

Vague 集之间的相似度量定义的再讨论

石玉强¹ 廖起彬² 王鸿绪³

(仲恺农业工程学院计算机科学与工程学院 广州 510225)¹

(仲恺农业工程学院教务处 广州 510225)² (琼州学院理工学院 三亚 572022)³

摘 要 从分析现有 Vague 值之间的相似度量的缺陷出发,给出一种 Vague 集相似度量新定义,试图解决该问题,并提出 Vague 集相似度量新公式。应用实例显示,该公式是实用的。

关键词 Vague 集,相似度量,新定义,新公式

中图法分类号 TP18 **文献标识码** A

Again Discussion for Definition of Similarity Measures between Vague Sets

SHI Yu-qiang¹ LIAO Qi-bin² WANG Hong-xu³

(College of Computer Science and Engineering, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)¹

(Dean's Office, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)²

(College of Science and Engineering, Qiongzhou University, Sanya 572022, China)³

Abstract Some faults of the similarity measures between vague sets were analyzed, and a new definition of the similarity measures between vague sets was given to attempt solving those faults. A new formula was put forward about the similarity measures between vague sets. An example shows that the proposed formulas is practicable.

Keywords Vague sets, Similarity measures, New definition, New formula

1 引言

迄今为止,相似度量不仅在 Vague 集理论中占有重要地位,而且在诸如 Vague 模式识别、Vague 决策、Vague 推理、Vague 集应用等许多应用领域都成为关键技术。但是,人们建立的许多 Vague 集的相似度量公式(比如文献[1-4]所提出的公式)皆有不足之处。本文希望用给出的 Vague 集之间的相似度量的新定义的方法来解决该问题。

2 现有 Vague 值之间的相似度量的不足之处

设 $x=[t_x, 1-f_x]$, $y=[t_y, 1-f_y]$ 为 Vague 值。文献[1-4]分别提出下列 Vague 值 x 和 y 之间的相似度量公式:

$$M_Q(x, y) = \frac{t_x \wedge t_y + f_x \wedge f_y + \pi_x \wedge \pi_y}{t_x \vee t_y + f_x \vee f_y + \pi_x \vee \pi_y} \quad (1)$$

$$M_F(x, y) = 1 - \frac{|t_x - t_y| + |f_x - f_y| + |\pi_x - \pi_y|}{2} \quad (2)$$

$$M_X(A, B) = 1 - \left(\frac{|t_x - t_y|}{1 + t_x + t_y} + \frac{|f_x - f_y|}{1 + f_x + f_y} \right) \quad (3)$$

$$M_J(x, y) = \begin{cases} 0.5, & \text{若 } x \text{ 为 } [0, 1] \text{ 或 } [1, 1], y \text{ 为 } [0, 1] \\ 1 - \frac{|t_x - t_y| + |f_x - f_y| + |\pi_x^2 - \pi_y^2| + |S_x - S_y|}{2 - (t_x - f_x) + 2 - (t_y - f_y)}, & \text{否则} \end{cases} \quad (4)$$

例 1 当 $x=y=[t_x, 1-f_x]$ 时,应用式(1)~式(4)计算,都能得出

$$M_Q(x, y) = M_F(x, y) = M_X(x, y) = M_J(x, y) = 1$$

这表明,通常事物都是自己与自己最相似。但是,仔细分析,从表 1 可见有些情况是不符合人们的直观认识的,是值得商榷的。

表 1 相似度量公式的分析表

x_i	$x_1=[0, 0]$	$x_2=[0.4, 0.4]$	$x_3=[0.4, 0.7]$	$x_4=[0, 1]$
y_i	$y_1=[0, 0]$	$y_2=[0.4, 0.4]$	$y_3=[0.4, 0.7]$	$y_4=[0, 1]$
$M_Q(x_i, y_i)$	1	1	1	1
$M_F(x_i, y_i)$	1	1	1	1
$M_X(x_i, y_i)$	1	1	1	1
$M_J(x_i, y_i)$	1	1	1	1
$M(x_i, y_i)$	1	1	0.7	0

在表 1 中, $M_Q(x_1, y_1) = M_F(x_1, y_1) = M_X(x_1, y_1) = M_J(x_1, y_1) = 1$, 这是正常的。因为不仅 $x_1 = y_1$, 而且 $\pi_{x_1} = \pi_{y_1} = 0$ 。即知普通集合的相似度量合理的数值为最大值 1。

$M_Q(x_2, y_2) = M_F(x_2, y_2) = M_X(x_2, y_2) = M_J(x_2, y_2) = 1$ 也是正常的。因为不仅 $x_2 = y_2$, 而且 $\pi_{x_2} = \pi_{y_2} = 0$ 。即知模糊集合的相似度量合理的数值也为最大值 1。

$M_Q(x_4, y_4) = M_F(x_4, y_4) = M_X(x_4, y_4) = M_J(x_4, y_4) = 1$ 是有悖于人们的认识的。因为, $x_4 = y_4 = [0, 1]$, 这个 Vague 值的真隶属度为零, 假隶属度也为零, 而不确定度却为最大值

到稿日期:2012-02-24 返修日期:2012-05-25 本文受广东省自然科学基金(S2011010003590), 海南省自然科学基金(610224), 海南省社会发展专项基金(2010SF004) 资助。

石玉强(1965—), 男, 教授, CCF 高级会员, 主要研究方向为信息技术、数据处理等, E-mail: yuqingshi@163.com; 廖起彬(1970—), 男, 硕士, 助理研究员, 主要研究方向为数据处理等; 王鸿绪(1946—), 男, 教授, 主要研究方向为模糊控制和数据处理等。

1. 人们对这个元素是完全不了解,一无所知的,无法比较它们是否相似。如果硬要求它们的相似度,则其至少不能等于1,最好等于零。

同理, $M_Q(x_3, y_3) = M_F(x_3, y_3) = M_X(x_3, y_3) = M_J(x_3, y_3) = 1$ 也是有悖于人们的认识的。因为, $x_3 = y_3 = [0.4, 0.7]$ 既不是普通集合的元素,也不是模糊集合的元素,它是一般的 Vague 元素值。其特点是 $x_3 = y_3$, 而且 $\pi_{x_3} = \pi_{y_3} \neq 0$ 。这意味着 x_3 和 y_3 都有一定的不确定性,所以它们的相似度量为 1 是不合理的,它应该等于某个非零数(由于 $\pi_{x_3} = \pi_{y_3} \neq 1$),但也不能等于 1。

基于这些分析,给出 Vague 集相似度量的新定义。

3 Vague 集相似度量的新定义及新公式

定义 1 设有 Vague 值 $x = [t_x, 1 - f_x]$, $y = [t_y, 1 - f_y]$ 。如果公式 $M(x, y)$ 满足下列准则,则称公式 $M(x, y)$ 是 Vague 值 x 和 y 之间的相似度量。

- (1) 模糊准则: $0 \leq M(x, y) \leq 1$;
- (2) 对称准则: $M(x, y) = M(y, x)$;
- (3) 最小准则: 当 $x = [0, 0]$, $y = [1, 1]$ 或者 $x = [1, 1]$, $y = [0, 0]$ 时, $M(x, y) = 0$;
- (4) 最大准则: $M(x, y) = 1 \Leftrightarrow x = y$ 且 $\pi_x = \pi_y = 0$;
- (5) 特殊准则: 当 $x = y = [0, 1]$ 时, 有 $M(x, y) = 0$ 。

说明: 准则(1)~(3)是一般相似度量都具备的准则, 它们的合理性是得到公认的; 准则(4)和准则(5)是弥补了在上述分析中所指出的现有相似度量的两点不足之处的条款。

定理 1 记

$$S_x = t_x - f_x, S_y = t_y - f_y, K_x = t_x + f_x, K_y = t_y + f_y$$

$$\text{则 } M(x, y) = 1 - \frac{|S_x - S_y| + |K_x - K_y| + \pi_x + \pi_y}{2} \quad (5)$$

是 Vague 值 $x = [t_x, 1 - f_x]$ 和 $y = [t_y, 1 - f_y]$ 之间的相似度量。

例 2 应用式(5)计算表 1 中的 Vague 值, 结果仍见表 1。计算结果满足了定义 1 的要求, 也满足了在上述分析中所指出的相似度量的合理数值的期望。这些结果正是我们所期待的。

定义 2 设论域 $X = \{x_1, x_2, \dots, x_n\}$, 其上有 Vague 集:

$$A = ([t_A(x_1), 1 - f_A(x_1)], [t_A(x_2), 1 - f_A(x_2)], \dots, [t_A(x_n), 1 - f_A(x_n)])$$

$$B = ([t_B(x_1), 1 - f_B(x_1)], [t_B(x_2), 1 - f_B(x_2)], \dots, [t_B(x_n), 1 - f_B(x_n)])$$

分别简记为:

$$A = ([t_{x_1}, 1 - f_{x_1}], [t_{x_2}, 1 - f_{x_2}], \dots, [t_{x_n}, 1 - f_{x_n}])$$

$$B = ([t_{y_1}, 1 - f_{y_1}], \dots, [t_{y_n}, 1 - f_{y_n}])$$

如果公式 $M(A, B)$ 满足下列准则, 则称公式 $M(A, B)$ 是 Vague 集 A 和 B 之间的相似度量。

- (1) 模糊准则: $0 \leq M(A, B) \leq 1$;
- (2) 对称准则: $M(A, B) = M(B, A)$;
- (3) 最小准则: 当 $A = ([0, 0], [0, 0], \dots, [0, 0])$, $B = ([1, 1], [1, 1], \dots, [1, 1])$ 或者 $A = ([1, 1], [1, 1], \dots, [1, 1])$, $B = ([0, 0], [0, 0], \dots, [0, 0])$ 时, $M(A, B) = 0$;
- (4) 最大准则: $M(A, B) = 1 \Leftrightarrow A = B$ 且 $\pi_{x_k} = \pi_{y_k} = 0, k = 1, 2, \dots, n$;

(5) 特殊准则: 当 $A = B = ([0, 1], [0, 1], \dots, [0, 1])$ 时, 有 $M(A, B) = 0$ 。

定理 2 下列公式 $M(A, B)$ 是 Vague 集 A 和 B 之间的相似度量:

$$M(A, B) = 1 - \frac{\sum_{i=1}^n \frac{|S_{x_i} - S_{y_i}| + |K_{x_i} - K_{y_i}| + \pi_{x_i} + \pi_{y_i}}{2}}{n} \quad (6)$$

定理 3 下列公式 $WM(A, B)$ 是 Vague 集 A 和 B 之间的加权相似度量:

$$WM(A, B) = 1 - \frac{\sum_{i=1}^n w_i \cdot \frac{|S_{x_i} - S_{y_i}| + |K_{x_i} - K_{y_i}| + \pi_{x_i} + \pi_{y_i}}{2}}{\sum_{i=1}^n w_i} \quad (7)$$

式中, 元素 x_i 的权重为 w_i ($0 \leq w_i \leq 1$) 且 $\sum_{i=1}^n w_i = 1$ 。

4 实例

许多文献都应用基于 Vague 集相似度量的 Vague 模式识别准则进行故障诊断^[6-8]和方案优选^[9-16]等方面的应用。

基于 Vague 集相似度量的 Vague 模式识别准则, 用 Vague 集表示的论域 $X = \{x_1, x_2, \dots, x_n\}$ 上有标准模式 A_i ($i = 1, 2, \dots, m$) 和待识别模式 B ; 应用 Vague 集之间的相似度量公式 $M(A, B)$ 计算 Vague 集 A_i ($i = 1, 2, \dots, m$) 和 B 之间的相似度量 $M(A_i, B)$ ($i = 1, 2, \dots, m$)。如果存在 $h \in \{1, 2, \dots, n\}$, 使得

$$M(A_h, B) = \max\{M(A_1, B), M(A_2, B), \dots, M(A_m, B)\}$$

则待识别模式 B 应归属于标准模式 A_h 。

例 3 设指标集为 $X = \{x_1, x_2, \dots, x_n\}$, 其上有标准模式 A_i ($i = 1, 2, 3$) 和待识别模式 B :

$$A_1 = ([0.8, 0.9], [0.3, 0.4], [0.6, 0.8])$$

$$A_2 = ([0.1, 0.6], [0.7, 0.8], [0.8, 0.9])$$

$$A_3 = ([0.5, 0.6], [0.5, 0.9], [0.5, 0.5])$$

$$B = ([0.4, 0.5], [0.6, 0.7], [0.8, 1.0])$$

应用相似度量式(6)计算, 结果为: $M(A_1, B) = 0.57$, $M(A_2, B) = 0.65$, $M(A_3, B) = 0.58$ 。

因为 $M(A_2, B) = \max\{M(A_1, B), M(A_2, B), M(A_3, B)\}$, 故由基于 Vague 集相似度量的 Vague 模式识别准则, 待识别模式 B 应归属于标准模式 A_2 。

结束语 提出的 Vague 集之间的相似度量的新定义, 不仅保留了原 Vague 集之间的相似度量的基本性质, 而且有效排除了当 $A = B = ([0, 1], [0, 1], \dots, [0, 1])$ 时, 有 $M(A, B) = 1$ 的困扰。提出的新公式(5)~(7)形式简单, 计算方便。这一思路有助于 Vague 集之间的相似度量的定义完善化。应用实例显示, 这样的公式是实用的和有效的。

参考文献

- [1] 邱卫根. Vague 集及其相似度量[J]. 计算机科学, 2007, 34(1): 156-158
- [2] 范九伦. Vague 值与 Vague 集上的相似度量[J]. 系统工程理论与实践, 2006, 8: 95-100
- [3] 徐凤生. Vague 集相似度量方法[J]. 计算机工程与应用, 2009, 45(5): 53-54, 58

(下转第 294 页)

图 10(a)反映了传统安全标签下主体访问资源的能力,图 10(b)反映了使用可变安全标签后主体访问资源的能力。相较于使用传统安全标签,可变安全标签的引入保证了主体对客体的访问能够得到一致、确定的访问授权,显著地提高了访问控制策略的适用范围;扩展了主体只读、只写能力范围,增添了主体可读写能力范围。

可变安全标签访问控制策略不仅显著减少了对系统可用性的影响,还能有效控制进程对系统资源的访问,实现操作系统安全加固。以机密性属性为例,进程安全标签随可信度的变化过程如图 11 所示。

原始数据子图即通过特征匹配得到的进程行为状态值曲线,其波动的剧烈程度反映了进程运行状态;进程行为状态值曲线经过 5 次迭代处理过程后得到较稳定的进程可信度曲线,如 5 次迭代子图所示;依据进程可信度曲线调整安全标签区间中点与区间半径,如安全标签半径子图所示;安全标签子图反映了不同时刻主体机密性属性上界与机密性属性下界。以机密性等级为 60 的文件为例,只有进程运行正常时才能执行对该文件的读、写操作,一旦程序异常,就失去对文件的访问操作,如安全标签子图中 API 数为 21 时所示。

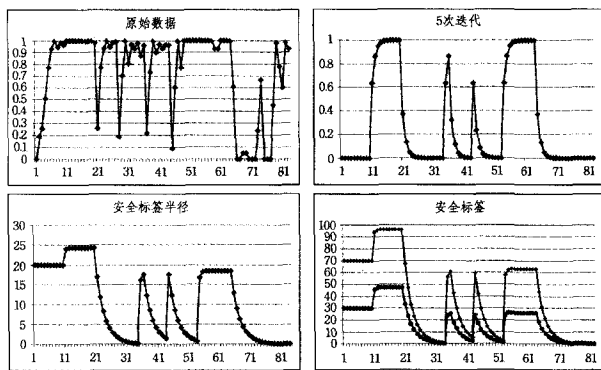


图 11 进程安全标签变化曲线

总之,进程安全标签随进程可信度的增加而扩张,随进程可信度的下降而收缩,有效限制了进程对系统资源的访问能力,降低了进程异常时的系统风险。

结束语 相对于经典的访问控制策略,可变标签的访问

控制策略以机密性和完整性两个不同但同时存在的属性为基础,在不影响系统可用性的前提下结合了 BLP 模型保障信息机密性与 Biba 模型保障信息完整性的优点,依据进程可信度动态调整主体安全标签,消除了因 BLP 与 Biba 作用域重叠导致的系统可用性低的问题。实现基于 Windows 文件系统过滤驱动的进程对文件访问的简单原型系统,有效控制了进程对资源的访问能力,加强了 Windows 操作系统安全防护能力。在今后的研究中,将进一步研究进程可信度与安全标签之间的变化关系,并解决因多次迭代而导致的访问控制作用延时问题。

参考文献

- [1] Net Application. 2011 年主机操作系统市场调查报告[OL]. <http://www.netmarketshare.com/>
- [2] 数据中心. 2011 年主机操作系统市场调查[OL]. <http://www.cnzz.com>
- [3] 黄强,沈昌祥,陈幼雷,等. 基于可信计算的保密和完整性统一安全策略[J]. 计算机工程与应用,2006,42(10):15-18
- [4] 马新强,黄羿,李丹宁. 基于扩充敏感标记的格理论模型研究[J]. 计算机工程,2009,35(21):171-173
- [5] 张俊,周正,李建,等. 基于 MLS 策略的机密性和完整性动态统一模型[J]. 计算机工程与应用,2008,44(12):19-21
- [6] 李益发,沈昌祥. 一种新的操作系统安全模型[J]. 中国科学(E 辑:信息科学),2006,36(4):347-356
- [7] 黄强,沈昌祥,陈幼雷,等. 基于可信计算的保密和完整性统一安全策略[J]. 计算机工程与应用,2006,42(10):15-18
- [8] 石文昌,孙玉芳. 多级安全性政策的历史敏感性[J]. 软件学报,2003,14(1):91-96
- [9] 郝东白,郭林,黄皓. 基于 Hook 的程序异常行为监测系统的设计与实现[J]. 计算机工程与设计,2007,28(18):4373-4376
- [10] 潘爱明. Windows 内核原理与实现[M]. 北京:电子工业出版社,2010:35-38
- [11] 冯力. 主机内核级入侵检测系统的建模与分析[D]. 西安:西安交通大学,2005

(上接第 256 页)

- [4] 江伟,梁家荣. Vague 集相似度量度的比较研究[J]. 计算机工程与应用,2011,47(25):30-33,64
- [5] 范平,梁家荣,李天志. Vague 集之间相似度量度的新方法[J]. 计算机工程与应用,2006,42(34):70-72
- [6] 叶军. 基于 Vague 集相似度量度的汽轮机故障诊断的研究[J]. 中国电机工程学报,2006,26(1):16-20
- [7] 赵忠伟. 基于 Vague 集相似度量度的柴油机磨擦类故障诊断的研究[J]. 绍兴文理学院学报,2006,26(9):68-72
- [8] 王鸿绪,张福金. Vague 故障诊断方法在振动故障诊断中的应用[J]. 计算机工程与应用,2011,47(26):225-227
- [9] 彭安华. Vague 集的相似度量度分析在材料选择中的应用[J]. 煤矿机械,2006,27(5):891-893
- [10] 叶军,徐永灿. 一种 Vague 相似度量度的方案决策方法[J]. 计算

- 机仿真,2006,23(4):90-93
- [11] 娄建国. Vague 集之间的相似度量及其在方案决策中的应用[J]. 工程设计学报,2005,12(6):325-328,333
- [12] 叶军,娄建国,李伟波. Vague 集间的相似度量度分析在机构方案决策中的应用[J]. 机械设计,2005,22(6):10-12
- [13] 王鸿绪. 应用 Vague 优化决策方法对小麦新品种综合评估[J]. 计算机工程与应用,2011,47(12):210-212
- [14] 杨洁,王鸿绪. 基于完整的 Vague 模式识别方法的战场目标识别[J]. 计算机工程与应用,2011,47(4):168-170
- [15] 江军,张慧坚,王鸿绪,等. 基于 Vague 相似度量度分析的椰子品种优选应用研究[J]. 江西农业学报,2011,23(2):55-57
- [16] 江军,张慧坚,王鸿绪,等. 基于 Vague 集的天然橡胶品种优化评估研究[J]. 广东农业科学,2011(6):200-201,206