Journal of Uncertain Systems
Vol.3, No.4, pp.252-256, 2009
Online at: www.jus.org.uk

Rough Set Model based on Uncertain Measure

Dazeng Tian*, Lin Wang, Jing Wu, Minghu Ha
College of Mathematics and Computer Sciences, Hebei University, Baoding 071002, China

Received 6 January 2009; Revised 12 March 2009

Abstract

Probabilistic rough set model based on probability measure is a new rough set model to deal with uncertain information systems. Uncertain measure is a generalization of probability measure. Based on the fundamental knowledge of rough set model and uncertain measure, a rough set model based on uncertain measure is established. Furthermore, by comparative study of the lower approximation and upper approximation, it is true that the rough set model based on uncertain measure is an extension of the probabilistic rough set model.

Keywords: rough set model, lower approximation, upper approximation, uncertain measure

1 Introduction

In 1982, Pawlak [1] introduced the theory of rough sets as an extension of set theory for the study of intelligent systems characterized by insufficient and incomplete information. In 1991, Pawlak [2] summarized the results of the theory and practices in the former period. In 1995, ACM communication classified it as emerging research topic in computer science. Because of its wide applications in machine learning, knowledge discovery, data mining, decision support and analysis [3,4], the research about rough set theory is taken seriously [5].

The study of rough set model is an important part of the research on rough set theory [6]. Pawlak rough set model is based on certainty knowledge database, therefore, it is helpless to rules extraction of incoordination decision tables. Probabilistic rough set model [7, 8] can offer a new model to study the uncertainty information systems. Zhang [9] and Gong [10] further discussed the probabilistic rough set model, and gave some developments and applications. But the new model is based on probability measure. In 2007, Liu [11] founded an uncertain measure, which is based on normality, monotonicity, self-duality, and countable subadditivity axioms. The measure is an important generalization of probability measure. In this paper, we will establish the rough set model based on uncertain measure, and discuss some properties of the model.

2 Preliminaries

In this section, the basic knowledge containing uncertain measure, conditional uncertain measure, and classical rough set model are introduced.

Definition 1 ([11]) Let Γ be a nonempty set and $\mathcal{L}$ be a $\sigma-$algebra on Γ. Each element $\Lambda \in \mathcal{L}$ is called an event. An extended real valued set function $\mathcal{M}$ defined on $\mathcal{L}$ is called an uncertain measure, if it satisfies:

(1) (Normality) $\mathcal{M}(\Gamma)=1$.

(2) (Monotonicity) $\mathcal{M}(\Lambda_1) \le \mathcal{M}(\Lambda_2)$ whenever $\Lambda_1, \Lambda_2 \in \mathcal{L}$, and $\Lambda_1 \subseteq \Lambda_2$.

(3) (Self-Duality) $\mathcal{M}(\Lambda)+\mathcal{M}(\Lambda^c)=1$ for every $\Lambda \in \mathcal{L}$.

(4) (Countable Subadditivity) There is

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \le \sum_{i=1}^{\infty} \mathcal{M}(\Lambda_i)$$

* Corresponding author. Email: tdz19651204@mail.hbu.edu.cn (D. Tian).

for every countable sequence (Λ_i) of sets in $\mathcal{L}$.

The triplet $(\Gamma,\mathcal{L},\mathcal{M})$ is called an uncertainty space.

Definition 2 ([11]) Let $(\Gamma,\mathcal{L},\mathcal{M})$ be an uncertainty space, $A,B\in\mathcal{L}$. Then the conditional uncertain measure of A given B is defined by

$$\mathcal{M}\{A\mid B\}=\begin{cases}\dfrac{\mathcal{M}\{A\cap B\}}{\mathcal{M}\{B\}}, & \text{if } \dfrac{\mathcal{M}\{A\cap B\}}{\mathcal{M}\{B\}}<0.5\\ 1-\dfrac{\mathcal{M}\{A^c\cap B\}}{\mathcal{M}\{B\}} & \text{if } \dfrac{\mathcal{M}\{A^c\cap B\}}{\mathcal{M}\{B\}}<0.5\\ 0.5 & \text{otherwise}\end{cases}$$

provided that $\mathcal{M}(\mathrm{B})>0$.

Definition 3 ([8]) Let U be a finite and nonempty set, R be an equivalent relation on U, U/R be all equivalence classes of R, and $[x]_R$ be R equivalence class containing x ($x\in U$). For each subset $\mathrm{X}\subseteq U$, we define two subsets:

$\underline{R}X=\cup\{Y\in U/R\mid Y\subseteq X\}=\{x\in U\mid [x]_R\subseteq X\}$ $\overline{R}X=\cup\{Y\in U/R\mid Y\cap X\neq\varnothing\}=\{x\in U\mid [x]_R\cap X=\varnothing\}$.

The two equations are called R lower approximation and R upper approximation of X, respectively.

R boundary region, positive region and negative region of X are defined respectively as follows:

$$bn_R(X)=\overline{R}X-\underline{R}X\ ,\ pos_R(X)=\underline{R}X\ ,\ neg_R(X)=U-\overline{R}X\ .$$

The approximation accuracy $\rho_R(X)$ of X with respect to R is defined by:

$$\rho_R(X)=\frac{|\underline{R}X|}{|\overline{R}X|},$$

where $\mathrm{X}\neq\varnothing$, and $|X|$ denotes the cardinal number of X.

Example 1: Let $U=\{x_1,x_2,x_3,x_4\}$ be the universe, R be an equivalence relation on U, (U,R) be the approximate space, of which equivalence class is $U/R=\{\{x_1,x_2\},\{x_3\},\{x_4\}\}$ and $X=\{x_1,x_4\}$ be one of the subsets of U. According to the above definitions, we can obtain the lower and upper approximate sets of X on (U,R) as followings: $\underline{R}X=\{x_4\}$, $\overline{R}X=\{x_1,x_2,x_4\}$, $\rho_R(X)=1/3$.

3 Rough Set Model based on Uncertain Measure

Definition 4: Let U be a finite and nonempty set, R be an equivalence relation on U, and the equivalence classes of R be $U/R=\{X_1,X_2,...,X_n\}$. The equivalence class containing x is denoted by $[x]$. Let $\mathcal{M}$ be an uncertain measure over $\sigma-$ algebra assembled by subsets of U. Then the triplet $A_{\mathcal{M}}=(U,R,\mathcal{M})$ is called uncertainty approximation space. Any subset of U is called notion, which means a random event. $\mathcal{M}\{X\mid Y\}$ is the conditional uncertain measure of X given Y.

Definition 5: Let $0\le\beta<\alpha\le 1$ and $\forall X\subseteq U$. Then its uncertain lower and upper approximations with respect to uncertainty approximation space $A_{\mathcal{M}}=(U,R,\mathcal{M})$ according to parameters α,β are defined as follows:

$$\underline{\mathcal{M}}_\alpha(\mathrm{X})=\{x\in U\mid \mathcal{M}(X\mid[x])\ge\alpha\};$$

$$\overline{\mathcal{M}}_\beta(X)=\{x\in U\mid \mathcal{M}(X\mid[x])>\beta\}.$$

The uncertain positive region, negative region and boundary region of X with respect to $A_{\mathcal{M}}$ according to parameters α,β are defined by:

$$pos(X,\alpha,\beta)=\underline{\mathcal{M}}_\alpha(X)=\{x\in U\mid\mathcal{M}(X\mid[x])\ge\alpha\},\ bn(X,\alpha,\beta)=\{x\in U\mid\beta<\mathcal{M}(X\mid[x])<\alpha\},$$

$$neg(X,\alpha,\beta)=U\setminus\overline{\mathcal{M}}_\beta=\{x\in U\mid\mathcal{M}(X\mid[x])\le\beta\}$$

Obviously, the above definitions are equivalent to the following equations:

$$\underline{\mathcal{M}}_{\alpha}(X)=pos(X,\alpha,\beta)=\cup\{[x]\mid\mathcal{M}(X\mid[x])\geq\alpha\}\ \overline{\mathcal{M}}_{\beta}(X)=\cup\{[x]\mid\mathcal{M}(X\mid[x])>\beta\},$$
$$bn(X,\alpha,\beta)=\cup\{[x]\mid\beta<\mathcal{M}(X\mid[x])<\alpha\},\ neg(X,\alpha,\beta)=\cup\{[x]\mid\mathcal{M}(X\mid[x])\leq\beta\}.$$

The uncertain positive region, negative region and boundary region of X with respect to A_M according to parameters α,β form a division of U , that is

$$\overline{\mathcal{M}}_{\beta}(X)=pos(X,\alpha,\beta)\cup bn(X,\alpha,\beta)\ \text{or}\ bn(X,\alpha,\beta)=\overline{\mathcal{M}}_{\beta}(X)\setminus pos(X,\alpha,\beta).$$

When $\underline{\mathcal{M}}_{\alpha}(X)=\overline{\mathcal{M}}_{\beta}(X)$, or $bn(X,\alpha,\beta)=\varnothing$, X is called uncertain defined set with respect to $A_{\mathcal{M}}$ according to parameters α,β, otherwise X is called uncertain rough set with respect to $A_{\mathcal{M}}$ according to parameters α,β.

Definition 6: If $0\leq\beta<\alpha\leq 1$, then the approximation accuracy $\eta(X,\alpha,\beta)$ of X with respect to uncertainty approximation space $A_{\mathcal{M}}$ according to α,β is defined by

$$\eta(X,\alpha,\beta)=\frac{|\underline{\mathcal{M}}_{\alpha}(X)|}{|\overline{\mathcal{M}}_{\beta}(X)|},$$

where the numbers of the elements contained in X are expressed by $|X|$.

Then we can see, X is definable about uncertainty approximation space A_M according to α,β if and only if $\eta(X,\alpha,\beta)=1$ or $\rho(X,\alpha,\beta)=0$.

Example 2: Let $U=\{x_1,x_2,x_3,x_4\}$ be the universe and R be an equivalence relation on U. Let its equivalence class be $U/R=\{\{x_1,x_2\},\{x_3\},\{x_4\}\}$, $X=\{x_1,x_4\}$ be one of the subsets of U, and let $\mathcal{M}\{x\}$ be the uncertain measure defined on a $\sigma-$algebra over U. We define

$$\begin{array}{lll}\mathcal{M}\{x_1\}=0.2,\mathcal{M}\{x_2\}=0.3,\mathcal{M}\{x_3\}=0.4, & \mathcal{M}\{x_1,x_2\}=0.4,\mathcal{M}\{x_1,x_3\}=0.5, & \mathcal{M}\{x_1,x_2,x_3\}=0.7,\mathcal{M}\{x_1,x_2,x_4\}=0.6,\\ \mathcal{M}\{x_4\}=0.3,\mathcal{M}\{\varnothing\}=0,\mathcal{M}\{U\}=1 & \mathcal{M}\{x_1,x_4\}=0.3,\mathcal{M}\{x_2,x_3\}=0.7, & \mathcal{M}\{x_1,x_3,x_4\}=0.7,\mathcal{M}\{x_4,x_2,x_3\}=0.8.\\ & \mathcal{M}\{x_2,x_4\}=0.5,\mathcal{M}\{x_3,x_4\}=0.6 & \end{array}$$

According to the above definitions, we can obtain the uncertain lower and upper approximate sets of X with respect to R according to the parameters α,β:

$$\underline{\mathcal{M}}_1\{X\}=\{x_4\}\ \text{and}\ \overline{\mathcal{M}}_0=\{x_1,x_2,x_4\}\ \eta(X,\alpha,\beta)=\frac{1}{3}\ \text{when}\ \alpha=1,\beta=0;$$

$$\underline{\mathcal{M}}_{\alpha}\{X\}=\{x_4\}\ \text{and}\ \overline{\mathcal{M}}_{\beta}=\{x_1,x_2,x_4\}\ \eta(X,\alpha,\beta)=\frac{1}{3}\ \text{when}\ 0\leq\beta<0.5\leq\alpha<1;$$

$$\underline{\mathcal{M}}_{\alpha}\{X\}=\{x_1,x_2,x_4\}\ \text{and}\ \overline{\mathcal{M}}_{\beta}=\{x_1,x_2,x_4\}\ \eta(X,\alpha,\beta)=1\ \text{when}\ 0\leq\beta<\alpha\leq 0.5;$$

$$\underline{\mathcal{M}}_{\alpha}\{X\}=\{x_4\}\ \text{and}\ \overline{\mathcal{M}}_{\beta}=\{x_4\}\ \eta(X,\alpha,\beta)=1\ \text{when}\ 0.5\leq\beta<\alpha<1.$$

From the Examples 1 and 2, we can obtain $\underline{R}(X)=\{x\in U\mid[x]\subseteq X\}\subseteq\underline{\mathcal{M}}_{\alpha}(X)$

$$\subseteq\overline{\mathcal{M}}_{\beta}(X)\subseteq\overline{R}(X)=\{x\in U\mid[x]\cap X\neq\varnothing\},$$

and the conclusion that the new model can improve the approximation accuracy.

Theorem 1: Let $0\leq\beta<\alpha\leq 1$, $X,Y\subseteq U$. Then the uncertain approximation operators satisfies the flowing dual properties:

(1) $\underline{\mathcal{M}}_{\alpha}(\varnothing)=\overline{\mathcal{M}}_{\beta}(\varnothing)=\varnothing$, $\underline{\mathcal{M}}_{\alpha}(U)=\overline{\mathcal{M}}_{\beta}(U)=U$;

(2) $\underline{\mathcal{M}}_{\alpha}(X)\subseteq\overline{\mathcal{M}}_{\beta}(X)$;

(3) $\underline{\mathcal{M}}_{\alpha}(X)=\sim\overline{\mathcal{M}}_{1-\alpha}(\sim X),\overline{\mathcal{M}}_{\beta}(X)=\sim\underline{\mathcal{M}}_{1-\beta}(\sim X)$;

(4) When $X\subseteq Y$, $\underline{\mathcal{M}}_{\alpha}(X)\subseteq\underline{\mathcal{M}}_{\alpha}(Y),\overline{\mathcal{M}}_{\beta}(X)\subseteq\overline{\mathcal{M}}_{\beta}(Y)$;

(5) $\overline{\mathcal{M}}_{\beta}(X\cup Y)\supseteq\overline{\mathcal{M}}_{\beta}(X)\cup\overline{\mathcal{M}}_{\beta}(Y)$, $\underline{\mathcal{M}}_{\alpha}(X\cap Y)\subseteq\underline{\mathcal{M}}_{\alpha}(X)\cap\underline{\mathcal{M}}_{\alpha}(Y)$;

(6) $\overline{\mathcal{M}}_{\beta}(X\cap Y)\subseteq\overline{\mathcal{M}}_{\beta}(X)\cap\overline{\mathcal{M}}_{\beta}(Y)$, $\underline{\mathcal{M}}_{\alpha}(X\cup Y)\supseteq\underline{\mathcal{M}}_{\alpha}(X)\cup\underline{\mathcal{M}}_{\alpha}(Y)$;

(7) If $\alpha_1\leq\alpha_2,\beta_1\leq\beta_2$, then $\underline{\mathcal{M}}_{\alpha_2}(X)\subseteq\underline{\mathcal{M}}_{\alpha_1}(X),\overline{\mathcal{M}}_{\beta_2}(X)\subseteq\overline{\mathcal{M}}_{\beta_1}(X)$

The proof can be obtained directly from the definition.

Remark: When $\alpha = 1, \beta = 0$, if $\mathcal{M}\{X \mid [x]\} = \dfrac{|X \cap [x]|}{|[x]|}$, then

$$\underline{\mathcal{M}}_\alpha(X) = \underline{\mathcal{M}}_1(X) = \{x \in U \mid [x] \subseteq X\} \text{ and } \overline{\mathcal{M}}_\beta(X) = \overline{\mathcal{M}}_0(X) = \{x \in U \mid [x] \cap X \neq \emptyset\}.$$

In this situation, the uncertain lower and upper approximations respectively become lower and upper approximations of Pawlak rough set model, respectively. Therefore, the uncertain rough set model is an extension form of Pawlak rough set model.

From (2) and (7) in Theorem 1, we can obtain:

$$\underline{R}(X) = \{x \in U \mid [x] \subseteq X\} \subseteq \underline{\mathcal{M}}_\alpha(X) \subseteq \overline{\mathcal{M}}_\beta(X) \subseteq \overline{R}(X) = \{x \in U \mid [x] \cap X \neq \emptyset\}$$

Therefore, the boundary region of the uncertainty rough set model is narrower than Pawlak's, but the positive region and the negative region are all wider than Pawlak's. The reason is that the uncertain rough set model allows to include some part of Pawlak rough set boundary region whose conditional uncertain measure is less than one and greater than zero into its positive region. From this, we know that the application of the uncertainty rough set model is more extensive than Pawlak's. It can be seen from (7), $\underline{\mathcal{M}}_\alpha(X)$ increases with the parameter α decreasing, while $\overline{\mathcal{M}}_\beta(X)$ decreases with the parameter β increasing. However, the following equations

$$\lim_{\alpha \downarrow \gamma} \underline{\mathcal{M}}_\alpha(X) = \underline{\mathcal{M}}_\gamma(X), \ \lim_{\beta \uparrow \gamma} \overline{\mathcal{M}}_\beta(X) = \overline{\mathcal{M}}_\gamma(X)$$

are generally not true.

Theorem 2: Let $0 < \gamma < 1$. Then $\forall X \subseteq U$, we have

(1) $\lim_{\alpha \downarrow \gamma} \underline{\mathcal{M}}_\alpha(X) = \bigcup_{\alpha > \gamma} \underline{\mathcal{M}}_\alpha(X) = \overline{\mathcal{M}}_\gamma(X)$;

(2) $\lim_{\beta \uparrow \gamma} \overline{\mathcal{M}}_\beta(X) = \bigcap_{\beta < \gamma} \overline{\mathcal{M}}_\beta(X) = \underline{\mathcal{M}}_\gamma(X)$.

Proof: (1) When $\alpha > \gamma$, then

$$\underline{\mathcal{M}}_\alpha(X) = \{x \in U \mid \mathcal{M}\{X \mid [x]\} \geq \alpha\} \subseteq \{x \in U \mid \mathcal{M}\{X \mid [x]\} > \gamma\} = \overline{\mathcal{M}}_\gamma(X)$$

and we know $\underline{\mathcal{M}}_\alpha(X)$ increases with the parameter α decreasing, so $\lim_{\alpha \downarrow \gamma} \underline{\mathcal{M}}_\alpha(X) = \bigcup_{\alpha > \gamma} \underline{\mathcal{M}}_\alpha(X) \subseteq \overline{\mathcal{M}}_\gamma(X)$.

If $x_0 \in \overline{\mathcal{M}}_\gamma(X) \setminus \bigcup_{\alpha > \beta} \underline{\mathcal{M}}_\alpha(X)$ exists, then $\mathcal{M}\{X \mid [x_0]\} > \gamma$. However for any $\alpha > \gamma$, we get $x_0 \notin \underline{\mathcal{M}}_\alpha(X)$, that is $\forall \alpha > \gamma$, $\mathcal{M}\{X \mid [x_0]\} < \alpha$, then $\mathcal{M}\{X \mid [x_0]\} \leq \gamma$ and $\mathcal{M}\{X \mid [x_0]\} > \gamma$ are contradiction! This means $\bigcup_{\alpha > \beta} \underline{\mathcal{M}}(X) = \overline{\mathcal{M}}_\gamma(X)$. Therefore, the conclusion is proved.

(2) For any $\beta < \gamma$,we have $\overline{\mathcal{M}}_\beta(X) = \{x \in U \mid Cr\{X \mid [x]\} > \beta\} \supseteq \{x \in U \mid M\{X \mid [x]\} \geq \gamma\} = \underline{M}_\gamma(X)$ and $\overline{\mathcal{M}}_\beta(X)$ decreases with the parameter β increasing, we get $\lim_{\beta \uparrow \gamma} \overline{\mathcal{M}}_\beta(X) = \bigcap_{\beta < \gamma} \overline{\mathcal{M}}_\beta(X) \supseteq \underline{\mathcal{M}}_\gamma(X)$. If $y \in \bigcap_{\beta < \gamma} \overline{\mathcal{M}}_\beta(X) \setminus \underline{\mathcal{M}}_\gamma(X)$ exists, then for $\forall \beta < \gamma, y \in \overline{\mathcal{M}}_\beta(X)$, but $y \notin \underline{\mathcal{M}}_\gamma(X)$, thus $\mathcal{M}\{X \mid [y]\} > \beta$, $\forall \beta < \gamma$, while $\mathcal{M}\{X \mid [y]\} < \gamma$. According to the above formula we know $\mathcal{M}\{X \mid [y]\} \geq \gamma$, which is contradictory. Therefore, conclusion (2) is proved.

Theorem 3: Let $0 < \gamma < 1$. Then $\forall X \subseteq U$,

(1) $\lim_{\alpha \uparrow \gamma} \underline{\mathcal{M}}_\alpha(X) = \bigcap_{\alpha < \gamma} \underline{\mathcal{M}}_\gamma(X) = \underline{\mathcal{M}}_\gamma(X)$;

(2) $\lim_{\beta \downarrow \gamma} \overline{\mathcal{M}}_\beta(X) = \bigcup_{\alpha > \gamma} \overline{\mathcal{M}}_\beta(X) = \overline{\mathcal{M}}_\gamma(X)$.

Theorem 4: Let $0 < \gamma < 1$. Then

$$\begin{aligned} \lim_{\beta \uparrow \gamma, \alpha \downarrow \gamma} bn(X, \alpha, \beta) &= \bigcap_{\beta < \gamma < \alpha} \left(\overline{\mathcal{M}}_\beta(X) \setminus \underline{\mathcal{M}}_\alpha(X) \right) \\ &= \underline{\mathcal{M}}_\gamma(X) \setminus \overline{\mathcal{M}}_\gamma(X) = \left\{ x \in U \mid \mathcal{M}\left(X \mid [x] \right) = \gamma \right\} \end{aligned}.$$

4 Conclusions

Probabilistic rough set model can solve the uncertainty of the information systems, but it is based on probability measure, which satisfies countable additivity axiom. However, in our life, there exist many measures that don't satisfy countable additivity axiom such as capacity measure, fuzzy measure and so on. Uncertain measure founded by Liu in 2007 is an important non-additive measure. In this paper, rough set model based on uncertain measure is established. This model is an extension of probabilistic rough set model. But how to define the different values of two parameters in the model needs to further discuss in the future.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 60773062), the Natural Science Foundation of Hebei Province of China (No. F2008000633), the Key Scientific Research Project of Education Department of Hebei Province of China(No.2005001D) and the Key Scientific and Technical Research Project of the Ministry of Education of China (No. 206012).

References

[1] Pawlak, Z., Rough sets, *International Journal of Computer and Information Sciences*, vol.11, pp.341-356, 1982.

[2] Pawlak, Z., *Rough Sets-Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Bosten, 1991.

[3] Jacksona, A.G., Z. Pawlak, and S.R. Leclaira, Rough sets applied to the discovery of materials knowledge, *Journal of Alloys and Compounds*, vol.279, pp.14-21, 1998.

[4] Pawlak, Z., Rough sets and intelligent data analysis, *Information Sciences*, vol.147, pp.1-12, 2002.

[5] Zhang, W.X., and W.Z. Wu, A review of introduction and research about rough set theory, *Fuzzy Systems and Mathematics*, vol.14, no.4, pp.1-12, 2000.

[6] Yao, Y.Y., S.K.M. Wong, and T.Y. Lin, *A Review of Rough Set Models*, Kluwer Academic Publishers, Bosten, 1991.

[7] Wang, J.Y., and L.M. Xu, Probabilistic rough set models, *Computer Science*, vol.29, no.8, pp.76-78, 2002.

[8] Zhang, W.X., and W.Z. Wu, *Rough Set Theory and Method*, Science Press, Beijing, 2001.

[9] Fu, C.L., X.H. Du, and Q.Z. Yao, An incremental rules learning algorithm based on probabilistic rough set model, *Computer Science*, vol.35, no.5, pp.143-146, 2008.

[10] Gong, Z.T., and Shi, Z.H., Probabilistic rough set model based on covering and it's Bayes decision, *Fuzzy Systems and Mathematics*, vol.22, no.4, pp.142-148, 2008.

[11] Liu, B., *Uncertainty Theory*, 2nd ed, Springer-Verlag, Berlin, 2007.