

基于钟型函数的基本信度指派生成方法

郭威,朱珂,吴毅涛,肖二宁,吕平,宋克

(国家数字交换系统工程技术研究中心,河南郑州 450002)

摘要: 针对认知目标属性为单点值情况下,利用模糊相似度生成证据理论的基本信度指派(Basic Belief Assignment, BBA)存在悖论的问题,本文提出一种基于钟型函数(Bell Function, Bell-F)的相似度确定方法和相应的BBA生成算法.该方法避免了上述悖论算例与实际相左的情况,且能够细粒度刻画相似度评估模型,进一步提高了认知描述的准确性.通过对鸢尾花数据集进行识别实验,结果证明了该方法的有效性,并且对已知样本数量不足的非确定环境具备较好的适应能力.

关键词: 证据理论;基本信度指派;钟型函数;相似度评估模型;认知描述

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2016)04-0926-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2016.04.025

A Generation Method of Basic Belief Assignment Based on Bell Function

GUO Wei, ZHU Ke, WU Yi-tao, XIAO Er-ning, LÜ Ping, SONG Ke

(National Digital Switching System R&D Center, Zhengzhou, Henan 450002, China)

Abstract: In this paper, a method to determine similarity and the basic belief assignment (BBA) generation based on bell function (Bell-F) is proposed in order to solve the problem that the fuzzy similarity used in BBA of evidence theory exist paradox when the cognitive objective is a single point. This method can avoid the paradox which is inconsistent with the actual result in common sense, and can more granularly describe similarity evaluation model to improve the accuracy of cognitive description further. The results of a recognition experiment performed for data set of Iris, demonstrate the effectiveness of this method. Furthermore, this method still has good adaptability in uncertain environment when the number of known samples is insufficient.

Key words: evidence theory; basic belief assignment; bell function; similarity evaluation model; cognitive description

1 引言

证据理论(Dempster Shafer Theory, DST)作为不确定环境下数据融合和结论推理的有力方法,广泛应用于信息融合、认知决策领域.在理论实际应用中,如何进行基本信度指派(Basic Belief Assignment, BBA)是首要问题.近年来,伴随模式识别和人工智能的迅速发展,结合模糊理论,通过机器系统实现BBA自主生成,既能保证信度分配的稳定性,又能保证数据处理的灵活高效.其相比于人为指派方式显出诸多优势,逐渐成为学者们的关注方向.

在运用模糊理论进行目标描述时,模型可分为非线性和线性两类.非线性模型描述准确性高,根据场景不同,研究者们已经提出了一些典型应用和改进算

法^[1-4].然而,随着认知域扩大,识别目标增多,非线性模型应用的复杂度骤增,可实现性很低.因此,在处理认知域为多目标问题时,借助线性模型进行目标描述,通过计算其相似度进而生成BBA的方法得到广泛应用.2009年,Deng等人^[5]提出了基于回转半径(Radius of Gyration, ROG)计算模糊数间相似度,归一化后得到BBA;2011年,文成林等人^[6]从广义梯形模糊数(GT-FN)距离和形状考虑,综合指数距离,周长和面积等因素,提出了一种新的相似性度量方法并用于BBA生成;2012年,康兵义等^[7]分析了区间数在描述目标属性时简单易行的优势,提出了基于区间数(Crisp Interval)距离生成BBA的方法.

在前人研究基础上,本文对一些算例进行了讨论,发现当认知目标属性为单点值时,现有算法的相似度计算

结果与实际认知存在相悖情况;在进行分析后,针对此问题提出了解决思路,对线性模型刻画能力的缺陷进行说明;然后,根据认知规律特点,对多目标,属性为单点这一类问题引入一种钟型函数(Bell Function, Bell-F)作为构建相似度评估模型(Similarity Evaluation Model, SEM)的基础,进而提出一种新的 BBA 生成算法.通过对 UCI 数据库中鸢尾花数据集进行识别实验,结果证明该方法能够有效处理 BBA 生成的一些重要问题,构建了可信程度较高的认知模型,且应用于非确定环境时,SEM 表现出较好的适应能力,保持了较高的识别水平.

2 相关背景与问题分析

首先,简要介绍 DST 相关概念和几种典型的模糊相似度计算方法;然后,列举算法悖论算例,阐述其错误原因.最后,给出相应思考和解决思路.

2.1 DST 相关概念

DST 由 Arthur Dempster 提出^[8],随后 Glenn Shafer 给出其形式化表示^[9].下面对相关概念进行简要介绍:

概念 1 设一个互斥完备集合

$$\Theta = \{F_1, F_2, \dots, F_n\} \quad (1)$$

为识别框架,代表了整个认知域.记 Θ 的幂集为 2^Θ ,囊括了框架下所有子集,构建起该认知域中所有状态.

概念 2 在 Θ 上的 BBA 就是 mass 函数的映射过程:

$$m:2^\Theta \rightarrow [0,1], \sum_{A \in \Theta} m(A) = 1 \quad (2)$$

BBA 结果 $m(A)$,定量表示了对 A 的信度支持,本文遵循标准 BBA 准则,即限定了 $m(\emptyset) = 0$ 条件.

概念 3 Dempster 组合公式将相互独立的 BBA 进行组合,聚合和同步多源信息,得到一个联合 BBA 结果,为进一步决策提供有效数据支撑.组合公式如下:

$$m(A) = \frac{\sum_{A=B \cap C} m_1(B) \cdot m_2(C)}{1 - k}, \forall A \subseteq \Theta \quad (3)$$

系数 $k = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C)$,反映了进入组合公式的两个证据间的冲突程度.

2.2 模糊相似度计算方法

在多目标框架下,使用非线性模型直接构造 BBA 生成曲线难以实现.对此,研究者引入了模糊相似度 S ,使用现有认知(Acknowledge)和待测实例(New Instance, NI)构造线性模糊数 \tilde{A} 和 \tilde{N} ,如图 1(a)所示,并根据不同算法^[10]计算得出 $S(\tilde{A}, \tilde{N})$,标准化后最终得到 BBA 结果.一些应用场景中目标属性值退化为单点实数,如图 1(b)所示.

记 $\tilde{A} = (a_1, a_2, a_3, a_4; w_A)$, $\tilde{N} = (b_1, b_2, b_3, b_4; w_N)$ 为所构造的一般线性模糊数.文献[5]的算法(ROG)定义了回转半径点 (r_x, r_y) ,且给出相应计算公式:

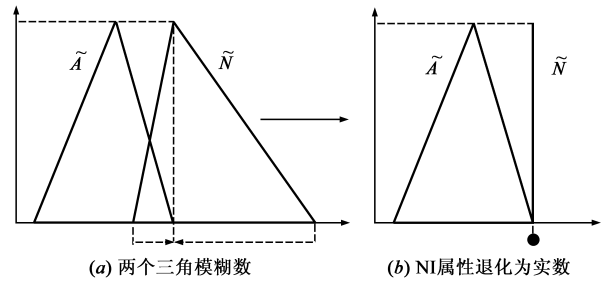


图1 用三角模糊数刻画现有认知和待测实例的某一属性

$$S(\tilde{A}, \tilde{N}) = \left[1 - \frac{\sum_{i=1}^4 |a_i - b_i|}{4} \right] \times (1 - |r_y^{\tilde{A}} - r_y^{\tilde{N}}|)^{B(S_1, S_2)} \times \frac{\min(r_x^{\tilde{A}}, r_x^{\tilde{N}})}{\max(r_x^{\tilde{A}}, r_x^{\tilde{N}})} \quad (4)$$

文成林^[6]方法(GTFN)考虑周长 a , 面积 P , 重心 (x^*, y^*) :

$$S(\tilde{A}, \tilde{N}) = \frac{\min(P(\tilde{A}), P(\tilde{N})) + \min(a(\tilde{A}), a(\tilde{N}))}{\max(P(\tilde{A}), P(\tilde{N})) + \max(a(\tilde{A}), a(\tilde{N}))} \times e^{-|x_i^* - x_j^*|} \quad (5)$$

康兵义^[7]将模糊数简化为区间数 (x_1, x_2) , 计算式(Crisp):

$$S(\tilde{A}, \tilde{N}) = \left[\frac{a_1 + a_2}{2} - \frac{b_1 + b_2}{2} \right]^2 + \frac{[(a_2 - a_1) + (b_2 - b_1)]^2}{12} \quad (6)$$

2.3 悖论算例

对不同应用场景下的算例进行挖掘,发现在 \tilde{N} 为实数条件下,现有算法的 S 计算结果与实际认知情况存在较大冲突,继续生成 BBA 将会导致错误发生.观察图 2 所示数据集,得到表 1 中的相悖对照结果.

表 1 数据集 S 计算结果与认知相悖

Method	No. A	Result S	No. B	Result S	Cognitive	Compute
ROG	Set1	0.2435	Set2	0.2656	$A = B$	$A \neq B$
ROG	Set3	0.6384	Set4	0.5349	$A < B$	$A > B$
ROG	Set5	0.5123	Set6	0.5091	$A < < B$	$A > B$
Crisp	Set7	0.2844	Set8	0.2592	$A < B$	$A > B$
Crisp	Set9	0.3660	Set10	0.3245	$A > > B$	$A > B$
Crisp	Set11	0.1176	Set12	0.0878	$A < < B$	$A > B$
GTFN	Set13	0.4713	Set14	0.7735	$A > B$	$A < B$
GTFN	Set15	0.6595	Set16	0.7452	$A > B$	$A < B$
GTFN	Set17	0.8155	Set18	0.8658	$A > B$	$A < B$

分析原因:当 \tilde{N} 退化为单点实数时,模糊区间长度,模糊数面积,隶属变化速度等其自身一些属性也随之退化变得没有意义.然而,这些退化属性恰恰是现有算法中评定 S 的必需因素,正是这种退化情况的发生,使得计算公式失效,最终导致了悖论结果的出现.

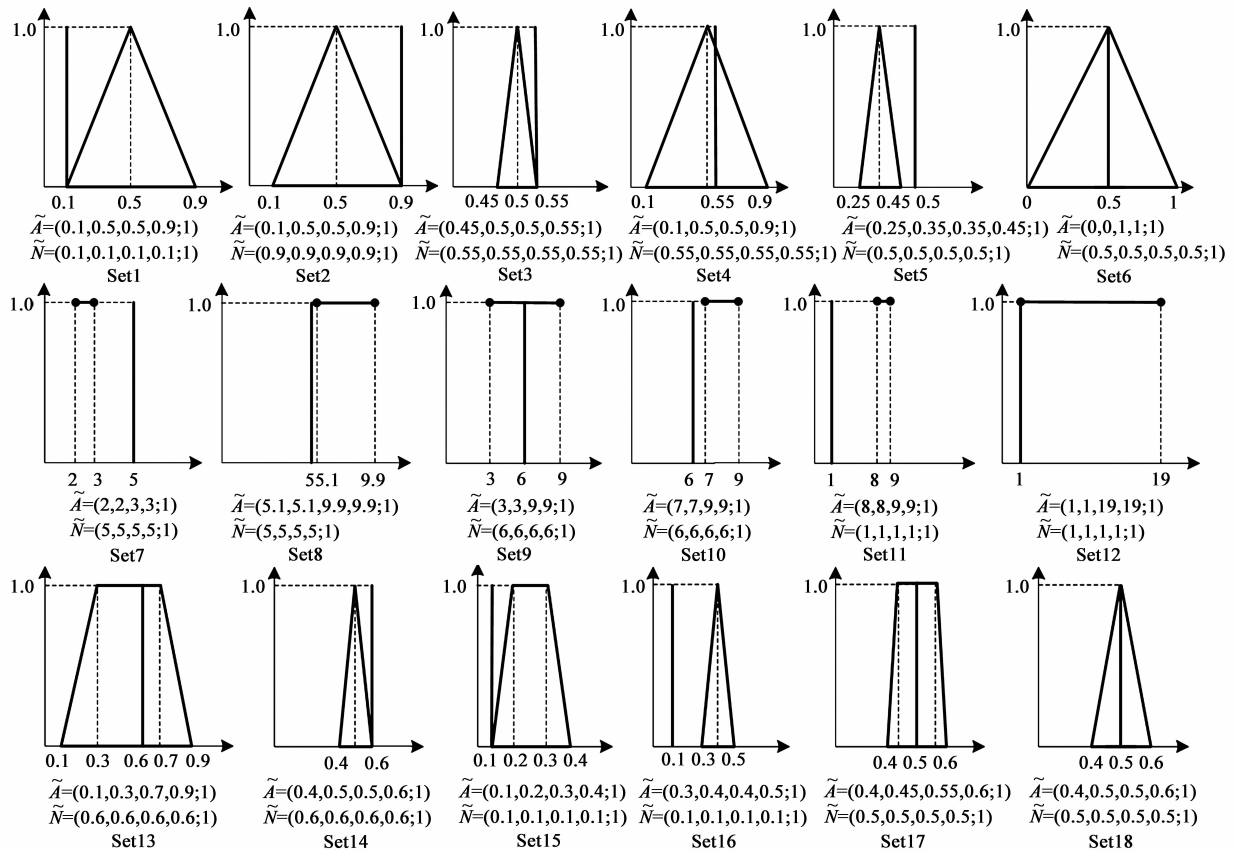


图2 悖论数据集

2.4 解决思路

鉴于上述情况,本文在 S 计算时,不再考虑退化属性因素,采用最大隶属原则得出 S 值^[11]. 由于该方式依赖于交点的隶属取值,这就对 \tilde{A} 的构建提出了更高的要求.

使用线性模型进行属性描述虽然简单易行,但其隶属度变化完全均匀,没有体现出认知规律的诸多特征,这种粗粒度的描述方式,结果会与实际情况发生偏差,难以满足 \tilde{A} 的构建要求,因此无法保证生成的 BBA 合理可靠.

3 基于 Bell-F 的相似度确定方法

本节将根据一般认知规律,引入一种 Bell-F 作为 SEM 的构建基础.

3.1 认知规律分析

考虑一应用场景,认知域为

$$\Theta = \{ \text{District A} | A \in \text{Beijing} \}$$

属性房价 (houseprice) 作为识别证据之一,以目标西城区 (Xicheng) 为例 (单位:元/平方米).

最初,认知主体通过一些见闻和猜测,认为 Xicheng 的 houseprice 可能最接近 40000,浮动 10000 上下. 落于

浮动区间内的 NI 目标,推测为 Xicheng 的肯定度较大,区间外则肯定度较小. 提炼关键信息,确定初始认知状态:

- (a) 中心点为 40000;
- (b) 判决边界为 [30000, 50000] ;

通过网上查阅资料,咨询购房公司等渠道,认知主体对 Beijing 的 houseprice 了解逐渐深入,认知状态发生变化:

- (a) 中心点为 38000;
- (b) 判决边界为 [26000, 52000], 边界发生移动和扩张;

(c) 另外,认知“决断性”会增强:对中心点附近和两侧远离的区域更加“确定”,肯定度变化更缓,而接近判决边界的目标,肯定度增降会更快.

对多个类似案例分析,总结相似性,能够得出一些认知规律:

- (a) 存在一个或几个肯定度最强的中心点;
- (b) 存在一个判决边界,边界附近肯定度变化最大;
- (c) 随着认识的深入,决断性增强,认知曲线会发生变化.

如果能将这些认知规律融入模型构建,可以提高 SEM 的合理性、严谨性,得到更准确的认知描述结果。

3.2 钟型函数

本文引入 Bell-F 来刻画认知目标属性,其函数表达式为:

$$f(x) = \frac{w}{1 + (\frac{x-c}{a})^{2b}} \quad (7)$$

如图 3 所示, Bell-F 通过参数 a, b, c, w 确定, c 处取得最大隶属值 w , 在 $c-a, c+a$ 处取得 $w/2$, b 可以控制曲线形状. w 为最大隶属值, 取值范围为 $[0, 1]$, 本文中限定 $w = 1$.

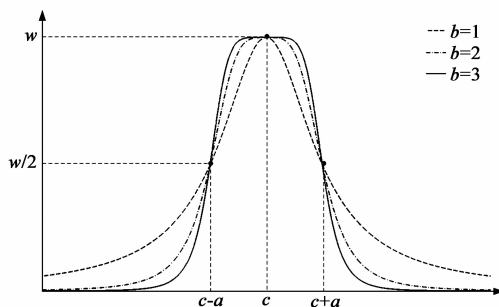


图3 Bell-F函数曲线及关键取值点

分析曲线, c 是函数中心点 (Central Point, CP), 其附近的隶属度变化缓慢. 认知过程中, 越临近某目标属性 CP, 推断是该目标的肯定性越强, S 越大; a 控制区间 $[c-a, c+a]$ 长度. 将 $S = 0.5$ 视为判决门限, a 定为单向判决长度, 则 $c-a, c+a$ 为判决点 (Judge Point, JP), JP 附近隶属度变化最为陡峭; b 则恰好刻画认知的“决断能力” (Decision Grade, DG), 随着 b 增大, 曲线会逐渐“收紧”, CP 点附近的变化更缓, JP 点附近的变化更陡, 表达出随着认知程度逐渐深入, 主体对判断的“信心”增强的认知特征。

分析可知, Bell-F 对 3.1 中认知规律的诸多特征进行了描述, 能够细粒度地对 SEM 进行构建, 解决了线性模型认知描述过程存在缺陷, 而无法满足建模要求的问题。

3.3 SEM 构建

记认知域 m 个目标

$$\Theta = \{C_1, C_2, \dots, C_m\}$$

有 n 个认知证据 (即目标属性), 取值限为单点实数

$$A = \{A_1, A_2, \dots, A_n\}$$

种类 i 对应的属性 j 记为

$$A_j^{C_i}, i \in [1, m] \text{ and } j \in [1, n]$$

那么, 属性 j 的 SEM 为

$$\text{model}_j = \{A_j^{C_1}, A_j^{C_2}, \dots, A_j^{C_m}\}, j \in [1, n]$$

根据已知样本情况, 提取特征参量: 均值 ξ , 标准差 D_{std} , 最大差 D_{max} 和样本数量 N , 如图 4 所示进行 SEM 构建。

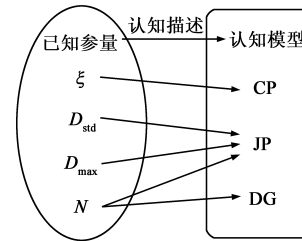


图4 SEM构建概念图

将已知参量通过认知描述过程, 确定 SEM 中各目标属性 Bell-F 参数:

定义 1 单向判决长度 a

$$a = \begin{cases} D_{std} + \frac{D_{max} - D_{std}}{2} \times \frac{N}{N_{ta}}, & N \leq N_{ta} \\ D_{std} + \frac{D_{max} - D_{std}}{2}, & N > N_{ta} \end{cases} \quad (8)$$

N_{ta} 为 SEM 达最大判决长时 N 的阈值, 可根据所在应用场景下, 判决点由标准差 (认知初始态) 向最大差 (认知成熟态) 转移过程所经历的样本量而定。

定义 2 决断强度 b

$$b = \begin{cases} \frac{N}{N_{tb}} + 1, & N \leq N_{tb} \\ 2, & N > N_{tb} \end{cases} \quad (9)$$

最小值设为 1, N_{tb} 为认知状态达最大决断强度时数量阈值, 由认知难易程度而定. 难度大的应用场景, 需要更多的已知数量增加“信心”, 因此 N_{tb} 增大; 反之亦然。

定义 3 中心点 c

$$c = \xi \quad (10)$$

c 是目标属性的中心位置, 相似度 S 取得最大值 1。

3.4 目标并集的隶属度确定

利用 SEM 进行相似度确定时, 多目标 $C_1 \sim C_k$ 在同一属性 A_x 上存在交叉情况, 称为目标并集 \cup . 值 $S(NI, \cup)$ 能反映在属性 A_x 上, 待测目标 NI 与 \cup 中各目标同时接近的程度, 用于后续对 $\{C_1, C_2, \dots, C_k\}$ 指派信度. 本文提出一种目标并集的相似度的计算方法:

$$S_{C_1 \cup C_2 \cup \dots \cup C_k}(NI) = \min \theta_{C_i \cup C_j} \times E_s \times \psi(NI) \quad (11)$$

式中 $E_s = (S_{C_1} + S_{C_2} + \dots + S_{C_k})/k, x, y, p, q \in [1, k], \psi(NI) = 1 - \max |S_{C_p} - S_{C_q}|$.

如图 5 示, θ 为 model 中目标交点 S 值, 并集中 C 的个数为 k . $S(NI, \cup)$ 结果越大, 说明从 $C_1 \sim C_k$ 中辨别出 NI 真实状态的难度越大; 反之则区分的难度越低。

4 一种新的 BBA 生成方法

本节在上一节基础上, 提出一种新的 BBA 生成方

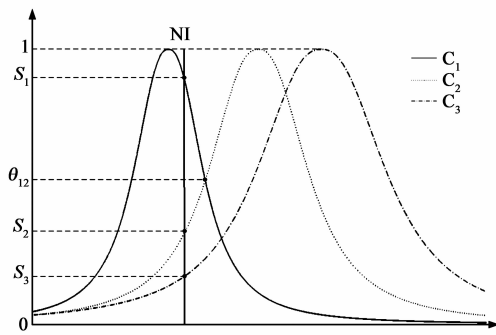


图5 目标并集的相似度确定

法,详细步骤如下:

Step1 从已知样本中提取所有目标证据信息 A_x , 基于 Bell-F 建立关于属性 x 的 $model_x$;

Step2 利用最大隶属原则确定 NI 与各目标相似度 $S(NI, model_x) = \{S_x^{C_1}, S_x^{C_2}, \dots, S_x^{C_n}\}, x \in [1, n]$;

Step3 计算目标并集相似度 $S(NI, \cup)$;

Step4 对所有 S 规则化处理,生成证据 A_x 提供的 BBA.

UCI 是用于机器学习、模式识别的开源数据库,其中的鸢尾花数据集^[12](Iris Data Set, IDS)符合多目标单点属性的应用场景,用于本文示例和实验.

IDS 认知域有三个目标: Setosa (Se)、Versicolor (Ve) 和 Virginica (Vi), 样本总量 150, 其中每个目标 50 个:

$$\Theta = \{C_1 = Se, C_2 = Ve, C_3 = Vi\},$$

$$N_{C_1} = N_{C_2} = N_{C_3} = 50$$

目标有四种属性特征,分别是花萼长度 (Sepal Length, SL), 花萼宽度 (Sepal Width, SW), 花瓣长度 (Petal Length, PL), 花瓣宽度 (Petal Width, PW), 属性单位: cm.

$$A = \{A_1 = SL, A_2 = SW, A_3 = PL, A_4 = PW\}$$

以 SL 为例,推导由 A_1 生成证据 BBA 的详细过程.

Step1 建立属性 SEM.

根据认知应用场景,确定 SEM 的场景参数,本文设定 N_{ta} 与 N_{tb} 值为 200, b 取值范围为 $[1, 2]$, 提取样本关于 A_1 的特征信息,带入式 (8) ~ (10) 得出 $model_1$ 参数,如表 2 所示.

表 2 属性 A_1 特征信息以及对应 $model_1$ 参数

	ξ	D_{std}	D_{max}	N	a	b	c
$A_1^{C_1}$	5.0060	0.3525	0.7940	50	0.3967	1.25	5.0060
$A_1^{C_2}$	5.9360	0.5162	1.0640	50	0.5710	1.25	5.9360
$A_1^{C_3}$	6.5880	0.6359	1.6880	50	0.7411	1.25	6.5880

Step2 抽取一个样本数据 NI_x 求相似度 $S(NI_x, model_1)$.

如图 6 所示,取 $NI_x = (A_1 = 6.2, A_2 = 2.2, A_3 = 4.5, A_4 = 1.5)$ 的属性 $A_1 = 6.2$ 进入 $model_1$ 评估,确定 $S(NI_x, model_1)$ 结果.

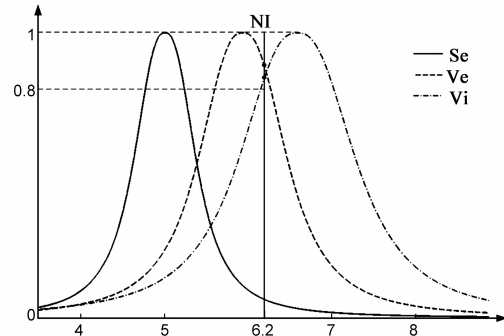


图6 确定 NI_x 与 $model_1$ 中各目标间相似度

Step3 根据式 (11) 计算 $S(NI_x, model_1 \cup)$.

将 Step2 中单集结果和目标并集结果整合如表 3 所示.

表 3 $S(NI_x, model_1)$ 结果

	$S_1^{C_1}$	$S_1^{C_2}$	$S_1^{C_3}$	$S_1^{C_1 \cup C_2}$	$S_1^{C_2 \cup C_3}$	$S_1^{C_1 \cup C_3}$	$S_1^{C_1 \cup C_2 \cup C_3}$
NI_x	0.0638	0.8795	0.8462	0.0706	0.8238	0.0573	0.0636

Step4 规则化处理生成 BBA, 如表 4 所示.

表 4 由 SL 属性得出的 BBA 结果

Notation	State	BBA
C_1	{Se}	0.0227
C_2	{Ve}	0.3136
C_3	{Vi}	0.3017
$C_1 \cup C_2$	{Se, Ve}	0.0252
$C_2 \cup C_3$	{Ve, Vi}	0.2937
$C_1 \cup C_3$	{Se, Vi}	0.0204
$C_1 \cup C_2 \cup C_3$	{Se, Ve, Vi}	0.0227

观察表 3, 表 4 数据, 分析算法结果.

NI_x 与 Ve, Vi 和 $\{Ve, Vi\}$ 相似程度都超过了 0.8, 所以 BBA 生成时被赋予了较高的信度. 还原认知推断情况: 当观测到一个花萼长度 (SL) 为 6.2cm 的 NI 时, 本文方法将信度指派于目标 Ve 和 Vi , 且两者间犹豫度较大. 考察指派结果合理性, 以 SL 属性为基准分析数量, 如表 5 所示.

表 5 显示出算法对支持信度的分配符合实际情况, 验证了本文方法的有效性. 需说明的是, 犹豫度高是由多种因素引起的. 其原因之一是证据的辨识度不够, 导致相应属性 SEM 建立时目标间重合程度较高, 因此在实际应用中, 选取高辨识度属性是提高结果正确率的一个重要手段.

表 5 准属性数量分析

Interval	Se	Ve	Vi	All
6.2 ± 0.0	0	2	2	4
6.2 ± 0.1	0	9	10	19
6.2 ± 0.2	0	15	17	32
6.2 *	0	16	13	29

注: * 为 NI 所在一侧,本表中对于 Se, Ve 为右侧,对于 Vi 为左侧

5 实验和结果分析

本节将对 IDS 进行全面实验,得出实验结果并统计识别正确率.具体步骤如下:

(1) 随机选取 Se, Ve, Vi 目标样本各 40 个,共 120 个作为认知样本集,用于信息提取;

(2) 构建 model₁ ~ model₄对剩余未知样本进行分类实验,利用本文方法生成各属性的证据 BBA₁ ~ BBA₄;

(3) 将 BBA₁ ~ BBA₄运用 Dempster 组合公式生成联合 BBA,查看识别正确与否,统计结果;

(4) 减少样本集数量 N,模拟非确定环境,再重复上述实验过程,观察正确率变化情况.

依照上述实验步骤,开始进行实验,这里抽取一个 NI 作为示例:

$$NI_1 = (A_1 = 5.1, A_2 = 3.5, A_3 = 1.4, A_4 = 0.2)$$

整个过程由 SEM 建立到 BBA 生成,再经过多次证据组合,得到最终的联合 BBA(如表 6 所示).组合结果表明 NI₁为目标 Se,查看 IDS,实验结果与事实情况相符.

表 6 实验中生成 BBA 和最终组合结果

State	BBA ₁ (SL)	BBA ₂ (SW)	BBA ₃ (PL)	BBA ₄ (PW)	Combined ₁	Combined ₂	Combined ₃
{Se}	0.6193	0.6772	0.9613	0.9502	0.8748	0.9968	1.0000
{Ve}	0.1502	0.0911	0.0187	0.0231	0.0510	0.0017	0.0000
{Vi}	0.0781	0.2089	0.0106	0.0141	0.0726	0.0016	0.0000
{Se, Ve}	0.0504	0.0055	0.0016	0.0037	0.0007	0.0000	0.0000
{Ve, Vi}	0.0778	0.0022	0.0065	0.0062	0.0005	0.0000	0.0000
{Se, Vi}	0.0134	0.0143	0.0007	0.0016	0.0006	0.0000	0.0000
{Se, Ve, Vi}	0.0109	0.0008	0.0005	0.0011	0.0000	0.0000	0.0000

重复实验步骤,对其余待测样本进行识别,统计结果得:对未知样本识别率为 93.3%,识别的两处错误如表 7 所示.

表 7 错误案例

NO.	SL	SW	PL	PW	Inference	Reality
1	4.9	2.5	4.5	1.7	Ve	Vi
2	6.9	3.1	4.9	1.5	Vi	Ve

上述实验表明,当 N 确保一定规模时,本文方法能够给出较可靠的信度支持.但在一些非确定应用场景下,N 可能远远不够,此时就需要模型具有良好的适应性,能够根据实际场景进行调整.

依次降低 N 的大小,重新对 IDS 进行识别实验,结果如图 7 所示,在 N 变化的过程中,识别实验仍能保持不低于 80% 的较高正确率,显示出模对已知信息不足的非确定环境具备很好的适应能力.

如图 7 所示,对部分实验实例分析可知,即便组合结果是正确的,但过程中各证据 BBA 没有给出可靠的信度支持,如表 8 所示.

对表 8 的 BBA 组合后,虽然判别结果 Vi > Ve 与实际相符,但在 SL, SW, PL 三种属性上 Ve 的信度支持反而更高,出现信度偏差情况.这带给研究的启示是:

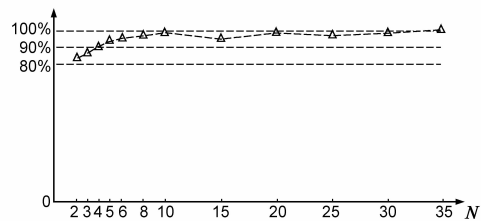


图 7 不同 N 下识别正确率

表 8 偏差 BBA 举例

State	SL	SW	PL	PW
{Se}	0.0692	0.0875	0.0008	0.0013
{Ve}	0.3959	0.3432	0.5090	0.0562
{Vi}	0.2306	0.3055	0.3474	0.9402
{Se, Ve}	0.0395	0.0609	0.0047	0.0010
{Ve, Vi}	0.2071	0.1020	0.1324	0.0009
{Se, Vi}	0.0295	0.0764	0.0025	0.0003
{Se, Ve, Vi}	0.0281	0.0244	0.0033	0.0001

(1) 由于不存在无任何交叉的理想判别证据,因此信度存在一定程度上的偏差是时常发生的;

(2) 利用多个判别证据生成 BBA 能够有效校正这种偏差现象,原因是从多维度对事物进行认知和推断

更容易得到正确结果;

(3)低辨识度的证据容易产生信度偏差(官方公认IDS中的属性SW),这在上述实验中也有明显体现.

6 结论及下一步工作

针对多目标,属性为单点实数的应用场景,本文提出一种基于 Bell-F 的 BBA 生成算法,避免了文中列举的悖论情况.通过 3.2 中的 Bell-F 曲线分析以及 3.3 中模型构建时的考虑,体现了本文在认知描述时的具体思考,说明了基于 Bell-F 在构建 SEM 和生成 BBA 时的独特优势.实验结果验证了本文算法的有效性,且能够适应认知不足的非确定环境,达到了预期效果.

本文还存在以下问题:模型方面,由于本文引入的 Bell-F 是对称的,没有考虑对非对称目标的刻画,在非对称属性场景下,BBA 可靠性会受到影响,需要进一步研究提出解决方案;算法方面,对于目标并集的相似度确定需要进一步的讨论,深入探讨其物理意义以及计算方法;应用方面,仍需在更多应用场景下实现,进一步检验算法的可用广度.另外,对于不同应用环境下,各场景参数的确定还需进一步探索.

参考文献

- [1] 诸葛建伟,王大为,陈昱,等.基于 D-S 证据理论的网络异常检测方法[J].软件学报,2006,17(3):463-471.
Zhuge J W, Wang D W, Chen Y, et al. A network anomaly detector based on the D-S evidence theory[J]. Journal of Software, 2006, 17(3):463-471. (in Chinese)
- [2] Nguyen-Thanh N, Koo I. An enhanced cooperative spectrum sensing scheme based on evidence theory and reliability source evaluation in cognitive radio context[J]. IEEE Communications Letters, 2009, 13(7):492-494.
- [3] Dezert J, et al. Edge detection in color images based on DSMT[A]. Proceedings of the 14th International Conference on Information Fusion[C]. Chicago:IEEE, 2011. 969-976.
- [4] 唐良瑞,等.基于证据理论的无线传感器网络链式路由算法[J].计算机研究与发展,2013,50(7):1362-1369.
Tang L R, et al. A chain routing algorithm based on evidence theory in wireless sensor networks[J]. Journal of Computer Research and Development, 2013, 50(7):1362-1369. (in Chinese)
- [5] Deng Y, Jiang W, Xu X, et al. Determining BPA under uncertainty environments and its application in data fusion[J]. Journal of Electronics, 2009, 26(1):13-17.
- [6] 文成林,周哲,徐晓滨.一种新的广义梯形模糊数相似性度量方法及在故障诊断中的应用[J].电子学报,2011,39(3A):1092-1096.
Wen C L, Zhou Z, Xu X B. A new similarity measure between generalized trapezoidal fuzzy numbers and its application to fault diagnosis[J]. Acta Electronica Sinica, 2011, 39(3A):1092-1096. (in Chinese)
- [7] 康兵义,李娅,等.基于区间数的基本概率指派生成方法及应用[J].电子学报,2012,40(6):1092-1096.
Kang B Y, Li Y, et al. Determination of basic probability assignment based on interval numbers and its application[J]. Acta Electronica Sinica, 2012, 40(6):1092-1096. (in Chinese)
- [8] Dempster A P. Upper and lower probabilities induced by a multiple valued mapping[J]. The Annals of Mathematical Statistics, 1967, 38(2):325-339.
- [9] Shafer G. A Mathematical Theory of Evidence[M]. Princeton:Princeton University Press, 1976.
- [10] B Farhadinia, A I Ban. Developing new similarity measures of generalized intuitionistic fuzzy numbers and generalized interval-valued fuzzy numbers from similarity measures of generalized fuzzy numbers[J]. Mathematical and Computer Modelling, 2013, 57(8):812-825.
- [11] 李弼程.模式识别原理与应用[M].陕西西安:西安电子科技大学,2008.154-159.
- [12] Iris Data Set. Famous Database for Pattern Recognition from Fisher[OL]. <http://archive.ics.uci.edu/ml/datasets/Iris>, 2011-3-20.

作者简介



郭威男,1990年8月出生于北京.国家数字交换系统工程技术研究中心在读硕士研究生,主要研究方向为信息科学、计算机体系结构、网络安全.
E-mail: guowjss@126.com



朱珂男,1975年8月出生于河南开封,博士.国家数字交换系统工程技术研究中心副教授、硕士生导师,主要研究方向为计算机体系结构.