_k. We have

(1) $k_H(x_i, x_k) \in [0, 1]$;

(2) $k_H(x_i, x_k) = k_H(x_k, x_i)$;

(3) $k_H(x_i, x_i) = 1$.

Using the new Gaussian kernel function, we can compute the T_{cos}-equivalence relation R_G in HIS. Furthermore, we can construct a Gaussian fuzzy rough set model.

Example 3. Base on Table 1, let $\delta^2 = 0.8$, each sample is a 4-D vector, the fuzzy relation between each two samples can be computed by Formula (12). For example, $R_G(x_1, x_3) = k_H(x_1, x_3) = exp(-\frac{HD^2(x_1,x_3)}{2\times 0.8}) = exp(-1.33^2/1.6) = 0.33$. Therefore,

$$R_G = \begin{pmatrix} 1 & 0.99 & 0.33 & 0.49 & 0.30 & 0.53 & 0.18 & 0.76 & 0.33 \\ & 1 & 0.33 & 0.53 & 0.30 & 0.57 & 0.18 & 0.79 & 0.35 \\ & & 1 & 0.26 & 0.15 & 0.29 & 0.13 & 0.33 & 0.14 \\ & & & 1 & 0.48 & 0.40 & 0.13 & 0.61 & 0.53 \\ & & & & 1 & 0.24 & 0.13 & 0.30 & 0.26 \\ & & & & & 1 & 0.22 & 0.80 & 0.26 \\ & & & & & & 1 & 0.22 & 0.08 \\ & & & & & & & 1 & 0.40 \\ & & & & & & & & 1 \end{pmatrix}.$$

3.3 Gaussian Kernelized Fuzzy Rough Set in HIS

Let HIS$=(U, C \cup D, V, f)$, $U/D = \{d_i\}, i = 1, 2, ..., |U/D|$. Here we suppose the following relationships hold: $\forall x \in d_i, d_i(x) = 1$; otherwise, $d_i(x) = 0$. Therefore, we can approximate the decision regions with the fuzzy granules induced by Gaussian function. Based on Definition 3, Hu et al. proposed the following proposition [11]:

Proposition 1. *HIS*=$(U, C \cup D, V, f)$, $\forall d_i \in U/D$,

$$\underline{R_G}d_i(x) = \inf_{y \notin d_i} \sqrt{1 - {R_G}^2(x,y)};$$
$$\overline{R_G}d_i(x) = \sup_{y \in d_i} R_G(x,y). \quad (13)$$

To simple the computing, we can generate the fuzzy lower and upper approximations by the follow proposition:

Proposition 2. *HIS*=$(U, C \cup D, V, f)$, $\forall x \in U, \forall d_i \in U/D$.

$$\underline{R_G}d_i(x) = \sqrt{1 - (\sup_{y \notin d_i} R_G(x,y))^2};$$
$$\overline{R_G}d_i(x) = \sup_{y \in d_i} R_G(x,y). \quad (14)$$

Proof. It is clear that function $y = \sqrt{1-x^2}, x \in [0,1]$ is a monotonically decreasing function. It is easy to prove that $\sqrt{1 - (\sup(x))^2} = \inf(\sqrt{1-x^2}), x \in [0,1]$. Therefore, $\underline{R_G}d_i(x) = \sqrt{1 - (\sup_{y \notin d_i} R_G(x,y))^2}$.

Example 4. Based on Examples 1 and 3, $U/D = \{d_1, d_2, d_3\}$, $d_1 = \{x_1, x_2, x_3, x_8\}$, $d_2 = \{x_4, x_5\}$, $d_3 = \{x_6, x_7, x_9\}$. According to Proposition 2,

$\underline{R_G}d_1(x_1) = \sqrt{1 - (\sup_{y \notin d_1} R_G(x_1,y))^2} = \sqrt{1 - (\sup\{0.49, 0.3, 0.53, 0.18, 0.33\})^2}$
$= \sqrt{1 - 0.53^2} = 0.85$.

Similarly, the other lower approximations can be computed as follows.
$\underline{R_G}d_1 = \{0.85/x_1, 0.82/x_2, 0.96/x_3, 0/x_4, 0/x_5, 0/x_6, 0/x_7, 0.84/x_8, 0/x_9\}$.
$\underline{R_G}d_2 = \{0/x_1, 0/x_2, 0/x_3, 0.79/x_4, 0.95/x_5, 0/x_6, 0/x_7, 0/x_8, 0/x_9\}$.
$\underline{R_G}d_3 = \{0/x_1, 0/x_2, 0/x_3, 0/x_4, 0/x_5, 0.61/x_6, 0.98/x_7, 0/x_8, 0.84/x_9\}$.

$\overline{R_G}d_1(x_1) = \sup_{y \in d_1} R_G(x_1,y) = \sup\{R_G(x_1,x_1), R_G(x_1,x_2), R_G(x_1,x_3),$
$R_G(x_1,x_8)\} = \sup\{1, 0.99, 0.33, 0.76\} = 1$.

Similarly, the other upper approximations can be computed as follows.
$\overline{R_G}d_1 = \{1/x_1, 1/x_2, 1/x_3, 0.61/x_4, 0.3/x_5, 0.8/x_6, 0.22/x_7, 1/x_8, 0.4/x_9\}$.
$\overline{R_G}d_2 = \{0.49/x_1, 0.53/x_2, 0.26/x_3, 1/x_4, 1/x_5, 0.4/x_6, 0.13/x_7, 0.61/x_8, 0.53/x_9\}$.
$\overline{R_G}d_3 = \{0.53/x_1, 0.57/x_2, 0.29/x_3, 0.53/x_4, 0.26/x_5, 1/x_6, 1/x_7, 0.8/x_8, 1/x_9\}$.

In next section, we apply the fuzzy rough set to design the incremental updating approximations under the variation of the attribute set.

4 A Fuzzy Rough Set Approach of Incrementally Updating Approximations under the Variation of the Attribute Set

We discuss the variation of approximations in HIS when the attribute set evolves over time. Given a HIS $= (U, C \cup D, V, f)$ at time t, $U \neq \emptyset$ and $C \cap D = \emptyset$. Suppose there are some attributes enter into HIS or get out of HIS at time $t+1$. The fuzzy equivalence relations will be changed. And then, the fuzzy lower and upper approximations will be changed too. Let $R_G^{(t)}$ be the fuzzy equivalence relation at time t. For each fuzzy set $X \subseteq U$, the fuzzy lower and upper approximations are denoted by $\underline{R_G^{(t)}}X$ and $\overline{R_G^{(t)}}X$ at time t, respectively. Let $P \subseteq C$ denote the attribute set at time t, R_G^P denotes the fuzzy equivalence relation under the attribute set P. Let $R_G^{(t+1)}$ be the fuzzy equivalence relation, Q_i be the immigrating attribute set and Q_e be the emigrating attribute set at time $t+1$. The fuzzy lower and upper approximations of X are denoted by $\underline{R_C^{(t+1)}}X$ and $\overline{R_G^{(t+1)}}X$, respectively. With these stipulations, we focus on the algorithms for updating approximations of the decision classes when (1) attributes enter into the HIS at time $t+1$; (2) attributes get out of the HIS at time $t+1$.

4.1 The Immigration of Attributes

Given a HIS $= (U, C \cup D, V, f)$, $\forall x_i, x_k \in U$. x_i, x_k can be seen as two m-dimension vectors, and $x_i =< x_i^{c_1}, x_i^{c_2}, ..., x_i^{c_m} >$, $x_k =< x_k^{c_1}, x_k^{c_2}, ..., x_k^{c_m} >$, $c_j \in C$, and $j = 1, ..., m, m = |C|$. $\forall P \subseteq C$, x_i, x_k can be seen as two m-dimension vectors denoted as x_i^P and x_k^P, respectively. $x_i^P =< x_i^{p_1}, x_i^{p_2}, ..., x_i^{p_l} >$, $x_k^P =< x_k^{p_1}, x_k^{p_2}, ..., x_k^{p_l} >$, $p_j \in P$, and $j = 1, ..., l, l = |P|$. According to formula (12), the following proposition holds.

Proposition 3. *$\forall P \subseteq C, \forall x_i, x_k \in U$, $x_i \neq x_k$.*

$$R_G^P(x_i, x_k) = \prod_{p_j \in P} R_G^{\{p_j\}}(x_i, x_k). \tag{15}$$

Proof. $R_G^P(x_i, x_k) = \exp(-\frac{||x_i^P - x_k^P||^2}{2\delta^2}) = \exp(-\frac{\sum_{j=1}^{|P|} d_{sa}^2(x_{ij}, x_{kj})}{2\delta^2})$
$= \prod_{p_j \in P} \exp(-\frac{d_{sa}^2(x_{ij}, x_{kj})}{2\delta^2}) = \prod_{p_j \in P} \exp(-\frac{||x_i^{\{p_j\}} - x_k^{\{p_j\}}||^2}{2\delta^2}) = \prod_{p_j \in P} R_G^{\{p_j\}}(x_i, x_k)$.

Proposition 4. *Let Q_i be an attribute set immigrating into HIS at time $t+1$. $\forall d_i \in U/D$, and $\forall x \in U$. The fuzzy approximations at time $t+1$ are as follows:*

$$\underline{R_G^{(t+1)}}d_i(x) = \sqrt{1 - (\sup_{y \notin d_i}\{R_G^{(t)}(x, y) \times \prod_{q \in Q_i} R_G^{\{q\}}(x_i, x_k)\})^2};$$
$$\overline{R_G^{(t+1)}}d_i(x) = \sup_{y \in d_i}\{R_G^{(t)}(x, y) \times \prod_{q \in Q_i} R_G^{\{q\}}(x_i, x_k)\}. \tag{16}$$

Table 2. Attribute a_5 is added into HIS

U	**Headache**(a_1)	**Muscle Pain**(a_2)	**Temperature**(a_3)	**Syndrome**(a_4)	**Cough**(a_5)	d
x_1	Sick	Yes	40	{C, R, A}	Yes	Flu
x_2	Sick	Yes	39.5	{C, R, A}	Yes	Flu
x_3	Middle	?	39	{C}	Yes	Flu
x_4	Middle	Yes	36.8	{R}	No	Rhinitis
x_5	Middle	No	?	{R}	No	Rhinitis
x_6	No	No	36.6	{R, A}	No	Health
x_7	No	?	?	{A}	No	Health
x_8	No	Yes	38	{C, R, A}	Yes	Flu
x_9	?	Yes	37	{R}	No	Health

Example 5. Based on Example 4, attribute a_5 is added into HIS (shown as Table 2). Therefore, $P = \{a_1, a_2, a_3, a_4\}$, $Q_i = \{a_5\}$.

According to Formula (12), we can compute the fuzzy relation between each two samples under the attribute set Q_i. For example, $R_G^{\{a_5\}}(x_1, x_4) = exp(-\frac{\frac{1}{3}(1+(\frac{2}{5})^2+(\frac{3}{5})^2)}{2\delta^2}) = 0.73$. Therefore,

$$R_G^{\{a_5\}} = \begin{pmatrix} 1 & 1 & 1 & 0.73 & 0.73 & 0.73 & 0.73 & 1 & 0.73 \\ & 1 & 1 & 0.73 & 0.73 & 0.73 & 0.73 & 1 & 0.73 \\ & & 1 & 0.73 & 0.73 & 0.73 & 0.73 & 1 & 0.73 \\ & & & 1 & 1 & 1 & 1 & 0.73 & 1 \\ & & & & 1 & 1 & 1 & 0.73 & 1 \\ & & & & & 1 & 1 & 0.73 & 1 \\ & & & & & & 1 & 0.73 & 1 \\ & & & & & & & 1 & 0.73 \\ & & & & & & & & 1 \end{pmatrix}.$$

Because $R_G^{(t)}$ has been generated in Example 3, $\forall x_i, x_k \in U, R_G^{(t+1)}(x_i, x_k) = R_G^{(t)}(x_i, x_k) \times R_G^{\{a_5\}}(x_i, x_k)$. Therefore,

$$R_G^{(t+1)} = \begin{pmatrix} 1 & 0.99 & 0.33 & 0.36 & 0.22 & 0.38 & 0.13 & 0.76 & 0.24 \\ & 1 & 0.33 & 0.39 & 0.22 & 0.41 & 0.13 & 0.79 & 0.25 \\ & & 1 & 0.19 & 0.11 & 0.21 & 0.09 & 0.33 & 0.10 \\ & & & 1 & 0.48 & 0.40 & 0.13 & 0.44 & 0.53 \\ & & & & 1 & 0.24 & 0.13 & 0.22 & 0.26 \\ & & & & & 1 & 0.22 & 0.58 & 0.26 \\ & & & & & & 1 & 0.16 & 0.08 \\ & & & & & & & 1 & 0.29 \\ & & & & & & & & 1 \end{pmatrix}.$$

According to Proposition 4, the approximations are as follows.

$\underline{R_G}d_1 = \{0.92/x_1, 0.91/x_2, 0.98/x_3, 0/x_4, 0/x_5, 0/x_6, 0/x_7, 0.81/x_8, 0/x_9\}$.
$\underline{R_G}d_2 = \{0/x_1, 0/x_2, 0/x_3, 0.84/x_4, 0.97/x_5, 0/x_6, 0/x_7, 0/x_8, 0/x_9\}$.
$\underline{R_G}d_3 = \{0/x_1, 0/x_2, 0/x_3, 0/x_4, 0/x_5, 0.81/x_6, 0.99/x_7, 0/x_8, 0.84/x_9\}$.
$\overline{R_G}d_1 = \{1/x_1, 1/x_2, 1/x_3, 0.44/x_4, 0.22/x_5, 0.58/x_6, 0.16/x_7, 1/x_8, 0.29/x_9\}$.
$\overline{R_G}d_2 = \{0.36/x_1, 0.39/x_2, 0.19/x_3, 1/x_4, 1/x_5, 0.4/x_6, 0.13/x_7, 0.44/x_8, 0.53/x_9\}$.
$\overline{R_G}d_3 = \{0.38/x_1, 0.41/x_2, 0.21/x_3, 0.53/x_4, 0.26/x_5, 1/x_6, 1/x_7, 0.58/x_8, 1/x_9\}$.

4.2 The Emigration of Attributes

Given two attribute sets $P, Q_e \subseteq C$, and $Q_e \subset P, Q_e \neq \emptyset$, fuzzy relation $R_G^{P-Q_e}(x_i, x_k)$ between x_i and x_k can be computed according to Proposition 3. And then, the following updating proposition of fuzzy approximations can be gotten.

Proposition 5. *Let $P \subseteq C$, and Q_e be the attributes emigrating from HIS at time $t+1$, and $Q_e \subset P$. $\forall d_i \in U/D$, and $\forall x \in U$. The fuzzy lower and upper approximations at time $t+1$ as follows:*

$$\underline{R_G^{(t+1)}} d_i(x) = \sqrt{1 - (\sup_{y \notin d_i}(R_G^{(t)}(x,y) / \prod_{q \in Q_e} R_G^{\{q\}}(x,y)))^2};$$
$$\overline{R_G^{(t+1)}} d_i(x) = \sup_{y \in d_i}(R_G^{(t)}(x,y) / \prod_{q \in Q_e} R_G^{\{q\}}(x,y)). \tag{17}$$

Table 3. The emigrating of attributes a_4, a_5

U	Headache(a_1)	Muscle Pain(a_2)	Temperature(a_3)	Syndrome(a_4)	Cough(a_5)	d
x_1	Sick	Yes	40	{C, R, A}	Yes	Flu
x_2	Sick	Yes	39.5	{C, R, A}	Yes	Flu
x_3	Middle	?	39	{C}	Yes	Flu
x_4	Middle	Yes	36.8	{R}	No	Rhinitis
x_5	Middle	No	?	{R}	No	Rhinitis
x_6	No	No	36.6	{R, A}	No	Health
x_7	No	?	?	{A}	No	Health
x_8	No	Yes	38	{C, R, A}	Yes	Flu
x_9	?	Yes	37	{R}	No	Health

Example 6. Based on Example 5, attribute set $\{a_4, a_5\}$ is deleted from HIS (shown as Table 3). Therefore $P = \{a_1, a_2, a_3, a_4, a_5\}$, $Q_e = \{a_4, a_5\}$. According to Proposition 3, we can compute the fuzzy relations under the attribute set Q_e. For example, $R_G^{Q_e}(x_1, x_4) = R_G^{\{a_4\}}(x_1, x_4) \times R_G^{\{a_5\}}(x_1, x_4) = 0.55$, $R_G^{(t+1)}(x_1, x_4) = R_G^{(t)}(x_1, x_4)/R_G^{Q_e}(x_1, x_4) = 0.36/0.55 = 0.65$. Therefore,

$$R_G^{(t+1)} = \begin{pmatrix} 1 & 0.99 & 0.43 & 0.65 & 0.40 & 0.56 & 0.24 & 0.76 & 0.43 \\ & 1 & 0.44 & 0.70 & 0.40 & 0.61 & 0.24 & 0.79 & 0.46 \\ & & 1 & 0.48 & 0.29 & 0.39 & 0.24 & 0.43 & 0.26 \\ & & & 1 & 0.48 & 0.74 & 0.24 & 0.80 & 0.53 \\ & & & & 1 & 0.44 & 0.24 & 0.40 & 0.27 \\ & & & & & 1 & 0.29 & 0.85 & 0.48 \\ & & & & & & 1 & 0.29 & 0.15 \\ & & & & & & & 1 & 0.52 \\ & & & & & & & & 1 \end{pmatrix}.$$

According to Proposition 5, the approximations are as follows:
$\underline{R_G}d_1 = \{0.76/x_1, 0.72/x_2, 0.88/x_3, 0/x_4, 0/x_5, 0/x_6, 0/x_7, 0.52/x_8, 0/x_9\}$.
$\underline{R_G}d_2 = \{0/x_1, 0/x_2, 0/x_3, 0.6/x_4, 0.9/x_5, 0/x_6, 0/x_7, 0/x_8, 0/x_9\}$.
$\underline{R_G}d_3 = \{0/x_1, 0/x_2, 0/x_3, 0/x_4, 0/x_5, 0.52/x_6, 0.96/x_7, 0/x_8, 0.84/x_9\}$.
$\overline{R_G}d_1 = \{1/x_1, 1/x_2, 1/x_3, 0.8/x_4, 0.4/x_5, 0.85/x_6, 0.29/x_7, 1/x_8, 0.52/x_9\}$.
$\overline{R_G}d_2 = \{0.65/x_1, 0.7/x_2, 0.48/x_3, 1/x_4, 1/x_5, 0.74/x_6, 0.24/x_7, 0.8/x_8, 0.53/x_9\}$.
$\overline{R_G}d_3 = \{0.56/x_1, 0.61/x_2, 0.39/x_3, 0.74/x_4, 0.44/x_5, 1/x_6, 1/x_7, 0.85/x_8, 1/x_9\}$.

5 Conclusions

In HIS, the attributes may be hybrid, and possible have unknown values. Based on this, a new HD formula was designed. Combined with the HD distance and the Gaussian kernel, a novel fuzzy rough set was constructed. In HIS, the attributes generally vary with time. The incremental updating principles of upper and lower approximations of fuzzy rough sets under the variation of the attribute set were discussed in this paper. Several examples were employed to illustrate the proposed methods. Our future research work will focus on the validation of the proposed algorithms in real data sets and the application on feature selection.

Acknowledgement. This work is supported by the National Science Foundation of China (Nos. 61175047, 61100117, U1230117), the Youth Social Science Foundation of the Chinese Education Commission (No. 11YJC630127), the Fundamental Research Funds for the Central Universities (SWJTU11ZT08, SWJTU12CX117, SWJTU12CX091), the Scientific Research Fundation of Sichuan Provincial Education Department (No. 13ZB0210), the 2013 Doctoral Innovation Funds of Southwest Jiaotong University.

References

1. Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11(5), 341–356 (1982)
2. Ananthanarayana, V.S., Narasimha, M.M., Subramanian, D.K.: Tree structure for efTcient data mining using rough sets. Pattern Recognit. Lett. 24, 851–862 (2003)
3. Skowron, A.: Extracting laws from decision tables: A rough set approach. Comput. Intell. 11, 371–388 (1995)
4. Peters, J.F., Skowron, A.: A rough set approach to knowledge discovery. International Journal of Computational Intelligence Systems 17(2), 109–112 (2002)
5. Tsumoto, S.: Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model. Information Sciences 162(2), 65–80 (2004)
6. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems 17(2-3), 191–209 (1990)
7. Yao, Y.Y.: Combination of rough and fuzzy sets based on α-level sets. In: Lin, T.Y., Cercone, N. (eds.) Rough Sets and Data Mining: Analysis for Imprecise Data, pp. 301–321. Kluwer Academic Publishers, Boston (1997)
8. Morsi, N., Mohamed, Y.: Axiomatics for Fuzzy Rough Set. Fuzzy Sets System 100(1-3), 327–342 (1998)

9. Mi, J., Zhang, W.: An Axiomatic Characterization of a Fuzzy Generalization of Rough Sets. Information Sciences 160, 235–249 (2004)
10. Yeung, D.S., Chen, D.G., Tsang, E.C.C., Lee, J.W.T., Wang, X.Z.: On the Generalization of Fuzzy Rough Sets. IEEE Trans. Fuzzy Systems 13(3), 343–361 (2005)
11. Hu, Q.H., Zhang, L., Chen, D.G., et al.: Gaussian Kernel based Fuzzy rough Sets: Model, Uncertainty Measures and Applications. International Journal of Approximate Reasoning 51, 453–471 (2010)
12. Zhang, J.B., Li, T.R., Ruan, D., Gao, Z.Z., Zhao, C.B.: A Parallel Method for Computing Rough Set Approximations. Information Sciences 194, 209–223 (2012)
13. Zhang, J.B., Li, T.R., Pan, Y.: Parallel Rough Set based Knowledge Acquisition using Map Reduce from Big Data. In: ACM SIGKDD 2012 Big Data Mining (BigMine 2012) Workshop, Beijing, China, pp. 20–27 (2012)
14. Li, T.R., Ruan, D., Wets, G., Song, J., Xu, Y.: A rough sets based characteristic relation approach for dynamic attribute generalization in data mining. Knowledge-Based Systems 20(5), 485–494 (2007)
15. Li, T.R., Xu, Y.: A generalized rough set approach to attribute generalization in data mining. Journal of Southwest Jiaotong University (English Edition) 8(1), 69–75 (2000)
16. Chan, C.C.: A rough set approach to attribute generalization in data mining. Information Sciences 107, 177–194 (1998)
17. Cheng, Y.: The incremental method for fast computing the rough fuzzy approximations. Data and Knowledge Engineering 70, 84–100 (2011)
18. Liu, D., Li, T.R., Ruan, D., Zhang, J.B.: Incremental learning optimization on knowledge discovery in dynamic business intelligent systems, Journal of Global Optimization. Journal of Global Optimization 51, 325–344 (2011)
19. Chen, H.M., Li, T.R., Hu, C.X., Ji, X.L.: An incremental updating principle for computing approximations in information systems while the object set varies with time. In: Proc. IEEE International Conference on Granular Computing, pp. 49–52. IEEE Press, Chengdu (2009)
20. Shusaku, T., Hiroshi, T.: Incremental learning of probabilistic rules from clinical database based on rough set theory. Journal of the American Medical Informations Association 4, 198–202 (1997)
21. Guan, Y., Wang, H.: Set-valued information systems. Information Sciences 176(17), 2507–2525 (2006)
22. Wang, L., Wu, Y., Wang, G.Y.: An incremental rule acquisition algorithm based on variable precision rough set model. Journal of Chongqing University of Posts and Telecommunications (Natural Science) 17(6), 709–713 (2005)
23. Zhang, J.B., Li, T.R., Ruan, D., Liu, D.: Neighborhood rough sets for dynamic data mining. International Journal of Intelligent Systems 27, 317–342 (2012)
24. Yong, L., Congfu, X., Yunhe, P.: A parallel approximate rule extracting algorithm based on the improved discernibility matrix. In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 498–503. Springer, Heidelberg (2004)
25. Guo, S., Wang, Z.Y., Wu, Z.C., Yan, H.P.: A novel dynamic incremental rules extraction algorithm based on rough set theory. In: Proc. Fourth International Conference on Machine Learning and Cybernetics, pp. 1902–1907. IEEE Press, Guang Dong (2005)
26. Zheng, Z., Wang, G.Y.: A rough set and rule tree based incremental knowledge acquisition algorithm. Fundamenta Informaticae 59(2-3), 299–313 (2004)
27. Randall Wilson, D., Martinez Tony, R.: Improved Heterogeneous Distance Functions. Journal of Artificial Intelligence Research 6, 1–34 (1997)