Vague 值(集)相似度量的研究

张清华
(重庆邮电大学计算机科学与技术学院 重庆 400065)
(西南交通大学信息科学与技术学院 成都 610031)

摘 要： 该文对现有 Vague 集(值)相似度量方法进行研究，发现目前 Vague 值(集)相似度量方法存在一些缺陷，给出了 Vague 集(值)相似度的一种定义，提出了两种 Vague 值(集)相似度量的改进方法，最后通过数据分析验证了改进方法的有效性和优越性。

关键词：Vague 值(集)；相似度量；隶属函数；模糊集
中图分类号：TP391 **文献标识码**：A **文章编号**：1009-5896(2007)08-1855-05

Study on Similarity Measure between Vague Values (Sets)

Zhang Qing-hua
(*College of Computer Science & Technology*, *Chongqing University of Posts and Telecommunications*, *Chongqing* 400065, *China*)
(*School of Information Science & Technology*, *Southwest Jiaotong University*, *Chengdu* 610031, *China*)

Abstract: The methods of similarity measure between vague values (sets) in existence, some errors and limitations in some former result are analyzed. A new definition of similarity measure between values (sets) is presented, and two ameliorative similarity measures between values (sets) are presented. Finally, the ameliorative similarity measures between values (sets) are validated with experimental data.

Key words: Vague values (sets); Similarity measure; Membership function; Fuzzy sets

1 引言

自从 1965 年Zadeh提出模糊集(Fuzzy Sets)理论[1]以来，利用模糊集理论来处理模糊性和不确定性的信息和数据已经取得重大进展。由于在处理实际问题(如投票模型)往往需要同时得知支持和反对的程度信息，这时模糊集理论就无法很好解决这类模糊性问题。Gau和Buehrer于 1993 年提出了Vague集理论[2]，它将模糊集理论中的隶属函数单值扩充为一个区间$[t_A(x),1-f_A(x)]$，其中$t_A(x)$表示真隶属函数，$f_A(x)$表示假隶属函数，且满足$t_A(x)\le\mu_A(x)\le 1-f_A(x)$。Vague集理论用这种真假隶属函数的定义体现了元素对模糊概念的属于与不属于的程度和证据，比模糊集有更强的表达不确定性的能力，进一步加强了对客观世界描述的真实程度。Vague集理论的提出得到许多研究者的关注，目前该理论在模糊控制、决策、故障诊断、近似推理和医疗诊断等方面取得了较好的应用[3-5]。近年来许多学者对Vague集(值)之间的相似度量问题进行了研究，本文在前人的研究基础上，分析了文献[6-8]对的相似度量研究的某些不足，给出了相似度的一般定义，并提出了两种改进的相似度量方法，最后通过数据分析比较验证了改进方法的优越性和合理性。

论文余下部分的安排如下：第 2 节介绍了 Vague 集(值)相似度量的一些基本方法，并指出其不足之处，第 3 节给出了 Vague 集(值)相似度的一般定义，第 4 节提出了两种改进度的相似度量方法并给出证明过程，第 5 节给出了数据来验证结论的有效性，最后是结束语。

2006-01-04 收到，2006-08-07 改回
重庆市高等学校优秀中青年骨干教师资助计划(D2005-29)资助课题

2 Vague 集(值)相似度量

设U是一个对象空间，其中的任意一个元素用x表示，U上的Vague集[2]A是用一个真隶属函数$t_A(x)$和一个假隶属函数$f_A(x)$表示。$t_A(x)$表示从支持x的证据所得出的隶属度下界，$f_A(x)$表示从反对x的证据所得出的隶属度下界，$t_A(x)$和$f_A(x)$均是[0,1]中的某个实数值，且满足$0\le t_A(x)+f_A(x)\le 1$。从Vague集A的定义可以看出，任意元素x的Vague值是一个[0,1]的子区间$[t_A(x),1-f_A(x)]$，如果$t_A(x)=1-f_A(x)$，则Vague集A就退化为模糊集，其隶属函数$\mu_A(x)=t_A(x)=1-f_A(x)$，如果$t_A(x)=1-f_A(x)=0$或者$t_A(x)=1-f_A(x)=1$，则Vague集A就退化为经典的Cantor集了。为叙述简介方便，在不产生混淆的情况下，本文用t_x表示$t_A(x)$，用f_x表示$f_A(x)$。

Chen 在文献[9]中给出了两个 Vague 集(值)的相似度量方法，具体如下：定义核函数$S(x)=t_x-f_x$，Vague 值x,y之间的相似度量公式：

$$M(x,y)=1-\left|\frac{S(x)-S(y)}{2}\right| \tag{1}$$

又设 A,B 分别是 U 上的两个Vague集，令 $V_A(u_i)=[t_A(u_i),1-f_A(u_i)]$，$V_B(u_i)=[t_B(u_i),1-f_B(u_i)]$，则Vague集 A,B 之间的相似度量公式定义为[9]：

$$T(A,B)=\frac{1}{n}\sum_{i=1}^{n}M(V_A(u_i),V_B(u_i))=\frac{1}{n}\sum_{i=1}^{n}\left(1-\left|\frac{S(V_A(u_i))-S(V_B(u_i))}{2}\right|\right) \tag{2}$$

Chen 在文献[9]中认为 $M(x,y)$ 越大，则两个 Vague 值之间的相似程度越大，$T(A,B)$ 越大，则两个 Vague 集之间的相似程度越大。为了方便叙述，记式(1)为 M_C。李凡等人在文献[6]中指出了用 M_C 测量 Vague 值之间的相似程度存在的缺陷，并在文献[6]提出了两个 Vague 值 x,y 之间新的相似度量方法，为了叙述方便，记为 M_L：

$$M_L(x,y)\equiv M(x,y)=1-\left|\frac{S(x)-S(y)}{4}\right|-\frac{|t_x-t_y|+|f_x-f_y|}{4} \tag{3}$$

在 M_L 的基础上，李艳红等人在文献[10]中提出了一种新的基于距离公式的相似度量公式，为了叙述方便，记为 M_O：

$$M_O(x,y)\equiv M(x,y)=1-\sqrt{\frac{(t_x-t_y)^2+(f_x-f_y)^2}{2}} \tag{4}$$

由于 M_O 对 Vague 值之间的相似度的区分力还是不尽如人意，闫德勤等人在文献[8]中认为 M_O 公式中只包含支持度(真隶属值) t_x 和反对度(假隶属值) f_x，对弃权因素 $m_x=1-t_x-f_x$ 没有考虑，并提出了基于欧氏距离的相似度量方法，为了方便，记为 M_R：

$$M_R(x,y)\equiv M(x,y)=1-\sqrt{\frac{(t_x-t_y)^2+(f_x-f_y)^2+(m_x-m_y)^2}{2}} \tag{5}$$

分析发现 M_R 虽然考虑了弃权因素 $m_x=1-t_x-f_x$，但对相似度的区分力并没有太大的改善。另外文献[8]中指出 $M_R(x,y)=0$ 的充要条件是 x,y 为[0,0]，[0,1]和[1,1]中的任意两个，然而Vague值[0,0]和 [1,1]之间的相似度为0是合理的，但[0,0]和 [0,1] 之间的相似度为 0 以及[1,1]和 [0,1] 之间的相似度为 0 不符合人们的直觉，因为在投票模型中[0,0]表示所有人都反对，[0,1]表示没有人支持，也没有人反对，[1,1]表示所有人都支持。刘华文在文献[7]考虑到弃权部分中有部分趋向投赞成票，也有部分趋向投反对票，因此对弃权部分 m_x 进一步细化成3个部分 (t_xm_x,f_xm_x,m_xm_x)，于是元素 x 的隶属情况用三维表示变为 $(t'_x,f'_x,m'_x)=(t_x+t_xm_x,f_x+f_xm_x,m_xm_x)$。并给出新的度量公式 M_W：

$$M_W(x,y)=\frac{|(t'_x-f'_x)-(t'_y-f'_y)|+2|(t'_x+f'_x)-(t'_y+f'_y)|}{4} \tag{6}$$

显然 $M_W(x,y)$ 提高了相似度的区分力。但人们在投票模型中的一般认为[0.4,0.8]与[0.3,0.8]的相似度应该比[0.4,0.8]与[0.3,0.7]的相似度大，但用 $M_W(x,y)$ 计算可知[0.4,0.8]与[0.3,0.8]的相似度应该比[0.4,0.8]与[0.3,0.7]的相似度小，这与实际的直觉不符合。

3 Vague 值(集)相似度的定义

我们在研究大量的文献时，发现人们对相似度的概念理解不一致，都是先给出一个公式，然后再证明它满足一些定理即可，而不同研究者对公式应该满足那些定理(或条件)才算是相似度量公式的认识不尽相同。为了更好的度量 Vague 值(集)之间的相似度，本文参考模糊集的相似度的定义方式[12]，给出了Vague值(集)之间相似度的一种定义。

定义 1 (Vague 值之间的相似度) 设 $x=[t_x,1-f_x]$，$y=[t_y,1-f_y]$ 是 Vague 集 A 上的两个 Vague 值，这里 $t_x\in[0,1]$，$f_x\in[0,1]$，$t_y\in[0,1]$，$f_y\in[0,1]$，且 $t_x+f_x\le 1$，$t_y+f_y\le 1$，定义映射：$M:A\times A\to[0,1]$，即 $(x,y)\to M(x,y)$。称 $M(x,y)$ 是 Vague 值 x,y 之间的相似度，如果 $M(x,y)$ 满足以下条件：

(1) 对任意的 $x,y\in A$，$0\le M(x,y)\le 1$ (有界性)；

(2) $M(x,y)=M(y,x)$ (对称性)；

(3) 对任意的 $x,y\in A$，$M(x,x)=1$；$M(x,y)=0$ 的充要条件是 x,y 分别是[0,0]，[1,1]或者[1,1],[0,0]；

(4) 用 $d(x,y)=1-M(x,y)$ 表示 x,y 之间的距离，则对任意的 $x,y,z\in A$，有

$$d(x,y)+d(y,z)\ge d(x,z)$$

定义 2 (Vague 集之间的相似度) 设 A,B 是论域 U 上的两个 Vague 集，

$$A=[t_A(u_1),1-f_A(u_1)]/u_1+[t_A(u_2),1-f_A(u_2)]/u_2+\cdots+[t_A(u_n),1-f_A(u_n)]/u_n$$

$$B=[t_B(u_1),1-f_B(u_1)]/u_1+[t_B(u_2),1-f_B(u_2)]/u_2+\cdots+[t_B(u_n),1-f_B(u_n)]/u_n$$

定义映射：$T:U\times U\to[0,1]$，即 $(A,B)\to T(A,B)$。称 $T(A,B)$ 是 Vague 集 A,B 的相似度，如果 $T(A,B)$ 满足如下条件：

(1) 对任意的 $A,B\in U$，$0\le T(A,B)\le 1$ (有界性)；

(2) 对称性：对任意的 $A,B\in U$，$T(A,B)=T(B,A)$ (对称性)；

(3) 对任意的 $A,B\in U$，$T(A,A)=1$；$T(A,B)=0$ 的充要条件是 $A=\sum_{i=1}^{n}[a_i,a_i]/u_i$，$B=\sum_{i=1}^{n}[b_i,b_i]/u_i$，$a_i+b_i=1$ 且 $a_i,b_i\in\{0,1\}$ ($i=1,2,\cdots,n$)。

4 改进的相似度量公式

4.1 两种改进的 Vague 值之间的相似度量

有了 Vague 值的相似度的一般定义，为了更好度量 Vague 值之间的相似程度，本文在总结前人研究的基础上，给出了两种改进方法。

方法 1 设 $x=[t_x,1-f_x]$，$y=[t_y,1-f_y]$ 是 Vague 集 A 上 Vague 值，这里 $t_x\in[0,1]$，$f_x\in[0,1]$，$t_y\in[0,1]$，$f_y\in[0,1]$，且 $t_x+f_x\le 1$，$t_y+f_y\le 1$。定义公式 M_*：

$$M_*(x,y)=1-\lambda|(t_x-f_x)-(t_y-f_y)|-(1/2-\lambda)\cdot(|t_x-t_y|+|f_x-f_y|),\ 0\le\lambda\le 1/2 \tag{7}$$

下面证明 M_* 是 Vague 值 x,y 之间的相似度。

证明 因为 $0 \le t_x \le 1, 0 \le f_x \le 1$，因此 $t_x - f_x \in [-1,1]$，$t_y - f_y \in [-1,1]$，所以 $|(t_x - f_x) - (t_y - f_y)| \in [0,2]$，$|(t_x - f_x)| + |(t_y - f_y)| \in [0,2]$，即 $M(x,y) = 1 - \lambda|(t_x - f_x) - (t_y - f_y)|$ $1 - \lambda[0,2] - (1/2 - \lambda)\ [0,2] \in [0,1]$，因此条件(1)(有界性)满足。$M(x,y) = M(y,x)$ 显然，条件(2)满足。对任意的 $x,y \in A$，$M(x,x) = 0$ 显然成立。如果 x,y 分别是 $[0,0]$，$[1,1]$ 或者 $[1,1],[0,0]$，显然有 $M(x,y) = 0$。反之如果 $M(x,y) = 0$，即 $\lambda|(t_x - f_x) - (t_y - f_y)| + (1/2 - \lambda)\ (|t_x - t_y| + |f_x - f_y|) = 1$。因为 $|(t_x - f_x) - (t_y - f_y)| \in [0,2]$，$|(t_x - f_x)| + |(t_y - f_y)| \in [0,2]$，且 $1/2 \ge \lambda \ge 0, 1/2 - \lambda \ge 0$，所以必有 $|(t_x - f_x) - (t_y - f_y)| = 2$，$|(t_x - f_x)| + |(t_y - f_y)| = 2$，即有 $[t_x, 1 - f_x] = [0,0]$ 且 $[t_y, 1 - f_y] = [1,1]$ 或者 $[t_x, 1 - f_x] = [1,1]$ 且 $[t_y, 1 - f_y] = [0,0]$，因此条件(3)满足。设 $d(x,y) = 1 - M(x,y) = \lambda|(t_x - f_x) - (t_y - f_y)| + (1/2 - \lambda)(|t_x - t_y| + |f_x - f_y|)$，对任意的 $x,y,z \in A$，有 $d(x,y) + d(y,z) = \lambda|(t_x - f_x) - (t_y - f_y)| + (1/2 - \lambda)(|t_x - t_y| + |f_x - f_y|) + \lambda|(t_y - f_y) - (t_z - f_z)| + (1/2 - \lambda)(|t_y - t_z| + |f_y - f_z|) = \lambda(|(t_x - f_x) - (t_y - f_y)| + |(t_y - f_y) - (t_z - f_z)|) + (1/2 - \lambda)(|t_x - t_y| + |f_x - f_y| + |t_y - t_z| + |f_y - f_z|) \ge \lambda|(t_x - f_x) - (t_z - f_z)| + (1/2 - \lambda)(|t_x - t_z| + |f_x - f_z|) = d(x,z)$，所以条件(4)成立。

证毕

在 M_* 中当 $\lambda = 0$ 时 $M_*(x,y) = 1 - (1/2)(|t_x - t_y| + |f_x - f_y|)$，它就变成是 Hong 等人在文献[11]中提出的相似度量公式 $M_H(x,y)$；同理，当 $\lambda = 1/2$ 时 $M_*(x,y)$ 就变成是 Chen 等人在文献[9]中提出的相似度量公式 $M_C(x,y)$；当 $\lambda = 1/4$ 时 $M_*(x,y)$ 就变成是李凡等人在文献[6]中提出的相似度量公式 $M_L(x,y)$。因此式(7)是 $M_H(x,y)$，$M_C(x,y)$，$M_L(x,y)$ 的推广，具有更广的使用范围，且相似度的区分力强于 $M_H(x,y)$，$M_C(x,y)$，$M_L(x,y)$ (比较数据见表 1)。这里不同 λ 值对应的 λ 和 $1/2 - \lambda$ 相当于 $|(t_x - f_x) - (t_y - f_y)|$ 和 $|t_x - t_y| + |f_x - f_y|$ 对应的权重，研究者可以根据问题的侧重点不同而灵活选择 λ 的值，只要满足条件($0 \le \lambda \le 1/2$)即可。

方法 2 设 $x = [t_x, 1 - f_x]$，$y = [t_y, 1 - f_y]$ 是 Vague 集 A 上的两个 Vague 值，这里 $t_x \in [0,1]$，$f_x \in [0,1]$，$t_y \in [0,1]$，$f_y \in [0,1]$，且 $t_x + f_x \le 1$，$t_y + f_y \le 1$。由于方法 1 只考虑了支持度 t_x，反对度 f_x，对踌躇度 m_x 没有体现出来，因此它的相似度的分辨能力有时还是无法满足实际需要。根据投票模型考虑到踌躇(弃权)部分 $m_x = 1 - t_x - f_x$ 中仍然有部分人趋向投赞成票，也有部分人趋向投反对票，由文献[7]得到启发，因此对弃权部分 m_x 进一步细化成 3 部分 $(t_x m_x, f_x m_x, m_x m_x)$，于是元素 x 的隶属情况用三维表示变为 $(t'_x, f'_x, m'_x) = (t_x + t_x m_x, f_x + f_x m_x, m_x m_x)$。设核函数 $S(x) = t_x - f_x$，$S'(x) = t'_x - f'_x$。定义公式 M_{**}：

$$M_{**}(x,y) = 1 - \lambda(|t_x - t_y| + |f_x - f_y|) - \alpha|S(x) - S(y)| - \beta|S'(x) - S'(y)| \tag{8}$$

其中($\lambda \ge 0, \alpha \ge 0, \beta \ge 0$ 且 $\lambda + \alpha + \beta = 1/2$)。下面证明式(8)是 Vague 值 x,y 之间的相似度。

证明 对任意的 $x,y \in A$，$|(t_x - f_x)| + |(t_y - f_y)| \in [0,2]$，$|(t_x - f_x) - (t_y - f_y)| \in [0,2]$，$|(t'_x - f'_x) - (t'_y - f'_y)| \in [0,2]$，所以 $M_{**}(x,y) = 1 - \lambda[0,2] - \alpha[0,2] - \beta[0,2] = 1 - (\lambda + \alpha + \beta)[0,2] = 1 - (1/2)[0,2] \in [0,1]$，条件(1)(有界性)满足。条件(2)(对称性)显然成立。对任意的 $x,y \in A$，$M(x,x) = 1$ 显然成立。如果 x,y 分别是 $[0,0]$，$[1,1]$ 或者 $[1,1]$，$[0,0]$，均 $M(x,y) = 1 - \lambda \cdot(|t_x - t_y| + |f_x - f_y|) - \alpha|S(x) - S(y)| - \beta|S'(x) - S'(y)| = 1 - 2\lambda - 2\alpha - 2\beta = 1 - 2(\lambda + \alpha + \beta) = 0$，反之如果 $M(x,y) = 1 - \lambda(|t_x - t_y| + |f_x - f_y|) - \alpha|S(x) - S(y)| - \beta|S'(x) - S'(y)| = 0$，即 $\lambda(|t_x - t_y| + |f_x - f_y|) + \alpha|S(x) - S(y)| + \beta|S'(x) - S'(y)| = 1$，因为 $|(t_x - f_x)| + |(t_y - f_y)| \in [0,2]$，$|S(x) - S(y)| \in [0,2]$，$|S'(x) - S'(y)| \in [0,2]$，且 $\lambda \ge 0, \alpha \ge 0, \beta \ge 0$，所以 $|(t_x - f_x)| + |(t_y - f_y)| = 2$，$|S(x) - S(y)| = 2$，$|S'(x) - S'(y)| = 2$，必有 $[t_x, 1 - f_x] = [0,0]$ 且 $[t_y, 1 - f_y] = [1,1]$ 或者 $[t_x, 1 - f_x] = [1,1]$ 且 $[t_y, 1 - f_y] = [0,0]$，因此条件(3)满足。

设 $d(x,y) = 1 - M(x,y) = \lambda(|t_x - t_y| + |f_x - f_y|) + \alpha \cdot |S(x) - S(y)| + \beta|S'(x) - S'(y)|$，对任意的 $x,y,z \in A$，有 $d(x,y) + d(y,z) = \lambda(|t_x - t_y| + |f_x - f_y|) + \alpha|S(x) - S(y)| + \beta|S'(x) - S'(y)| + \lambda(|t_y - t_z| + |f_y - f_z|) + \alpha|S(y) - S(z)| + \beta|S'(y) - S'(z)| \ge \lambda(|t_x - t_z| + |f_x - f_z|) + \alpha|S(x) - S(z)| + \beta|(S'(x) - S'(z)| = d(x,z)$，所以条件(4)满足。

证毕

在式(8)中，当 $\alpha = 0, \beta = 0$ 时，M_{**} 就退化为 Hong 等人在文献[11]中提出的相似度量公式 $M_H(x,y)$；当 $\lambda = 0$，$\beta = 0$ 时，M_{**} 就退化为 Chen 等人在文献[9]中提出的相似度量公式 M_C。$\lambda = 1/4, \alpha = 1/4, \beta = 0$，式(8)中的 M_{**} 就退化为李凡等人在文献[6]中提出的相似度量公式 M_L；当 $\beta = 0$ 时，M_{**} 就退化为方法 1 提出的相似度量公式 M_*。因此式(8)包括了支持度 t_x，反对度 f_x 和踌躇度 m_x 因素，它是 $M_H(x,y)$，$M_C(x,y)$，$M_L(x,y)$，$M_*(x,y)$ 更一般的形式，它的应用范围和相似度的区分力都优于 $M_H(x,y)$，$M_C(x,y)$，$M_L(x,y)$ 和 $M_*(x,y)$ 等公式(比较数据见表 1)。

4.2 改进的 Vague 集之间的相似度量

对应 4.1 节中的两种改进Vague值之间相似度量公式，本文给出两种相应的Vague集之间相似度量公式。设 A,B 是论域 U 上的两个Vague集[9]：

$$A = [t_A(u_1), 1 - f_A(u_1)]/u_1 + [t_A(u_2), 1 - f_A(u_2)]/u_2 + \cdots + [t_A(u_n), 1 - f_A(u_n)]/u_n$$

$$B=[t_B(u_1),1-f_B(u_1)]/u_1+[t_B(u_2),1-f_B(u_2)]/u_2+\cdots+[t_B(u_n),1-f_B(u_n)]/u_n$$

令 $S_A(u_i)=t_A(u_i)-f_A(u_i)$，$S_B(u_i)=t_B(u_i)-f_B(u_i)$，$S'_A(u_i)=(t_A(u_i)-f_A(u_i))(2-t_A(u_i)-f_A(u_i))$，$S'_B(u_i)=(t_B(u_i)-f_B(u_i))(2-t_B(u_i)-f_B(u_i))$ ($i=1,2,\cdots,n$)，定义两个公式：

$$T_*(A,B)=\frac{1}{n}\sum_{i=1}^{n}\left(1-\lambda\left(\left|S_A(u_i)-S_B(u_i)\right|-\left(\frac{1}{2}-\lambda\right)\cdot\left(\left|t_A(u_i)-t_B(u_i)\right|+\left|f_A(u_i)-f_B(u_i)\right|\right)\right)\right),0\le\lambda\le\frac{1}{2} \quad (9)$$

$$T_{**}(A,B)=\frac{1}{n}\sum_{i=1}^{n}\Big(1-\lambda(\left|t_A(u_i)-t_B(u_i)\right|+\left|f_A(u_i)-f_B(u_i)\right|)-\alpha\left|S_A(u_i)-S_B(u_i)\right|-\beta\left|S'_A(u_i)-S'_B(u_i)\right|\Big),$$
$$\lambda\ge0,\alpha\ge0,\quad\beta\ge0\text{且}\lambda+\alpha+\beta=\frac{1}{2} \quad (10)$$

与第 4.1 节中改进的 Vague 值之间的相似度量公式 M_* 和 M_{**} 的证明相似，它们都是 Vague 集 A,B 之间的相似度，由于篇幅有限这里不再证明。

值得指出的是文献[6]中定理 6 认为 $T(A,B)=0$ 的充要条件是 $A=\sum_{i=1}^{n}[0,0]/u_i$，$B=\sum_{i=1}^{n}[1,1]/u_i$ 或者 $A=\sum_{i=1}^{n}[1,1]/u_i$，$B=\sum_{i=1}^{n}[0,0]/u_i$ ($i=1,2,\cdots,n$)。通过分析发现该定理有不完整之处。例如如果 $A=[1,1]/u_1+[0,0]/u_2$，$B=[0,0]/u_1+[1,1]/u_2$ 根据文献[6]中的公式(5)可知 $T(A,B)=0$。因此 $A=\sum_{i=1}^{n}[0,0]/u_i$，$B=\sum_{i=1}^{n}[1,1]/u_i$ 或者 $A=\sum_{i=1}^{n}[1,1]/u_i$，$B=\sum_{i=1}^{n}[0,0]/u_i$ ($i=1,2,\cdots,n$) 是 $T(A,B)=0$ 充分非必要条件。通过分析不难证明 $T_*(A,B)=0$ 的充要条件是 $A=\sum_{i=1}^{n}[a_i,a_i]/u_i$，$B=\sum_{i=1}^{n}[b_i,b_i]/u_i$，满足 $a_i+b_i=1$，且 $a_i,b_i\in\{0,1\}$ ($i=1,2,\cdots,n$)。

5 数据分析

为了进一步说明 M_* 和 M_{**} 在度量两个 Vague 值之间的相似度的优越性，我们通过一组实验数据来对这些公式进行比较(在 M_* 中取 $\lambda=1/3$，在 M_{**} 中取 $\lambda=0.25$，$\alpha=0.125$，$\beta=0.125$)，这里 λ，α 和 β 的值相当于 $|(t_x-f_x)-(t_y-f_y)|$，$|S(x)-S(y)|$ 和 $|S'(x)-S'(y)|$ 在公式中对应的权重，读者可以根据问题的侧重点不同而灵活选择 λ，α 和 β 的值，只要满足条件 $\left(\lambda\ge0,\alpha\ge0,\beta\ge0\text{且}\lambda+\alpha+\beta=\frac{1}{2}\right)$ 即可。具体数据如表 1。

表 1　Vague 值 x,y 之间的相似度

	1	2	3	4	5	6	7	8	9	10	11	12
x	[0.4,0.8]	[0.4,0.8]	[0.4,0.8]	[0.4,0.8]	[0.4,0.8]	[0.4,0.8]	[0.4,0.8]	[0,0]	[0,1]	[0,1]	[0,1]	[1,1]
y	[0.3,0.7]	[0.3,0.8]	[0.3,0.9]	[0.4,0.7]	[0.4,0.9]	[0.5,0.7]	[0.5,0.8]	[1,1]	[1,1]	[0,0]	[0.5,0.5]	[0.5,0.5]
M_C	0.9	0.95	1	0.95	0.95	1	0.95	0	0.5	0.5	1	0.5
M_H	0.9	0.95	0.9	0.95	0.95	0.9	0.95	0	0.5	0.5	0.5	0.5
M_L	0.9	0.95	0.95	0.95	0.95	0.95	0.95	0	0.5	0.5	0.75	0.5
M_O	0.9	0.9293	0.9	0.9293	0.9293	0.9	0.9293	0	0.2929	0.2929	0.5	0.5
M_R	0.9	0.9	0.8268	0.9	0.9	0.8268	0.9	0	0	0	0.1340	0.1340
M_W	0.93	0.9225	0.89	0.9275	0.9125	0.93	0.9375	0	0.25	0.25	0.5	0.75
M_*	0.9	0.95	0.9667	0.95	0.95	0.9667	0.95	0	0.5	0.5	0.8333	0.5
M_{**}	0.89	0.9463	0.945	0.9438	0.9413	0.94	0.9488	0	0.5	0.5	0.75	0.5

从表 1 中的数据得到：M_C 虽然公式简洁，但对相似度的度量会产生与人们自觉不符合的情况，如从表 1 中得知[0.4,0.8]与[0.3,0.9], [0.4,0.8]与[0.5,0.7]以及[0,1] 与[0.5,0.5]的相似度均为 1，这明显与实际不符合。M_H 考虑因素太少，所以对相似度的区分能力很弱。M_O 与 M_R 对相似度的区分能力有所提高，但 M_R 认为[0,1]与[1,1]，[0,1]与[0,0]的相似度为 0，这显然不符合在投票模型中的直觉认识，M_L 与 M_* 对相似度的区分力进一步增强，但还是不能满足实际需要，如用公式 M_L 和 M_* 计算 Vague 值 [0.4,0.8]分别与[0.3,0.8]，[0.4,0.7]，[0.4,0.9]，[0.5,0.8]的相似度均为 0.95，这说明调节 M_* 中的系数 λ 的值还是不能根本解决相似度的分辨能力，但适当选取 M_* 中 λ 的值可以有效提高 M_* 的相似度区分力。M_W 和 M_{**} 对两个 Vague 值 x,y 之间的相似度的度量上分辨能力大大提高，但用 M_W 在计算[0.4,0.8]与[0.3,0.7]以及[0.4,0.8]与[0.3,0.8]的相似度时，得出[0.4,0.8]与[0.3,0.7]相似程度高于[0.4,0.8]与[0.3,0.8]的相似度，在计算[0,1]与[0.5,0.5]以及[1,1]与[0.5,0.5]的相似度时，得出[0,1]与[0.5,0.5]的相似度低于[1,1]与[0.5,0.5]的相似度，这些都与用其它公式的计算的结果违背，也与实际的投票模型不相符。相反用 M_{**} 得到的相似结果更加符合实际，原因在于它同时考虑了支持度 t_x，反对度 f_x 和踌躇度 m_x 对相似程度的影响，综合了其它公式的优点，避免了其它公式的考虑不全的弊端。由于篇幅有限这里不再举例验证 $T_*(A,B)$ 和 $T_{**}(A,B)$ 在度量 Vague 集

相似度的优越性了，读者可以自己举例验证。

6 结束语

模糊集之间的相似度量研究得比较成熟，随着 Vague 集理论研究的不断深入，Vague 值(集)之间的相似度量的研究也越来越受到人们的重视。本文在总结前人的研究基础上，给出了 Vague 值(集)之间的相似度的一种一般定义，并提出了两种改进的相似度量方法。实验数据表明改进的方法更符合实际，特别是度量公式 M_{**}，它同时考虑了支持度 t_x，反对度 f_x 和踌躇度 m_x 对相似度的影响，对 Vague 值之间的相似性的区分力大大增强。这些方法对 Vague 集的研究和应用都具有一定的实际意义。

参 考 文 献

[1] Zadeh L A. Fuzzy Sets. *Information and Control*, 1965, 8(3): 338-353.

[2] Gau Wen-Lung and Buehrer Daniel J. Vague sets. *IEEE Transactions on Systems, Man and Cybernetics*, 1993, 23(2): 610-614.

[3] 马志锋，刑汉承. Vague 决策表中的含糊规则获取策略. 计算机学报, 2001, 24(4): 382-389.
Ma Zhi-feng and Xing Han-cheng. Strategies of ambiguous rule acquisition from Vague decision table. *Chinese Journal Computer*, 2001, 24(4): 382-389. (in Chinese)

[4] 李凡，田应忠，吕泽华. 基于 Vague 集的近似推理方法. 华中科技大学学报(自然科学版),2004, 32(4): 44-46.
Li Fan, Tian Ying-zhong, and Lu Ze-hua. An approximation reasoning approach based on the measures of similarity between Vague sets. *J. Huazhong Univ. of Sci.& Tech.*(*Nature Science Edition*), 2004, 32(4): 44-46. (in Chinese)

[5] 李凡，蔡立晶，田应忠. 基于 Vague 集的医疗诊断系统. 华中科技大学学报(自然科学版), 2002, 30(10): 47-49.
Li Fan, Cai Li-jing, and Tian Ying-zhong. Medical diagnosis system based on Vague sets. *J.Huazhong Univ. of Sci.& Tech.*(Nature Science Edition), 2002, 30(10): 47-49. (in Chinese)

[6] 李凡， 徐章艳. Vague 集之间的相似度量. 软件学报, 2001, 12(6): 922-927.
Li Fan and Xu Zhang-yan. Measures of similarity between Vague Sets. *Journal of Software*, 2001,12(6): 922－927.(in Chinese)

[7] 刘华文. 模糊模式识别的基础——相似度量. 模式识别与人工智能, 2004, 17(2): 141-145.
Liu Hua-wen. Basis of fuzzy pattern recognition: Similarity measures. *Journal of Pattern Recognition and Artificial Intelligence*, 2004, 17(2): 141-145. (in Chinese)

[8] 闫德勤，迟忠先，艳红.关于 Vague 集的相似度量. 模式识别与人工智能, 2004, 17(1): 22-26.
Yan De-qin, Chi Zhong-xian, and Yan Hong. Measures of similarity between vague sets. *Journal of Pattern Recognition and Artificial Intelligence*, 2004,17(1): 22-26. (in Chinese)

[9] Chen S M. Measures of similarity between vague sets. *Fuzzy Sets and Systems*, 1995, 74(2): 217-223.

[10] 李艳红，迟忠先，阎德勤. Vague 相似度量与 Vague 熵. 计算机科学, 2002, 29(12): 129-132.
Li Yan-hong,Chi Zhong-xian,and Yan De-qin. Similarity measures and entropy for vague sets. *Computer Science*, 2002, 29(12): 129-132. (in Chinese)

[11] Hong D H and Kim C. A note on similarity measures between vague sets and between elements. *Information Sciences*,1999, 115(1): 83-96.

[12] Wen Y Z and Hong X L. Relationship between measure of fuzziness and measure of similarity. *The Journal of Fuzzy Mathematics* 2004,12(1): 207-213.

张清华： 男，1974 年生，博士生，研究方向包括粗糙集理论、不确定信息处理.