

文章编号: 1000-5862(2016)03-0285-05

一种基于大焦元分解的信任函数逼近方法

黄梅, 吴根秀*, 刘邱云, 吴恒, 毛临川

(江西师范大学数学与信息科学学院, 江西 南昌 330022)

摘要: 针对证据合成提出一种基于大焦元分解的信任函数逼近方法, 首先将基数过大的焦元进行分解, 将焦元基数控制在不大于 k 的范围内, 然后再按照改进的能量函数删减焦元, 这样不仅减少焦元的个数, 也控制了焦元基数, 更优化了在证据合成时的计算复杂度, 并且试验结果也表明了该方法的有效性.

关键词: D-S 证据理论; 基本概率赋值; 能量函数; 平均能量函数

中图分类号: O 236 文献标志码: A DOI: 10.16357/j.cnki.issn1000-5862.2016.03.14

0 引言

D-S 证据理论^[1]是信息融合的有力工具, 但是其计算量问题限制了证据理论在工程中的广泛应用. 自从 20 世纪 80 年代以来, 学者们着手于这一问题的研究, 到目前为止, 计算量问题仍然是研究的热点. 计算量主要产生于焦元的组合过程, 因此, 信任函数的逼近主要从 2 方面入手: (i) 改进组合规则^[2-5], 如 J. Gordon 等^[2]提出的层次假设空间和 GS 算法; (ii) 控制焦元, 加快运算速度^[6-10], 如基于能量函数^[9]和改进的能量函数^[10]的逼近算法. 但是对于改进的能量函数的逼近方法虽然看似解决了基于传统能量函数在删减焦元时的的问题, 但是此方法当证据焦元的基数差距较大时, 保留基数大的大信度焦元的同时基数小的小信度焦元也会过分保留, 从而降低了其减少计算复杂度的有效性. 因此, 对于这一问题, 本文在删减焦元前先对焦元进行处理, 即将证据焦元按基数大小排序, 把基数大的焦元进行分解, 并将大基数焦元的基本概率赋值分配给其单子集焦元中, 然后再按照改进的能量函数删减焦元. 这样不仅减少焦元的个数, 也控制了焦元基数, 从而更优化了在证据合成时的计算复杂度, 并且试验结果也表明了该方法的有效性.

1 证据理论的基础知识

设 Θ 是表示所有问题的可能情况的一个论域

集合, 称 Θ 为识别框架, 2^Θ 表示识别框架中所有命题的可能性组合.

定义 1^[11-13] 设 Θ 为一识别框架, 函数 $m: 2^\Theta \rightarrow [0, 1]$ 满足

$$\begin{cases} m(\emptyset) = 0, \\ \sum_{A \subseteq \Theta} m(A) = 1, \end{cases}$$

称 m 为 Θ 上的基本信任指派函数(BPA). 若 $\forall A \in \Theta, m(A) > 0$, 则称 A 为焦元, 所有焦元的并称为核. $m(A)$ 反映了证据对识别框架中命题 A 的支持程度.

若 $Bel(A) = \sum_{B \subseteq A} m(B), \forall A \subseteq \Theta$, 则称 $Bel: 2^\Theta \rightarrow [0, 1]$ 为信任函数^[14]. 若 $Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$,

$\forall A \subseteq \Theta$ 则称 $Pl: 2^\Theta \rightarrow [0, 1]$ 为似然函数^[15].

定理 1^[1] 设 m_1 和 m_2 为识别框架 Θ 上 2 个证据, 其焦元分别 $B_i, i = 1, 2, \dots, n; j = 1, 2, \dots, m$, 则 Dempster 的组合规则为

$$m(A) = \begin{cases} 0, & A = \emptyset, \\ \frac{\sum_{B_i \cap C_j = A} m_1(B_i) m_2(C_j)}{1 - K}, & A \subseteq \Theta, A \neq \emptyset, \end{cases} \quad (1)$$

其中 $K = \sum_{B_i \cap C_j = \emptyset} m_1(B_i) m_2(C_j), 0 \leq K < 1$, 它反映了对同一识别框架下各证据间的冲突程度, 称为冲突系数.

2 已有的信任函数的逼近方法

2.1 概括逼近法

J. D. Lowrance^[6] 首先提出了一种普通的方法——

收稿日期: 2015-12-30

基金项目: 江西省自然科学基金(20151BAB207030) 和江西省教育厅科学技术课题(GJJ14244) 资助项目.

通信作者: 吴根秀(1965-), 女, 江西南丰人, 教授, 主要从事数据挖掘、不确定性推理、信息融合方面的研究.

概括逼近法. 通过去掉基本概率赋值较小的焦元做逼近, 使得 m' 只剩下 k 个焦元. 这里令 m 表示给定 BPA 函数, 带有 s 个焦元 A_1, A_2, \dots, A_s , 且满足 $m(A_i) \geq m(A_{i+1})$, $i \in \{1, 2, \dots, s-1\}$. 将 $i \geq k$ 的焦元合并后形成的新焦元作为第 k 个焦元, 其基本概率赋值为参与合并的焦元的基本概率赋值的和, 而 $i < k$ 的焦元的基本概率赋值不变, 即 m 的概括逼近 m' 的具体公式为

$$m'(A) = \begin{cases} m(A), & A = A_i, i \in \{1, 2, \dots, k-1\}, \\ \sum_{j=k}^s m(A_j), & A = \bigcup_{j=k}^s A_j. \end{cases}$$

2.2 双逼近方法

双逼近方法^[8]是 David Harmanec 提出的 5 种启发式的逼近方法中的一种, 它是通过每一步合并 2 个焦元来减少当前的 BPA 函数的焦元数, 直到剩下 k 个焦元为止.

定理 2 m 为识别框架 Θ 上的 BPA 函数, A, B 为 m 的 2 个不同焦元, m' 为 m 合并焦元 A, B 而得到的新的 BPA 函数, 即

$$m'(C) = \begin{cases} 0, & C = \{A, B\}, \\ m(A) + m(B), & C = A \cup B, \\ m(C), & \text{其他}, \end{cases}$$

则

$$O(Bel, Bel') = 2^{|\Theta-A|} m(A) + 2^{|\Theta-B|} m(B) - 2^{|\Theta-A \cup B|} (m(A) + m(B)), \quad (2)$$

其中 Bel 和 Bel' 分别是 m 和 m' 对应的信任函数.

由定理 2 可知, 要得到 Bel 的最佳逼近, 只要通过每一步选取 2 个不同的焦元 A, B 合并, 使 (2) 式达到最小值, 直至剩下 k 个焦元为止.

2.3 基于能量函数的逼近方法

定义 2 证据的焦元能量函数为 $E(A) = m(A)/|A|$, 其中 A 为识别框架 Θ 上的焦元, $|A|$ 为焦元的基数.

定义 3 证据的平均能量函数为 $\bar{E}_i = (\sum_{j=1}^{N_j} m_i(A_j)/|A_j|)/N_j$, 其中 N_j 为第 i 个证据所支持的焦元的总个数.

设 m 为识别框架 Θ 上的 BPA 函数, 通过分别计算 m 的焦元能量值和平均能量值, 将焦元能量值不

小于平均能量值的焦元保留下来, 其他焦元进行分配, 分配后的基本概率赋值叠加到对应证据的保留焦元上, 从而得到新的 BPA 函数 m' , 即

$$\begin{cases} m'(R_j) = m(R_j) + \sum_{k=1}^{N_d} m(D_k) c_{kj}/N_k, & j = 1, 2, \dots, N_r, \\ m'(\Theta) = 1 - \sum_{j=1}^{N_r} m'(R_j), \end{cases}$$

其中 D_k, R_j 分别表示抛弃焦元和保留焦元, N_d, N_r 分别表示抛弃焦元和保留焦元的个数, N_k 表示保留焦元中与抛弃焦元 D_k 有关的焦元个数, $c_{kj} = |D_k \cap R_j|/|D_k|$, $j = 1, 2, \dots, N_r$.

但是当焦元基数过大时, 传统的焦元能量函数将会导致大信度焦元被遗弃的后果, 从而使得合成结果严重的不合理. 因此, 文献 [10] 提出了一种改进的能量函数的逼近方法, 焦元能量函数为

$$E(A) = f(\lambda) m(A) / (1 + \lambda |A|), \quad (3)$$

其中 A 为识别框架 Θ 上的焦元, $|A|$ 为焦元的基数, λ 为修正因子, $f(\lambda) = \begin{cases} 1, & \lambda < 1 \\ \lambda, & \lambda \geq 1 \end{cases}$ 为修正因子函数.

平均能量函数为

$$\bar{E}_i = (\sum_{j=1}^{N_j} f(\lambda) m_i(A_j) / (1 + \lambda |A_j|)) / N_j, \quad (4)$$

其中 N_j 为第 i 个证据所支持的焦元的总个数.

3 一种基于大焦元分解的信任函数逼近方法

虽然改进的能量函数能够有效地处理基数过大造成的大信度焦元被抛弃的情况, 但是当 λ 选取得不恰当时仍会造成焦元删减不合理. 即当焦元基数相差较大时, 若 λ 取值较大, 则大信度焦元仍然存在被抛弃的可能, 但若 λ 取较小的值, 则在保证基数大的大信度焦元保留下来的同时, 基数小的小信度焦元也被保留下来了, 使得焦元删减不彻底, 从而并不能有效地减少计算量.

例 1 假设识别框架为 $\Theta = \{a, b, c, d, e, f, g, h, i, j, k\}$, 某一证据的基本概率赋值为

$$\begin{aligned} m(a) &= 0.12, \quad m(b) = 0.12, \quad m(abcdef) = 0.18, \\ m(acdefg) &= 0.18, \quad m(abcdefghij) = 0.10, \\ m(abcdefghik) &= 0.30, \end{aligned}$$

则计算焦元的能量函数值和平均能量函数值如表1所示.

表1 焦元的能量函数值和平均能量

λ	焦元的能量函数值						平均能量
	a	b	$abcdef$	$acdefg$	$abcdefghij$	$abcdefghik$	
0.4	0.085 7	0.085 7	0.052 9	0.052 9	0.020 0	0.060 0	0.059 5
0.1	0.109 1	0.109 1	0.112 5	0.112 5	0.050 0	0.150 0	0.107 2

为了解决这个问题,本文将在文献[10]的基础上进行改进.通过在删减焦元前先对焦元进行处理,将基数过大的焦元分解,然后再利用改进的焦元能量函数删减焦元,具体步骤如下:

Step 1 给定 m 为 Θ 上的证据,分解焦元.将焦元按基数的大小进行排序,设定阈值 $k = \lfloor |\Theta|/2 \rfloor + 1$,然后把焦元所含命题的个数大于 k 的视为大焦元,并将 m 的大焦元按

$$\begin{cases} m'(\theta_i) = m(\theta_i) + \sum_{\theta_i \in B, B \subset \Theta, |B| > k} m(B)/|B|, \\ i = 1, 2, 3, \dots, |B|, \\ m'(A) = m(A), |A| < k, |A| \neq 1 \end{cases}$$

进行分解,从而把证据的焦元控制在基数小于 k 的范围内;

Step 2 由(3)式和(4)式分别计算 m' 的焦元能量函数和平均能量函数,并将焦元分为保留焦元 R_j 和抛弃焦元 D_k ;

Step 3 对满足条件

$$\begin{cases} D_k \cap R_j \neq \emptyset, R_j \not\subseteq D_k, R_j \cap D_k \neq \Theta, \\ D_k \subseteq R_j, j = 1, 2, \dots, N_r \end{cases} \quad (5)$$

的抛弃焦元按

$$\begin{cases} M_{kj} = m'(D_k) c_{kj}/N_k, j = 1, 2, \dots, N_r, \\ M_{kj}(\Theta) = 1 - \sum_{j=1}^{N_r} M_{kj} \end{cases}$$

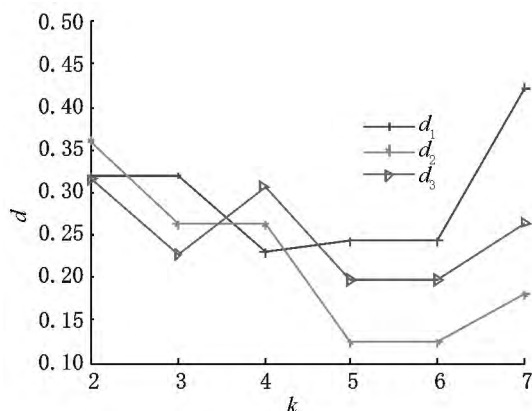


图1 当 $|\Theta| = 8$ 时 m' 与 m 的距离

由图1和图2可知,其中 d_1, d_2, d_3 分别表示3个证据的逼近结果 m' 与 m 的距离,而且当 $k =$

进行提取,其中 N_k 表示保留焦元中与抛弃焦元 D_k 有关并满足(5)式的焦元的个数, N_r 表示保留焦元的个数, c_{kj} 为保留焦元和抛弃焦元相交的程度与抛弃焦元之间的比值,即 $c_{kj} = |D_k \cap R_j|/|D_k|, j = 1, 2, \dots, N_r$;

Step 4 将提取的信度进行分配,得到 m' 的逼近 m'' 为

$$\begin{cases} m''(R_j) = m'(R_j) + \sum_{k=1}^{N_d} M_{kj}, \\ m''(\Theta) = 1 - \sum_{j=1}^{N_r} m''(R_j), \end{cases}$$

其中 N_d 表示证据 m' 中抛弃焦元的个数;

Step 5 利用证据合成规则(1)式对修正后的证据进行融合.

4 算例分析

例2 在 Matlab 环境下,分别随机生成 $|\Theta| = 8$ 和 $|\Theta| = 10$ 的3个证据.这里定义逼近结果 m' 与 m 的距离为 $d = \sqrt{(m - m')^2}/\sqrt{2}$, d 值越小则逼近效果越好.因此当阈值 k 取不同值时, m' 与 m 的距离如图1和图2所示.

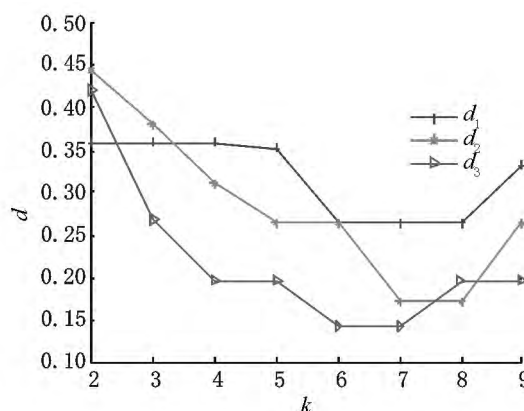


图2 当 $|\Theta| = 10$ 时 m' 与 m 的距离

$\lfloor |\Theta|/2 \rfloor + 1$ 时,逼近结果 m' 与 m 的距离相对较小,因此,本文 k 的取法是可行的.

例3 假设识别框架 $\Theta = \{a, b, c, d, e\}$, m_1, m_2, m_3 为该识别框架下的 3 个证据,其基本概率赋值如表 2 和表 3 所示. 通过将 m_1, m_2, m_3 直接用 D-S 合成与先将 m_1, m_2, m_3 按本文方法作逼近再按 D-S 合成作比较分析(见表 4),其中 $k = 3, \lambda = 0.5$.

由实验结果可知本文方法最终的结果和传统的 D-S 合成算法一样,都是命题 e 的可能性最大,并且其基本概率赋值的提高更有利于最终的决策. 另外,从计算速度上看,本文方法也快于 D-S 合成算法,有

表 2 基本概率赋值

焦元	a	a, b	c, d	d, e	a, b, c, d	a, c, d, e	e
m_1	0.10	0.50	0.05	0.05	0.30	0	0
m_2	0.20	0	0	0.10	0	0.40	0.30
m_3	0	0.05	0	0	0	0	0.95

表 3 修正后的剩余焦元的基本概率赋值

焦元	a	c	d	e	a, b	c, d	d, e	Θ
m_1	0.175	0	0	0	0.575	0	0	0.250
m_2	0.300	0	0	0.450	0	0	0	0.250
m_3	0	0	0	0.950	0	0	0	0.050

表 4 融合结果

方法	合成结果	相乘次数
D-S 合成规则	$m(a) = 0.420, m(e) = 0.580$	30 次
本文方法	$m(a) = 0.086, m(e) = 0.862$ $m(ab) = 0.036, m(\Theta) = 0.015$	17 次

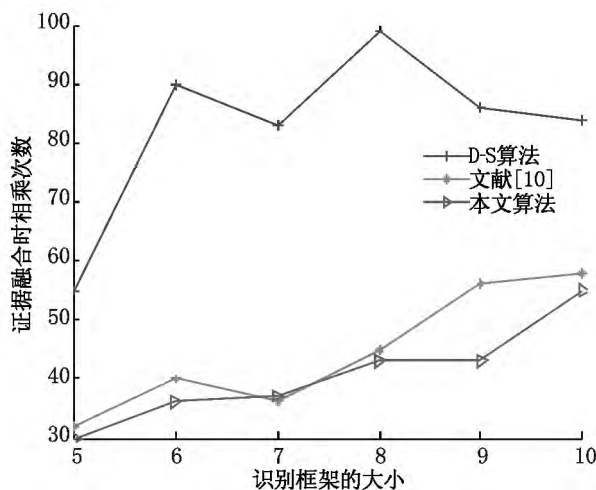


图 3 在融合证据上的相乘次数对比

5 结束语

D-S 证据理论是多传感器信息融合中常用的算法,但其计算复杂度也是人们一直重视的一个问题. 本文算法通过合理地删减证据焦元,不仅有效地减少证据合成过程中的计算量,而且不影响合成结果

效地减少了在证据合成时的计算量,并且随着识别框架中元素的增多,这种优势更加明显.

接下来,在 Matlab 环境下,随着识别框架中元素的增多,随机生成 3 个证据,分别计算 D-S 算法、文献[10]和本文方法在融合证据上的相乘次数,由图 3 可以发现随着识别框架中元素的增多,本文方法不仅明显快于 D-S 算法,而且优于文献[10]的方法,因此本文方法是行之有效的.

的准确性,使得 D-S 证据理论在实际中得到更广泛的应用.

6 参考文献

- [1] Dempster A P. Upper and lower probabilities induced by a multivalued mapping [J]. Annals of Mathematical Statistics, 1967, 38(2): 325-339.
- [2] Gordon J, Shortliffe E H. A method of Managing evidential reasoning in a hierarchical hypothesis space [J]. Artificial Intelligence, 1985, 26(3): 323-357.
- [3] 苗壮,程咏梅,潘泉,等. 快速 mass 函数收敛 [J]. 湖南大学学报: 自然科学版, 2011, 38(1): 89-92.
- [4] 李新德, Dezert J, 黄心汉, 等. 一种快速分层递阶 DSmt 近似推理融合方法 [J]. 电子学报, 2010, 38(11): 2566-2572.
- [5] 郭强, 何友, 李新德. 一种快速 DSmt-DS 近似推理融合方法 [J]. 电子与信息学报, 2015, 37(9): 2040-2046.
- [6] Lowrance J D, Garvey T D, Strat T M. A framework for evidential-reasoning system [C]//Kehler Int, Rosenschein S, Filman R, et al. Proceedings of the 5th National Conference of the American Association for Artificial Intelligence (AAAI-86), 1986(2): 896-903.
- [7] Harmanec D. Faithful approximations of belief functions [C]//Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers, 1999: 271-278.

- [8] Han Deqiang ,Dezert J ,Han Chongzhao. New basic belief assignment approximations based on optimization [C]// 15th International Conference on Information Fusion. Singapore: IEEE ,2012: 286-293.
- [9] 叶清 ,吴晓平 ,翟定军. 一种基于能量函数的证据合成算法 [J]. 系统工程与电子技术 ,2010 ,32(3) : 566-569.
- [10] 金宏斌 ,蓝江桥 ,李鸿飞. 基于改进能量函数的证据推理理论近似计算方法 [J]. 空军雷达学院学报 ,2012 ,26(6) : 427-430.
- [11] 邓勇 ,王栋 ,李齐 ,等. 一种新的证据冲突分析方法 [J]. 控制理论与应用 ,2011 ,28(6) : 839-844.
- [12] 李建军 ,吴根秀 ,张冬梅. 一种新的冲突证据合成方法 [J]. 江西师范大学学报: 自然科学版 ,2007 ,31(2) : 122-126.
- [13] 蒋雯 ,彭进业 ,邓勇 ,等. 一种新的证据冲突表示方法 [J]. 系统工程与电子技术 ,2010 ,32(3) : 562-565.
- [14] Chebbah M ,Martin A ,Yaghlane B B. Combining partially independent belief functions [J]. Decision Support Systems ,2015 ,73: 37-46.
- [15] 程子成 ,吴根秀 ,宋姝婷. 基于融合信息熵性质的信任函数概率逼近 [J]. 江西师范大学学报: 自然科学版 ,2014 ,38(5) : 534-538.

The Approximation Method of Belief Function Based on Decomposing Large Focal Elements

HUANG Mei ,WU Genxiu* ,LIU Qiuyun ,WU Heng ,MAO Linchuan

(College of Mathematics and Informatics ,Jiangxi Normal University ,Nanchang Jiangxi 330022 ,China)

Abstract: An improved method of belief function based on decomposing large focal elements is proposed for the theory of evidence. First ,by decomposing the focal elements with big cardinality ,the cardinality of focal elements are controlled within no more than k . Then focal elements based on the improved energy function are reduced. So the improved method not only reduces the number of focal elements ,but also controls the cardinality of focal elements ,and optimizes the computational complexity ,and the example analysis shows that the method takes on better effectiveness.

Key words: D-S theory of evidence; basic probability assignment; energy function; the average energy function

(责任编辑: 曾剑锋)