关于模糊集的粗糙度*)

闫德勤　张丽平

（辽宁师范大学计算机系　大连 116029）

摘　要　模糊集粗糙度的研究对于粗糙集和模糊集的理论和应用都有着重要作用。最近，Huynh 等改进了由 Banerjee 等给出的模糊集的粗糙度，提出了一种新的关于模糊集粗糙度的度量方法。本文对 Huynh 和 Banerjee 提出的粗糙度量进行了分析，指出了其中的不足，提出了一种新的改进方法。对 Huynh 度量方法存在缺陷的分析以及对新提出的改进方法性质的研究证明了新改进方法的合理性。

关键词　模糊集，粗糙集，粗糙度

Roughness Measure for Fuzzy Sets

YAN De-Qin　ZHANG Li-Ping

(Department of Computer Science, Liaoning Normal University, Dalian 116029)

Abstract　In this paper, the roughness measure for fuzzy sets is approached. By analyzing the defect of the measures for fuzzy sets proposed by Huynh and Banerjee, a new modified measure is proposed. To demonstrate the measure, an example is presented.

Keywords　Fuzzy sets, Rough sets, Roughness

1　引言

自模糊集(Fuzzy Sets)概念提出以来[6]，由于其在知识表达、系统分析与控制等诸多领域起着重要的作用，其相关理论和应用研究一直被人们所关注。同时，很多理论方法与模糊集的结合也产生着丰富的有重要理论和应用价值的研究结果。粗糙集(Rough Sets)就是其中的一个方面[1,2,7]。

Banerjee 等在文[3]中结合粗糙集粗糙度的思想方法，利用粗糙模糊集[4](Rough Fuzzy Sets)的概念，提出了计算模糊集的粗糙度的方法。以此为基础，一些学者进行了相关的研究。最近，Huynh 等在文[5]中指出 Banerjee 关于模糊集粗糙度的度量方法存在缺陷，并结合模糊截集，利用粗糙集粗糙度的方法，提出了一种新的模糊集粗糙度的度量方法。本文对 Huynh 和 Banerjee 提出的粗糙度进行了分析，指出了其中的不足，提出了一种对 Huynh 方法改进的度量方法，同时给出了应用例子。

2　相关概念

2.1　粗糙集

设 $U=\{x_1,x_2,\cdots,x_n\}$ 是一有限集，称为论域，R 是 U 上的一个等价关系，$\langle U,R\rangle$ 表示有限近似空间，$U/R=\{X_1,X_2,\cdots,X_m\}$ 表示在 U 上导出的所有等价类；$[x]_R$ 表示包含元素 x 的 R 的等价类。对任集合 $X\subseteq U$

$R_-(X)=\{x\in U\mid[x]_R\subseteq X\}$

$R^-(X)=\{x\in U\mid[x]_R\cap X\neq\varnothing\}$（$\varnothing$ 为空集）

分别称 $R_-(x)$ 与 $R^-(X)$ 为 X 的 R 下近似和 X 的 R 上近似，并构成关于 X 的粗糙集。

在文[2]中，Pawlak 分别给出了 X 在粗糙集表示下的精度和粗糙度：

$$\alpha_R(X)=\frac{|R_-(X)|}{|R^-(X)|}$$

$$\rho_R(X)=1-\alpha_R(X)=1-\frac{|R_-(X)|}{|R^-(X)|}$$

其中，$\alpha_R(X)$ 为精度，$\rho_R(X)$ 为粗糙度，$|\cdot|$ 表示集合的基数。

2.2　模糊集和粗糙模糊集

关于模糊集有如下定义：

定义 1　设在论域 U 上给定了一个映射 $F:U\to[0,1]$，$x\mapsto\mu_F(x)$，则称 F 为 U 上的模糊(Fuzzy)集，$\mu_F(x)$ 称为 F 的隶属度函数(或称为 x 对 F 的隶属度)。

对于给定的 $\alpha\in(0,1]$，模糊集 F 的 α 截集表示为：$F_\alpha=\{x\in U\mid\mu_F(x)\geqslant\alpha\}$。

在文[4]中，Dubois 等给出了粗糙模糊集的概念：

设 $\langle U,R\rangle$ 为有限近似空间，F 为 U 上的模糊集，在等价类 R 下 F 的上下近似 $R^-(F)$ 和 $R_-(F)$ 在等价类 $U/R=\{X_1,X_2,\cdots,X_m\}$ 上的隶属度分别为：

$\mu_{R^-(F)}(X_i)=\max_{x\in X_i}\{\mu_F(x)\}$

$\mu_{R_-(F)}(X_i)=\min_{x\in X_i}\{\mu_F(x)\}$

$i=1,2,\cdots,m$

$(R^-(F),R_-(F))$ 称为粗糙模糊集。

一般，为简单起见 $(R^-(F),R_-(F))$ 写为 (F^*,F_*)。

3　模糊集的粗糙度

在文[3]中，Banerjee 等基于粗糙集粗糙度的思想对模糊集粗糙度进行了研究。

设 $\langle U,R\rangle$ 为有限近似空间，F 为 U 上的模糊集，(F^*,F_*) 为相应的粗糙模糊集。对于给定的参数 $0<\beta\leqslant\alpha\leqslant1$ 可得到 F_* 的 α 截集 $(F_*)_\alpha$ 和 F^* 的 β 截集 $(F^*)_\beta$，分别称为 F 在有限近似空间 $\langle U,R\rangle$ 上的 α 下近似和 β 上近似。由此定义模糊集 F 的粗糙度为：

$$\rho_R^{\alpha,\beta}(F)=1-\frac{|(F_*)_\alpha|}{|(F^*)_\beta|}\tag{1}$$

对于这个粗糙度的度量方法，Huynh[5] 认为过于强地依赖参数 α 和 β 的选取，即人为因素很大。并指出了不足之处：

*)基金项目：国家自然科学基金(No. 60372071)；辽宁省教育厅高等学校科学研究基金(2004C031)；辽宁师范大学校基金。闫德勤　博士，教授，主要从事模式识别、数据挖掘和粗糙集等方面的研究。

(a)如果模糊集 F 在每个等价类 $X_i(i=1,2,\cdots,m)$ 中都有一个元素 x 使得 $\mu_F(x)<\alpha$，则 $\rho_R^{\alpha\beta}(F)=1$

(b)如果模糊集 F 在每个等价类 $X_i(i=1,2,\cdots,m)$ 中隶属度值都是常数，并且 $\alpha=\beta$，则 $\rho_R^{\alpha\beta}(F)=0$。

更进一步地，Huynh 在文[5]中以模糊集 F 在论域 U 上为常模糊集(即对于所有的 $x\in U,\mu_F(x)=\delta$ 常数)为特例说明利用文[3]给出的粗糙度量(1)式对 α 和 β 的不同选取得到差别较大的结果：对于 $\beta<\delta<\alpha,\rho_R^{\alpha\beta}(F)=1$；其它情况下，$\rho_R^{\alpha\beta}(F)=0$。

为此，Huynh 等在文[5]中给出了新的研究，并提出了新的粗糙度度量方法：

设模糊集 F 在论域 U 上隶属度的取值范围为 $rng(\mu_F)=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，其中 $\alpha_i>\alpha_{i+1}>0(i=1,2,\cdots,n-1)$。对于

$$F_i=\{x\in U|\mu_F(x)\geqslant\alpha_i\}$$

令 $m_F(F_i)=\alpha_i-\alpha_{i+1},(\alpha_{n+1}=0)$。

$(i=1,2,\cdots,n)$

当 F_i 为空集时，$m_F(F_i)=1-\alpha_1$。

粗糙度的计算方法为：

$$\hat{\rho}_R(F)=\sum_{i=1}^{n}m_F(F_i)\left(1-\frac{|R_-(F_i)|}{|R^-(F_i)|}\right)=\sum_{i=1}^{n}m_F(F_i)\rho_R(F_i) \tag{2}$$

在文[5]中 Huynh 认为模糊集粗糙度度量公式(2)不依赖人为给定的参数，具有很好的客观性，并用例子与公式(1)做了对比。

模糊集粗糙度度量公式(2)对公式(1)的改进不依赖人为给定的参数，因此具有很好的理论和应用意义。然而，公式(2)也存在着不足。以下我们将指出其不足之处，并提出改进的度量公式。

4 粗糙度量的改进

公式(2)的不足之处在于其中的 $R_-(F_i)$：如果模糊集 F 在某个等价类 $X_j(j=1,2,\cdots,m)$ 中有一个元素 x 使得 $\mu_F(x)=\alpha_k(1\leqslant k\leqslant n)$，则对于 $i\leqslant k$ 总有 $R_-(F_i)=0$。这意味着当上近似 $R^-(F_i)$ 一定时，决定模糊集 F 粗糙度大小的因素只有两个：每个等价类 $X_j(j=1,2,\cdots,m)$ 中的隶属度最小值和最大值，与模糊集在等价类其它元素上的隶属度没关系。这一点与 Huynh 对公式(1)所指出的不足之处(a)(见上一节)相类似。

下面以一个简单的例子来说明：

例 1 设论域 $U=\{x_1,x_2,\cdots,x_5\}$，U 上的等价类划分为：$X_1=\{x_1\}$；$X_2=\{x_2,x_2,x_4,x_5\}$。模糊集 F 在论域 U 上的隶属度分别为：$\mu_F(x_1)=0.8,\mu_F(x_2)=0.1,\mu_F(x_3)=0.9,\mu_F(x_4)=0.9,\mu_F(x_5)=0.9$。另一个模糊集 G 在论域 U 上的隶属度分别为：$\mu_G(x_1)=0.8,\mu_G(x_2)=0.1,\mu_G(x_3)=0.1,\mu_G(x_4)=0.1,\mu_G(x_5)=0.9$。由于模糊集 F 和模糊集 G 在等价类 X_2 上的隶属度具有相同的最大和最小值，因此利用公式(2)计算出的粗糙度相同。

从例 1 的计算结果可以看到，公式(2)的度量过于粗糙。为此，我们以下提出一种改进的方法(这里所用符号与本文中上面章节所用符号意义相同)：

$$\text{令 } R_*(F_i)=\sum_{x_j\in R(F_i)}^{n}\mu_F(x_j)$$

则改进的粗糙度量公式为

$$\tilde{\rho}_R(F)=\sum_{i=1}^{n}m_F(F_i)\left(1-\frac{|R_*(F_i)|}{|R^-(F_i)|}\right) \tag{3}$$

公式(3)具有以下性质：

(1)$0\leqslant\tilde{\rho}_R(F)\leqslant1$。

(2)当 F 是普通集合的时候，公式(3)与粗糙集的粗糙度相同，即 $\tilde{\rho}_R(F)=\rho_R(F)$。

(3)当且仅当模糊集 F 在论域 U 上为常模糊集(即对于所有的 $x\in U,\mu_F(x)=\delta$ 常数)时，其粗糙度为 0，即 $\tilde{\rho}_R(F)=0$。

以下以例子说明公式(3)的应用，同时也给出与公式(2)的比较结果。

例 2 设论域 $U=\{x_1,x_2,\cdots,x_6\}$，U 上的等价类划分为：

$X_1=\{x_1,x_2,x_3\}$，

$X_2=\{x_4,x_5,x_6\}$。

模糊集 F 在论域 U 上的隶属度分别为：$\mu_F(x_1)=1,\mu_F(x_2)=1,\mu_F(x_3)=0.75,\mu_F(x_4)=0.5,\mu_F(x_5)=0.25,\mu_F(x_6)=0$。

可以得到模糊集 F 在论域 U 上隶属度的取值范围为

$$rng(\mu_F)=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}=\{1,0.75,0.5,0.25\}。$$

$m_F(F_i)=0.25(i=1,2,3,4)$。

由 $\alpha_i(i=1,2,3,4)$ 得到的截集分别为

$F_1=\{x_1,x_2\}$，

$F_2=\{x_1,x_2,x_3\}$，

$F_3=\{x_1,x_2,x_3,x_4\}$，

$F_4=\{x_1,x_2,x_3,x_4,x_5\}$。

因此得到

$R^-(F_1)=R^-(F_2)=X_1$，

$R^-(F_3)=R^-(F_4)=\{X_1,X_2\}=U$，

$|R^-(F_1)|=|R^-(F_2)|=3$，

$|R^-(F_3)|=|R^-(F_4)|=6$，

$|R_*(F_1)|=|R_*(F_2)|=2.75$，

$|R_*(F_3)|=|R_*(F_4)|=3.5$。

代入公式(3)得到

$\tilde{\rho}_R(F)=0.479$。

而使用公式(2)得到的计算结果是 0.75。

例 3 使用和例 1 相同的条件。和例 2 的方法一样，利用公式(3)分别计算模糊集 F 和模糊集 G 的粗糙度得到

$\tilde{\rho}_R(F)=0.254$，

$\tilde{\rho}_R(G)=0.575$。

由例 1 知，使用公式(2)计算出模糊集 F 和模糊集 G 的粗糙度是相同的。而用公式(3)计算的结果就可以看出两个粗糙集粗糙度的差别。

结合例 2、例 3 的结果，同时结合对所计算的粗糙集隶属度分布的分析，可以看出公式(3)是合理的，是对公式(2)的一个有价值的改进。

结论 模糊集粗糙度的研究对于粗糙集与模糊集的结合具有理论和应用意义。本文深入分析了 Huynh 和 Banerjee 提出的粗糙度量存在缺陷的原因，提出了一种改进方法。新提出的改进方法克服了原有方法的不足，对于进一步的研究和应用具有重要性。

参考文献

1 Pawlak Z. Rough sets. International Journal of Computer and Information sciences，1982，11：341～356
2 Pawlak Z. Rough sets：theoretical aspects of reasoning about data. Kluwer Academic Publishers，Dordrecht，1991
3 Banerjee M，Pal S K. Roughness of a fuzzy set. Information sciences，1996，93：235～246
4 Dubois D，Prade H. Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems，1990，17：191～209
5 Huynh V H，Nakamori Y. A roughness measure for fuzzy sets. Information Sciences，2005，173：255～275
6 Zadeh L A. Fuzzy sets. Information and Control，1965，8：338～353
7 曾黄麟. 粗集理论及其应用(修订版). 重庆：重庆大学出版社，1998