

一种新区间值 vague 集及其在模式识别中的应用

张振华^{1,2} 杨静宇¹ 叶有培¹ 张倩生²

(南京理工大学计算机科学与技术学院 南京 210094)¹ (广东外语外贸大学思科信息学院 广州 510006)²

摘 要 基于现有定义证明了区间值 vague 集并非 vague 集的扩展,并提出了一种新的区间值 vague 集定义。证明了 vague 集和区间值 vague 集都是新区间值 vague 集的特例。进而提出了带参数区间值 vague 集的概念,并从已知隶属度和非隶属度出发,重点分析了带参数区间值 vague 集的构造。将模式识别与医疗诊断案例进行了对比。仿真实验结果表明,带参数区间值的 vague 集方法比传统的 vague 集方法更全面且更具灵活性。

关键词 vague 集,区间值 vague 集,新区间值 vague 集,带参数区间值 vague 集,模式识别

中图法分类号 TP391 **文献标识码** A

Novel Interval Valued Vague Sets and its Application in Pattern Recognition

ZHANG Zhen-hua^{1,2} YANG Jing-yu¹ YE You-pei¹ ZHANG Qian-sheng²

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)¹

(Cisco School of Informatics, Guangdong University of Foreign Studies, Guangzhou 510006, China)²

Abstract First, it was proved that interval valued vague sets is not the generalization of vague sets based on the conventional definition in this paper, and a novel interval valued vague sets were presented. Then, we proved that vague sets and interval valued vague sets are both special cases of novel interval valued vague sets. And then the concept of interval-valued vague sets with parameters (IVVSP) was proposed. Based on membership and non-membership, this paper focused on the construction of IVVSP. Finally, a pattern recognition example and a medical diagnosis decision-making example were given to demonstrate the application of IVVSP. The results of simulation show that the method of IVVSP is more comprehensive and flexible than that of the traditional vague sets.

Keywords Vague sets, Interval-valued vague sets, Novel interval-valued vague sets, Interval-valued vague sets with parameters, Pattern recognition

1 引言

1993 年 Gau 和 Buehrer 提出 vague 集 (vague sets, VS)^[1]。1996 年 Bustine 和 Burillo 指出 vague 集等同于直觉模糊集,二者本质上是一样的,仅仅在表达方式上有细微差别^[2]。接着,李凡、Chen、Hong 等将 vague 集理论运用于决策分析与模式识别领域^[8-11]。2001 年马志峰、邢汉承等在 vague 集基础上,结合区间值模糊集,提出了区间值 vague 集 (Interval-valued vague sets, IVVS)^[3]。随后,部分学者开始研究区间值 vague 集。万树平等将区间值 vague 集运用于模式识别^[12,13], Amit Kumar 等将其运用于模糊系统的可靠性分析^[14],周珍、吴祈宗、王卫平、张倩生等将其运用于决策分析^[15-19]。

事实上,基于现有文献中的定义,区间值 vague 集一直被认为是 vague 集的自然推广^[3,4,12-19]。从理论角度,任何区间

值模糊集也理应是相应模糊集合的扩展。但是,本文基于现有定义,证明了区间值 vague 集不是 vague 集的扩展,vague 集也不是区间值 vague 集的特例,因此有必要重新定义区间值 vague 集。

本文提出了新的区间值 vague 集定义,从理论上证明了现有文献中的 vague 集、区间值 vague 集都是新区间值 vague 集的特例。接着,基于新区间值 vague 集的定义,提出了带参数的区间值 vague 集定义及构造方法,并重点分析了从隶属度和非隶属度出发,构造双参数、三参数的区间值 vague 集模型。最后,利用传统的距离公式,将所构造的三参数区间值 vague 集用于模式识别和医疗诊断。仿真实验结果表明,本文的方法比传统的 vague 集和区间值 vague 集方法更全面且更灵活。本文的方法可以为当前区间值 vague 的实际应用研究领域,特别是需要构建隶属度的应用领域提供有价值的结论,对于将 vague 推理推广到区间值 vague 推理也有一定的

到稿日期:2010-12-11 返修日期:2011-04-15 本文受国家自然科学基金资助项目(61070061,70801020),广东外语外贸大学校级项目(GW20052013),广东省自然科学基金资助项目(9151026005000002),广东省高等院校学科建设专项资金(396-GK100018)资助。

张振华(1972-),男,博士生,讲师,CCF 会员,主要研究方向为模糊信息处理与模式识别, E-mail: zhangzhenhua@mail.gdufs.edu.cn; 杨静宇(1941-),男,教授,博士生导师,主要研究方向为模式识别、计算机视觉与图像处理、智能系统与智能机器人、数据融合; 叶有培(1944-),男,教授,博士生导师,主要研究方向为模糊信息处理与机器智能、密码学与加密算法; 张倩生(1975-),男,博士,副教授,CCF 会员,主要研究方向为模糊推理与决策。

理论意义^[20]。

2 区间值 vague 集的新定义

定义 1 U 是论域, $\forall x \in U$, 称区间 $[t_A(x), 1-f_A(x)]$ 为 $x \in A$ 的 vague 集隶属度区间, 即 vague 集 A 在点 x 的 vague 值。其中一对隶属度函数 $t_A(x), f_A(x)$ 满足: $t_A(x): U \rightarrow [0, 1], x \rightarrow t_A(x), f_A(x): U \rightarrow [0, 1], x \rightarrow f_A(x)$ 且 $t_A(x) + f_A(x) \in [0, 1]$ 。 $t_A(x)$ 称为 A 的真隶属函数, 表示支持 $x \in A$ 的隶属度下界; $f_A(x)$ 称为 A 的假隶属函数, 表示反对 $x \in A$ 的隶属度下界, 相应地, $\pi_A(x) = 1 - t_A(x) - f_A(x)$ 表示 $x \in A$ 的犹豫度, 也称为不确定程度。

定义 2 对于论域 U , 称满足下述条件的集合 $A = \langle x, M_A(x), N_A(x) \mid x \in U \rangle$ 为区间值 vague 集:

(1) $M_A(x) = [t_A^-(x), t_A^+(x)], N_A(x) = [f_A^-(x), f_A^+(x)], H_A(x) = [\pi_A^-(x), \pi_A^+(x)]$, 其中 $M_A(x), N_A(x), H_A(x)$ 分别表示真隶属度、假隶属度、不确定度范围;

(2) $M_A(x) \subseteq [0, 1], N_A(x) \subseteq [0, 1], H_A(x) \subseteq [0, 1]$;

(3) $\pi_A^-(x) = 1 - t_A^+(x) - f_A^+(x), \pi_A^+(x) = 1 - t_A^-(x) - f_A^-(x)$;

(4) $\text{Sup } M_A(x) + \text{Sup } N_A(x) = t_A^+(x) + f_A^+(x) \leq 1$ 。

这是马志峰、邢汉承等提出的区间值 vague 集, 也是现有文献中公认的区间值 vague 集定义^[3,4,12-19]。

命题 1 vague 集不是区间值 vague 集的特例。

证明: 对于任意 vague 集 A , 若 $H_A(x) = [0, 0]$, 则犹豫度为 0, 此时 $\pi_A(x) = 0$, 则 $\pi_A^-(x) = \pi_A^+(x) = 0, M_A(x) = [t_A^-(x), t_A^+(x)] = [\mu_A(x), \mu_A(x)], N_A(x) = [f_A^-(x), f_A^+(x)] = [1 - \mu_A(x), 1 - \mu_A(x)]$, 区间 $[a, a]$ 表示单元素集 $\{a\}$, 此模型就是 Zadeh 模糊集。

当 $H_A(x) \neq [0, 0]$ 时, $\pi_A^+(x) > 0$, 依据定义 1, 令 $t_A^-(x) = t_A(x), t_A^+(x) = 1 - f_A(x) = t_A(x) + \pi_A(x), f_A^-(x) = f_A(x), f_A^+(x) = 1 - t_A(x) = f_A(x) + \pi_A(x), \pi_A^+(x) = \pi_A(x)$, 且 $\pi_A(x) > 0$, 因此有:

$\text{Sup } M_A(x) + \text{Sup } N_A(x) = t_A^+(x) + f_A^+(x)$

$= t_A(x) + \pi_A(x) + f_A(x) + \pi_A(x) = 1 + \pi_A(x) > 1$

显然, 与定义 2(4) $\text{Sup } M_A(x) + \text{Sup } N_A(x) \leq 1$ 相反。即当 $\pi_A^+(x) > 0$ 时, 定义 1 与定义 2(4) 矛盾。

综上所述, 当 $H_A(x) = [0, 0]$ 时, vague 集与区间值 vague 集等价, 也等同于模糊集。只要 $H_A(x) \neq [0, 0]$, 若 A 是 vague 集, 则一定不是区间值 vague 集, 反之若 A 是区间值 vague 集, 则 A 一定不是 vague 集。因此 vague 集与区间值 vague 集互不包含。

定义 3 对于论域 U , 称满足下述条件的集合 $A = \langle x, M_A(x), N_A(x) \mid x \in U \rangle$ 为新区间值 vague 集:

(1) $M_A(x) = [t_A^-(x), t_A^+(x)], N_A(x) = [f_A^-(x), f_A^+(x)], H_A(x) = [\pi_A^-(x), \pi_A^+(x)]$, 其中 $M_A(x), N_A(x), H_A(x)$ 分别表示真隶属度、假隶属度、不确定度范围。

(2) $M_A(x) \subseteq [0, 1], N_A(x) \subseteq [0, 1], P_A(x) \subseteq [0, 1]$;

(3) $\pi_A^-(x) = \max\{0, 1 - t_A^+(x) - f_A^+(x)\}, \pi_A^+(x) = 1 - t_A^-(x) - f_A^-(x)$;

(4) $\text{Inf } M_A(x) + \text{Sup } N_A(x) = t_A^-(x) + f_A^+(x) \leq 1, \text{Sup } M_A(x) + \text{Inf } N_A(x) = t_A^+(x) + f_A^-(x) \leq 1$;

(5) $\pi_A^+(x) - \pi_A^-(x) \leq t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x)$ 。

根据定义, 下面将证明此类集合就是 vague 集在区间值上的推广, 因此以下仍统称之为新区间值 vague 集, 即 NIV-VS。

命题 2 vague 集和区间值 vague 集都是新区间值 vague 集的特例。

证明: 由 2.1 节定义, 对于任意 vague 集 A , 因 $t_A(x) + f_A(x) + \pi_A(x) = 1$, 则有:

$M_A(x) = [t_A^-(x), t_A^+(x)] = [t_A(x), 1 - f_A(x)]$

$= [t_A(x), t_A(x) + \pi_A(x)] \subseteq [0, 1]$

$N_A(x) = [f_A^-(x), f_A^+(x)] = [f_A(x), 1 - t_A(x)]$

$= [f_A(x), f_A(x) + \pi_A(x)] \subseteq [0, 1]$

$\pi_A^-(x) = \max\{0, 1 - t_A^+(x) - f_A^+(x)\}$

$= \max\{0, 1 - t_A(x) - f_A(x) - 2\pi_A(x)\}$

$= \max\{0, -\pi_A(x)\} = 0$

$\pi_A^+(x) = 1 - t_A^-(x) - f_A^-(x) = \pi_A(x)$

$H_A(x) = [\pi_A^-(x), \pi_A^+(x)] = [0, \pi_A(x)] \subseteq [0, 1]$

满足定义 3(1)(2)(3)。

$\text{Inf } M_A(x) + \text{Sup } N_A(x) = t_A^-(x) + f_A^+(x) = 1$

$\text{Sup } M_A(x) + \text{Inf } N_A(x) = t_A^+(x) + f_A^-(x) = 1$

满足定义 3(4)。

$t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x) = 2(1 - t_A(x) - f_A(x)) = 2\pi_A(x)$

$\pi_A^+(x) - \pi_A^-(x) = \pi_A(x) \leq t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x)$

满足定义 3(5), 故 A 满足定义 3 所有条件。

对于任意区间值 vague 集 A , 证明 A 也满足定义 3。由于定义 2 与定义 3 的(1)(2)定义完全相同, 只需证(3)和(4)。

由于定义 2(3)与定义 3(3)中 $\pi_A^+(x)$ 完全相同, 因此只需要证明定义 2 所定义的 $\pi_A^-(x)$ 也满足定义 3 即可。由定义 2(4), 有:

$\text{Sup } M_A(x) + \text{Sup } N_A(x) = t_A^+(x) + f_A^+(x) \leq 1$

故 $1 - t_A^+(x) - f_A^+(x) \geq 0$, 因此显然有 $\max\{0, 1 - t_A^+(x) - f_A^+(x)\} = 1 - t_A^+(x) - f_A^+(x)$, 则 $\pi_A^-(x) = \max\{0, 1 - t_A^+(x) - f_A^+(x)\} = 1 - t_A^+(x) - f_A^+(x) \geq 0$, 因此, 满足定义 2 的 $\pi_A^-(x)$ 同样满足定义 3(3)。

又由 $\text{Sup } M_A(x) + \text{Sup } N_A(x) = t_A^+(x) + f_A^+(x) \leq 1$, 显然有:

$\text{Sup } M_A(x) + \text{Inf } N_A(x) = t_A^+(x) + f_A^-(x) \leq t_A^+(x) + f_A^+(x) \leq 1$,

$\text{Inf } M_A(x) + \text{Sup } N_A(x) = t_A^-(x) + f_A^+(x) \leq t_A^-(x) + f_A^+(x) \leq 1$

因此, 若满足定义 2(4), 则满足定义 3(4)。

根据定义 2(3), 两式相减易得满足定义 3(5):

$\pi_A^+(x) - \pi_A^-(x) = t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x)$

综上所述, 任意区间值 vague 集 A 满足新区间值 vague 集所有条件。

命题 3(必要条件) 区间值 vague 集满足:

$t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x) = \pi_A^+(x) - \pi_A^-(x)$

而新区间值 vague 集满足:

$\max\{t_A^+(x) - t_A^-(x), f_A^+(x) - f_A^-(x)\} \leq \pi_A^+(x) - \pi_A^-(x)$

且 $\pi_A^+(x) - \pi_A^-(x) \leq t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x)$ 。

证明: 根据定义 2(3), 两式相减, 显然可证。

根据定义 3(3), 若 $\pi_A^-(x) = 1 - t_A^+(x) - f_A^+(x) \geq 0$, 同上, 有 $\pi_A^+(x) - \pi_A^-(x) = t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x)$, 也满足 $\max\{t_A^+(x) - t_A^-(x), f_A^+(x) - f_A^-(x)\} \leq \pi_A^+(x) - \pi_A^-(x)$; 若 $1 - t_A^+(x) - f_A^+(x) < 0$, $\pi_A^-(x) = \max\{0, 1 - t_A^+(x) - f_A^+(x)\} = 0$, 又依据定义 3(4), $t_A^+(x) + f_A^-(x) \leq 1$, $\rightarrow t_A^+(x) \leq 1 - f_A^-(x)$, $\rightarrow t_A^+(x) - t_A^-(x) \leq 1 - f_A^-(x) - t_A^-(x) = \pi_A^+(x)$, $\rightarrow t_A^+(x) - t_A^-(x) \leq \pi_A^+(x) = \pi_A^+(x) - \pi_A^-(x)$; 同理可证 $f_A^+(x) - f_A^-(x) \leq \pi_A^+(x) - \pi_A^-(x)$ 。

而另一条就是定义 3(5), 证毕。

命题 2 与命题 3 说明对于新区间值 vague 集, $\pi_A^-(x) = 0$, 且 $t_A^+(x) - t_A^-(x) = f_A^+(x) - f_A^-(x) = \pi_A^+(x) - \pi_A^-(x)$, 就是 vague 集; 若 $t_A^+(x) - t_A^-(x) + f_A^+(x) - f_A^-(x) = \pi_A^+(x) - \pi_A^-(x)$, 就是区间值 vague 集。因此, 区间值 vague 集与 vague 集都是新区间值 vague 集的两种极端情形, 且其中若 $\pi_A^+(x) = \pi_A^-(x) = 0$, 就是 Zadeh 模糊集。

3 带参数区间值 vague 集的构造

定义 4 对于论域 U , 称满足下述条件的集合 $A = \{\langle x, M_A(x), N_A(x) \mid x \in U \rangle\}$ 为双参数区间值 vague 集:

(1) $M_A(x) = [t_A^-(x), t_A^+(x)]$, $N_A(x) = [f_A^-(x), f_A^+(x)]$, $H_A(x) = [\pi_A^-(x), \pi_A^+(x)]$; $M_A(x) \subseteq [0, 1]$, $N_A(x) \subseteq [0, 1]$, $H_A(x) \subseteq [0, 1]$;

(2) $t_A^+(x) = t_A^-(x) + \alpha_A(x)$, $f_A^+(x) = f_A^-(x) + \beta_A(x)$, $\pi_A^+(x) = 1 - t_A^-(x) - f_A^-(x)$, $\pi_A^-(x) = 1 - t_A^+(x) - f_A^+(x)$ 。

其中 $0 \leq \alpha_A(x) \leq \pi_A^+(x) - \pi_A^-(x)$, $0 \leq \beta_A(x) \leq \pi_A^+(x) - \pi_A^-(x)$, $\pi_A^+(x) - \pi_A^-(x) \leq \alpha_A(x) + \beta_A(x) \leq 2(\pi_A^+(x) - \pi_A^-(x))$ 。

由定义 2, 设将所有样本数据分为三部分, 第一部分 $t_A^-(x)$ 为某事件 A 的坚定支持派, 第二部分 $f_A^-(x)$ 表示事件 A 的坚定反对派, $\pi_A^+(x)$ 表示全部弃权派, $\pi_A^-(x)$ 是其中坚决弃权的弃权派, $\pi_A^+(x) - \pi_A^-(x)$ 为其中可转化的弃权派, 只有这部分可能变成支持派或反对派之一。若有 $\alpha_A(x)$ 的样本可能支持事件 A , $\beta_A(x)$ 的样本可能反对事件 A , 则可以得到基于定义 4 的双参数区间值 vague 集。

若令其中转化为支持的比例为 $\lambda_{A1}(x)$, 反对的比例为 $\lambda_{A2}(x)$, 则有 $\alpha_A(x) = \lambda_{A1}(x)(\pi_A^+(x) - \pi_A^-(x))$, $\beta_A(x) = \lambda_{A2}(x)(\pi_A^+(x) - \pi_A^-(x))$, 即为双参数区间值模型的另一种表达形式。又由定义 4, 若满足 $\pi_A^-(x) = (1 - \lambda_{A0}(x))\pi_A^+(x)$, $\pi_A^+(x) - \pi_A^-(x) = \lambda_{A0}(x)\pi_A^+(x)$, 则可得到下述三参数区间值 vague 集。

定义 5 对于论域 U , 满足下述条件的集合 $A = \{\langle x, M_A(x), N_A(x) \rangle \mid x \in U \rangle\}$ 是三参数区间值 vague 集:

$H_A(x) = [\pi_A^-(x), \pi_A^+(x)] = [\lambda_{A0}(x)\pi_A^+(x), \pi_A^+(x)]$

$M_A(x) = [t_A^-(x), t_A^+(x) + \lambda_{A0}(x)\lambda_{A1}(x)\pi_A^+(x)]$

$N_A(x) = [f_A^-(x), f_A^-(x) + \lambda_{A0}(x)\lambda_{A2}(x)\pi_A^+(x)]$

且 $0 \leq \lambda_{Ai}(x) \leq 1, i=0, 1, 2, 1 \leq \lambda_{A1}(x) + \lambda_{A2}(x) \leq 2$; 其中 $M_A(x)$ 、 $N_A(x)$ 、 $H_A(x)$ 分别表示真隶属度、假隶属度、不确定度范围。

下面对上述双参数区间值 vague 集进行具体分类: 若 $\lambda_{A0}(x) = \lambda_{A1}(x) = \lambda_{A2}(x) = 1$, 三参数区间值 vague 集是 vague

集; 若 $\lambda_{A1}(x) + \lambda_{A2}(x) = 1$, 三参数区间值 vague 集等价于传统的区间值 vague 集; 若 $\lambda_{Ai}(x) = \lambda_i$, 上述模型为固定参数区间值 vague 集; 当 $\lambda_{Ai}(x)$ 随着 x 变化而变化时, 模型为变参数区间值 vague 集。

4 距离公式

以海明距离和欧氏距离为例进行距离分析,

$$d_{A,B} = \sqrt[p]{\frac{\sum_{i=1}^n (|t_A^-(x_i) - t_B^-(x_i)|^p + |f_A^-(x_i) - f_B^-(x_i)|^p + |t_A^+(x_i) - t_B^+(x_i)|^p + |f_A^+(x_i) - f_B^+(x_i)|^p)}{4n}}$$

不妨设 $d_{A,B} = \sqrt[p]{\frac{\sum_{i=1}^n d_{A,B}(x_i)}{4n}}$; 若 $p=1$, 是海明距离, $p=2$, 即欧氏距离。显然, 距离仅与下式有关: 当 $p=2$ 时,

$$\begin{aligned} d_{A,B}(x_i) &= (t_A^-(x_i) - t_B^-(x_i))^2 + (f_A^-(x_i) - f_B^-(x_i))^2 + \\ &\quad (t_A^+(x_i) - t_B^+(x_i))^2 + (f_A^+(x_i) - f_B^+(x_i))^2 \\ &= 2[(t_A^-(x_i) - t_B^-(x_i))^2 + (f_A^-(x_i) - f_B^-(x_i))^2 + \\ &\quad (\pi_A^+(x_i) - \pi_B^+(x_i))^2 + (\lambda_{A0}\lambda_{A1})^2 + \\ &\quad (\pi_A^+(x_i) - \pi_B^+(x_i))^2 + (\lambda_{A0}\lambda_{A2})^2 + 2(t_A^-(x_i) - \\ &\quad t_B^-(x_i))(\pi_A^+(x_i) - \pi_B^+(x_i))\lambda_{A0}\lambda_{A1} + 2(f_A^-(x_i) - \\ &\quad f_B^-(x_i))(\pi_A^+(x_i) - \pi_B^+(x_i))\lambda_{A0}\lambda_{A2} \end{aligned}$$

当 $p=1$ 时,

$$\begin{aligned} d_{A,B}(x_i) &= |t_A^-(x_i) - t_B^-(x_i)| + |f_A^-(x_i) - f_B^-(x_i)| + \\ &\quad |t_A^+(x_i) - t_B^+(x_i)| + |f_A^+(x_i) - f_B^+(x_i)| \\ &= |t_A^-(x_i) - t_B^-(x_i)| + |f_A^-(x_i) - f_B^-(x_i)| + \\ &\quad |(t_A^-(x_i) - t_B^-(x_i)) + \lambda_{A0}\lambda_{A1}(\pi_A^+(x_i) - \pi_B^+(x_i))| + \\ &\quad |(f_A^-(x_i) - f_B^-(x_i)) + \lambda_{A0}\lambda_{A2}(\pi_A^+(x_i) - \pi_B^+(x_i))| \end{aligned}$$

进行模式识别, 只需要比较 $\sum_{i=1}^n d_{A,B}(x_i)$ 的大小即可。

5 在医疗诊断和模式识别中的应用

例 1 基于文献[5, 6]的医疗诊断案例, 表 1 是基于 vague 集的诊断结果症状描述, 表 2 是基于 vague 集的待诊病人症状描述。其中疾病诊断集(决策属性) {Viral fever, Malaria, Typhoid, Stomach problem, Chest problem}、症状集(条件属性) {temperature, headache, stomach pain, cough, chest pain}, 上述为已有模式。待诊断病人集(待识别模式) {AI, Bob, Joe, Ted} 与已有模式条件属性相同。表中每一格两个数据, 前者表示隶属度, 后者为非隶属度。例如对于 Viral fever 病症患者, 症状 temperature, 隶属度 $t_A^-(x) = 0.4$, 假隶属度 $f_A^-(x) = 0$, $\pi_A^+(x) = 1 - t_A^-(x) - f_A^-(x) = 0.6$ 。

表 1 诊断结果症状描述_vague 集隶属度与非隶属度

| | temperature | headache | stomachpain | cough | chestpain |
|----------------|-------------|----------|-------------|---------|-----------|
| Viralfever | 0.4,0 | 0.3,0.5 | 0.1,0.7 | 0.4,0.3 | 0.1,0.7 |
| Malaria | 0.7,0 | 0.2,0.6 | 0,0.9 | 0.7,0 | 0.1,0.8 |
| Typhoid | 0.3,0.3 | 0.6,0.1 | 0.2,0.7 | 0.2,0.6 | 0.1,0.9 |
| Stomachproblem | 0.1,0.7 | 0.2,0.4 | 0.8,0 | 0.2,0.7 | 0.2,0.7 |
| Chestproblem | 0.1,0.8 | 0,0.8 | 0.2,0.8 | 0.2,0.8 | 0.8,0.1 |

表 2 待诊病人症状描述_vague 集隶属度与非隶属度

| | temperature | headache | stomachpain | cough | chestpain |
|-----|-------------|----------|-------------|---------|-----------|
| AI | 0.8,0.1 | 0.6,0.1 | 0.2,0.8 | 0.6,0.1 | 0.1,0.6 |
| Bob | 0,0.8 | 0.4,0.4 | 0.6,0.1 | 0.1,0.7 | 0.1,0.8 |
| Joe | 0.8,0.1 | 0.8,0.1 | 0,0.6 | 0.2,0.7 | 0,0.5 |
| Ted | 0.6,0.1 | 0.5,0.4 | 0.3,0.4 | 0.7,0.2 | 0.3,0.4 |

根据文献[5,6],依据海明距离和欧氏距离计算,可得基于 vague 集的模式识别结果,如表 3 所列。

表 3 vague 集诊断结果

| | AI | Bob | Joe | Ted |
|------|-------------|----------------|---------|-------------|
| 海明距离 | Malaria | Stomachproblem | Typhoid | Viral fever |
| 欧氏距离 | Viral fever | Stomachproblem | Typhoid | Viral fever |

根据定义 5 所定义三参数区间值 vague 集,利用欧氏距离计算可得表 4。

表 4(a) AI 的双参数区间值 vague 集诊断结果

| coefficient | Constant | $\lambda_{A0}\lambda_{A1}$ | $\lambda_{A0}\lambda_{A2}$ | $\lambda_{A0}^2(\lambda_{A1}^2+\lambda_{A2}^2)$ |
|----------------|-------------|----------------------------|----------------------------|---|
| Viral fever | 1.06 | -0.38 | -0.24 | 0.31 |
| Malaria | 1.08 | 0 | -0.2 | 0.1 |
| Typhoid | 1.6 | -0.22 | -0.18 | 0.2 |
| Stomachproblem | 5.28 | 0.14 | -0.42 | 0.14 |
| Chestproblem | 6.44 | 0.08 | -0.36 | 0.14 |

如表 4(a)所列,例如对于 AI 与 Viral fever 之间距离, $\sum_{i=1}^n d_{AI, Viral\ fever}(x_i) = 1.06 - 0.38\lambda_{A0}\lambda_{A1} - 0.24\lambda_{A0}\lambda_{A2} + 0.31\lambda_{A0}^2(\lambda_{A1}^2 + \lambda_{A2}^2)$

$$d_{AI, Viral\ fever} = \sqrt{\sum_{i=1}^n d_{AI, Viral\ fever}(x_i) / 4n}, \text{ 其余同理可得。}$$

表 4(b) Bob 的双参数区间值 vague 集诊断结果

| coefficient | Constant | $\lambda_{A0}\lambda_{A1}$ | $\lambda_{A0}\lambda_{A2}$ | $\lambda_{A0}^2(\lambda_{A1}^2+\lambda_{A2}^2)$ |
|----------------|-------------|----------------------------|----------------------------|---|
| Viral fever | 3.38 | 0.48 | -0.86 | 0.19 |
| Malaria | 6.12 | 0.5 | -0.62 | 0.06 |
| Typhoid | 2.04 | 0.32 | -0.52 | 0.1 |
| Stomachproblem | 0.28 | -0.14 | 0.02 | 0.06 |
| Chestproblem | 3.96 | 0.18 | -0.46 | 0.14 |

表 4(c) Joe 的双参数区间值 vague 集诊断结果

| coefficient | Constant | $\lambda_{A0}\lambda_{A1}$ | $\lambda_{A0}\lambda_{A2}$ | $\lambda_{A0}^2(\lambda_{A1}^2+\lambda_{A2}^2)$ |
|----------------|-------------|----------------------------|----------------------------|---|
| Viral fever | 1.7 | -0.52 | -0.34 | 0.43 |
| Malaria | 3.12 | -0.04 | -0.64 | 0.34 |
| Typhoid | 1.12 | -0.6 | -0.36 | 0.48 |
| Stomachproblem | 4.76 | -0.98 | 0.38 | 0.3 |
| Chestproblem | 6 | -0.96 | 0.28 | 0.34 |

表 4(d) Ted 的双参数区间值 vague 集诊断结果

| coefficient | Constant | $\lambda_{A0}\lambda_{A1}$ | $\lambda_{A0}\lambda_{A2}$ | $\lambda_{A0}^2(\lambda_{A1}^2+\lambda_{A2}^2)$ |
|----------------|-------------|----------------------------|----------------------------|---|
| Viral fever | 0.92 | -0.2 | -0.12 | 0.16 |
| Malaria | 1.46 | 0.14 | -0.4 | 0.13 |
| Typhoid | 2.06 | 0.04 | -0.42 | 0.19 |
| Stomachproblem | 3.42 | -0.14 | -0.16 | 0.15 |
| Chestproblem | 4.54 | 0.06 | -0.44 | 0.19 |

显然, $\sum_{i=1}^n d_{AI, Viral\ fever}(x_i)$ 越小,则 $d_{AI, Viral\ fever}$ 越小。因此,最小的 $\sum_{i=1}^n d_{AI, Viral\ fever}(x_i)$ 所对应的模式即为识别的结果。根据上述表 4(a)一表 4(d),因 $0 \leq \lambda_{Ai} \leq 1, i=0,1$,经计算得表 5。

表 5 双参数区间值 vague 集诊断结果

| | AI | Bob | Joe | Ted |
|------|---------------------|----------------|---------|-------------|
| 欧氏距离 | Malaria 条件 1 | Stomachproblem | Typhoid | Viral fever |
| 欧氏距离 | Viral fever 条件 2 | Stomachproblem | Typhoid | Viral fever |

如表 5 所列,除 AI 外,其余识别结果与表 3 一致。当 $\sum_{i=1}^n d_{AI, Viral\ fever}(x_i) > \sum_{i=1}^n d_{AI, Malaria}(x_i)$ 时, AI 识别为 Malaria; 即当

$0.02 + 0.04\lambda_{A0}\lambda_{A2} + 0.38\lambda_{A0}\lambda_{A1} - 0.21\lambda_{A0}^2(\lambda_{A1}^2 + \lambda_{A2}^2) < 0$ 时, AI 识别为 Malaria,此即为表 5 中的条件 1;反之,即为条件 2,此时 AI 识别为 Viral fever。

文献[5,6]采用隶属度与非隶属度计算,即得 vague 集诊断结果,当采用海明距离时, AI 诊断为 Malaria,采用欧氏距离, AI 诊断为 Viral fever。本文基于三参数 vague 集欧氏距离,结果为若满足条件 1, AI 诊断为 Malaria; 满足条件 2, 诊断为 Viral fever。若采用海明距离,结果更丰富,其中 AI、Bob、Joe 的识别与欧氏距离类似,而 Ted 在一定条件下分别识别为 Viral fever 与 malaria,在此不赘述。上述实验结果表明,本文所提出的带参数区间值 vague 集方法识别结果与文献[5,6]有细微差异。传统的 vague 集虽然方法简单,但缺陷在于利用传统距离公式识别时,识别结果固定,若对样本数据识别效果不好,则方法失效,因此灵活性不够。本文将变参数区间值 vague 集方法用于模式识别,通过参数变化调节识别结果,因此它比传统的 vague 集方法更为灵活,且运算比较简便。新方法考虑了对弃权派的分割,因此比传统 vague 集方法更全面。

例 2 文献[7]中, Hung 和 Yang 给出 15 个专家对某工程决策方案的 vague 集评价, Hung 和 Yang 用 SCM 聚类算法得到所有专家的 3 个模式,如表 6 所列。

表 6 专家模式 1

| mode | $t_{B_i}(x)$ | $f_{B_i}(x)$ | $\pi_{B_i}(x)$ |
|----------------|--------------|--------------|----------------|
| B ₁ | 0.875 | 0.119 | 0.006 |
| B ₂ | 0.089 | 0.813 | 0.098 |
| B ₃ | 0.444 | 0.499 | 0.057 |

设某专家 A 评分为 $t_A(x) = 0.6, f_A(x) = 0.38$, 根据下述欧氏距离计算可得距离公式计算结果,如表 7 所列。显然,对于任意 $0 \leq \lambda_{Ai} \leq 1, i=0,1$, 专家 A 都识别成 B₃。上述结论与文献[7]相同。根据本文提出的方法,还可以根据参数的取值范围计算出模式 A 与模式 B_i 距离的上限、下限区间范围。此例也说明,双参数区间值 vague 集方法在实际应用中比传统的 vague 集方法更全面、更灵活,适用范围更广。

表 7 专家 A 双参数区间值 vague 集识别结果

| coefficient | Constant | $\lambda_{A0}\lambda_{A1}$ | $\lambda_{A0}\lambda_{A2}$ | $\lambda_{A0}^2(\lambda_{A1}^2+\lambda_{A2}^2)$ |
|----------------|--------------|----------------------------|----------------------------|---|
| B ₁ | 0.288 | -0.00385 | 0.00731 | 0.000196 |
| B ₂ | 0.896 | -0.04037 | 0.06826 | 0.006241 |
| B ₃ | 0.077 | -0.00577 | 0.00881 | 0.001369 |

结束语 综上所述,本文重新定义了区间值 vague 集,并证明了新集合是传统 vague 集与区间值 vague 集的共同推广。基于新定义,提出了一系列带参数区间值 vague 集的构造方法,重点分析了多参数区间值 vague 集。并利用传统的距离公式,将所构造的三参数区间值 vague 集用于模式识别与医疗诊断。实验结果和文献[5-7]的基本一致,但方法比文献[5-7]的更灵活。新方法中的距离公式不仅考虑了隶属度与非隶属度,也考虑了对犹豫度的分解,因此比传统 vague 集方法更全面。

在本文基础上可进行一定的后续研究。本文重点阐述的区间值 vague 集模型是在已知 vague 集隶属度和非隶属度基础上进行的。事实上,研究人员可以根据自己领域的相关知识构造其他合适的带参数区间值 vague 集模型,并通过参数变化观察贴近度的变化规律及其对模式识别结果的影响。

参 考 文 献

- [1] Gau W L, Buehrer D J. Vague sets [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1993, 23: 610-614
- [2] Bustince H, Burillo P. Vague sets are intuitionistic fuzzy sets [J]. Fuzzy Sets and Systems, 1994, 61: 137-142
- [3] 马志峰, 邢汉承, 郑晓妹. 区间值 Vague 决策系统及其规则提取方法 [J]. 电子学报, 2001, 5: 585-589
- [4] 周晓光, 谭春桥, 张强. 基于 Vague 集的决策理论与方法 [M]. 北京: 科学出版社, 2009
- [5] Szmjdt E, Kacprzyk J. Distances between intuitionistic fuzzy sets and their applications in reasoning [J]. Studies in Computational Intelligence, 2005, 2: 101-116
- [6] Szmjdt E, Kacprzyk J. An intuitionistic fuzzy set based approach to intelligent data analysis; an application to medical diagnosis [M]// Abraham A, Jain L, Kacprzyk J, eds. Recent advances in intelligent paradigms and applications. Studies in Fuzziness and Soft Computing. Germany: Physica-Verlag GmbH Heidelberg, 2003: 57-70
- [7] Hung W L, Yang M S. Similarity measure of intuitionistic fuzzy sets based on L_p metric [J]. International Journal of Approximate Reasoning, 2007, 46(1): 120-136
- [8] Chen S M. Similarity measures between vague sets and between elements [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1997, 27: 153-157
- [9] Hong D H, Kim C. A note on Similarity measures between vague sets and between elements [J]. Information Sciences, 1999, 115: 83-96
- [10] 李凡, 徐章艳. Vague 集之间的相似度量 [J]. 软件学报, 2001, 12(6): 922-927
- [11] 朱振国, 王国胤. Vague 集之间的相似度量 [J]. 计算机科学, 2008, 35(9): 220-225
- [12] 万树平. 区间值 Vague 集在多传感器信息融合中的应用 [J]. 计算机应用研究, 2009, 10: 3818-3820
- [13] Wan Shu-ping. Applying interval-value vague set for multi-sensor target recognition [J]. International Journal of Innovative Computing, Information and Control, 2011, 7(2): 955-963
- [14] Kumar A, Yadav S P, Kumar S. Generalization of fuzzy sets and its use for analyzing the fuzzy system reliability [C]// IEEE International Conference on Industrial Technology, ICIT, 2006. Mumbai, India, 2006: 2675-2680
- [15] 周珍, 吴祈宗, 兰晓婷. 基于记分函数的区间值 Vague 集的相似度量 [J]. 小型微型计算机系统, 2006, 6: 1007-1012
- [16] 周珍, 吴祈宗, 等. 基于距离的区间值 Vague 集的多准则决策方法 [J]. 北京理工大学学报, 2006, 8: 693-696
- [17] Wang Wei-ping, Wu Qi-zong, Shi Yu. Similarity measures between interval value vague sets [C]// IEEE international conference on Chinese control and decision conference, CCDC, Yantai, China, 2008: 2300-2305
- [18] Liu fang, Zheng Guo-lei, Xu Da-ju, et al. Decision-making of mechanical design scheme based on the distance between interval value vague sets [C]// IEEE international conference on Chinese control and decision conference, CCDC. Xuzhou, China, 2010: 1157-1160
- [19] 张倩生, 蒋盛益. 基于区间值 Vague 集的多属性模糊决策方法 [J]. 小型微型计算机系统, 2010, 5: 942-946
- [20] Zhang Qian-sheng, Yao Hai-xiang, Zhang zhen-hua. An interval-valued fuzzy reasoning approach based on weighted similarity measure [J]. Advanced Materials Research, 2011, 143-144: 161-165
- [21] Zhang Qian-sheng, Yao Hai-xiang, Zhang zhen-hua. Some similarity measures of interval-Valued intuitionistic fuzzy sets and application to pattern recognition [J]. International Journal of Applied Mechanics and Materials, 2011, 44-47: 3888-3892

(上接第 193 页)

- [9] 张飞, 陈萍, 张丽. P-集合的 P-分离与应用 [J]. 山东大学学报: 理学版, 2010, 45(3): 18-22
- [10] 史开泉, 张丽. 内 P-集合与数据外-恢复 [J]. 山东大学学报: 理学版, 2009, 44(4): 8-14
- [11] Zhang Li, Cui Yu-quan. Outer P-sets and data internal recovery [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 189-199
- [12] Zhang Li, Xiu Ming, Shi Kai-quan. P-sets and applications of internal- outer data circle [J]. Quantitative Logic and Soft Computing, 2010, 1(2): 581-592
- [13] Li Yu-ying, Zhang Li, Shi Kai-quan. Generation and recovery of compressed data and redundant data [J]. Quantitative Logic and Soft Computing, 2010, 1(2): 661-672
- [14] Xiu Ming, Shi Kai-quan, Zhang Li. P-sets and \bar{F} -data selection-discovery [J]. Quantitative Logic and Soft Computing, 2010, 1(2): 791-800
- [15] 周玉华, 张冠宇, 张丽. 内-外数据圆与动态数据-恢复 [J]. 山东大学学报: 理学版, 2010, 45(8): 21-26
- [16] Lin Hong-kang, Li Yu-ying. P-sets and its P-separation theorems [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 209-215
- [17] 周玉华, 张冠宇, 史开泉. P-集合与双信息规律生成 [J]. 数学的实践与认识, 2010, 40(13): 71-80
- [18] Wang Yang, Geng Hong-qin, Shi Kai-quan. The mining of dynamic information based on P-sets and its application [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 234-240
- [19] Zhang Guan-yu, Li En-zhong. Information gene and identification of its information Knock-out/Knock-in [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 308-315
- [20] Huang Shun-liang, Wang Wei, Geng Dian-you. P-sets and its internal P-memory characteristics [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 216-222
- [21] Liu Jiqin. P-probabilities and its application [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 200-222
- [22] 张冠宇, 周厚勇, 史开泉. P-集合与双 P-数据恢复-辨识 [J]. 系统工程与电子技术, 2010, 32(9): 1919-1924
- [23] 于秀清. P-集合的识别与筛选 [J]. 山东大学学报: 理学版, 2010, 45(1): 94-98
- [24] 汤积华, 陈保会, 史开泉. P-集合与 (\bar{F}, F) -数据生成-辨识 [J]. 山东大学学报: 理学版, 2009, 44(11): 83-92
- [25] 李豫颖, 谢维奇, 史开泉. \bar{F} -残缺数据的辨识与恢复 [J]. 山东大学学报: 理学版, 2009, 45(9): 57-64
- [26] 刘若慧, 刘保仓, 史开泉. 外 P-集合与 F-信息伪装 [J]. 系统工程与电子技术, 2011, 33(1): 116-119
- [27] 耿红琴, 张冠宇, 史开泉. F-信息伪装与伪装-还原辨识 [J]. 计算机科学, 2011, 38(2): 241-245
- [28] 汪洋, 张冠宇, 史开泉. P-集合与 \bar{F} -记忆信息特征-应用 [J]. 计算机科学, 2011, 38(2): 246-249
- [29] 于秀清. P-集合与 F-外嵌入信息辨识-发现 [J]. 计算机科学, 2011, 38(2): 250-253
- [30] 史开泉, 赵建立. 函数 S-粗集与隐藏规律安全-认证 [J]. 中国科学 E: 信息科学, 2008, 38(8): 1234-1243
- [31] Shi Kai-quan, Zhao Jian-li. Function S-rough sets and security-authentication of hiding law [J]. Science in China, F: Information Sciences, 2008, 51(7): 924-935
- [32] 史开泉, 崔玉泉. 双枝模糊决策与决策加密-认证 [J]. 中国科学 E: 技术科学, 2003, 33(2): 154-163
- [33] Shi Kai-quan, Cui Yu-quan. Both-branch fuzzy decision and decision encryption-authentication [J]. Science in China, F: Information Science, 2003, 46(2): 90-103
- [34] Koblitz N. Elliptic curve cryptosystems [J]. Mathematics and Computer, 1987, 21(1): 203-209
- [35] Stallings W. Cryptography and Network Security Principles and Practice [M]. London: Prentice-Hall Int Ltd, 1999: 193-206