

基于截断型 D-S 的快速证据组合方法¹

王 壮 胡卫东 郁文贤 庄钊文

(国防科技大学电子工程学院 ATR 国家重点实验室 长沙 410073)

摘 要 如何解决 Dempster 组合规则的计算复杂性问题是目前研究的热点。该文从提高决策的准确度出发,对截断型 D-S 算法进行了改进,提出了一种新的近似计算方法;并通过仿真实验证明了,该方法在提高基于基本概率分配函数的决策的准确度方面要明显优于截断型 D-S 算法。

关键词 证据组合方法, 近似计算, 截断型 D-S

中图分类号 TN911.72

1 引 言

在数据融合领域中, D-S 证据理论已经成为处理不确定信息的有效方法。然而,由于 Dempster 组合规则的计算复杂度随识别框架 Θ 的基数 $|\Theta|$ 成指数增长,使得巨大的计算量问题妨碍了 D-S 方法在数据融合领域中更加广泛地应用。为了克服这一缺陷,人们先后从两个主要方面来考虑证据理论中的快速计算问题。在 80 年代,主要是针对特殊的证据组织结构,构造相应的快速算法。例如 Barnett 提出的基本证据函数^[1],文献 [2] 中提出的层次假设空间法以及文献 [3] 使用的改进方法等。从 90 年代初至今,主要是侧重于近似计算方法的研究,即如何减少焦点个数。具有代表性的近似计算方法有以下几种: Voorbraak^[4] 提出了贝叶斯近似方法,用元素赋值代替集合赋值,可使得焦点个数 $\leq |\Theta|$; Dubois 和 Prade^[5] 提出了一致近似方法,通过近似计算后的焦点是嵌套的,且焦点个数 $\leq |\Theta|$; Tessem^[6] 提出了 (k, l, x) 近似方法,通过选择一些具有较大基本概率分配 (BPA) 值的焦点,而选择的个数或由焦点个数 k , l 控制,或由剩余的 BPA 值 x 控制;文献 [7] 对 (k, l, x) 近似方法进行了改进,不仅考虑了焦点 BPA 值的大小,而且考虑了焦点本身的大小。文献 [8] 给出了剩余的 BPA 值的分配方法。从本质上说,这些近似方法都是将 D-S 方法中特有的表达“不确定”和“不知道”的 BPA 值向贝叶斯概率转换,并且追加到相应的单子集焦点的概率中去。它们的区别在于如何将“不确定”和“不知道”的 BPA 值近似变换成“确定”和“知道”的概率的机制。这虽然使计算得到了简化,但也失去了证据理论最吸引人的优点。为此,文献 [9] 提出了截断型 D-S 算法 (Truncated D-S)。该算法既减少了焦点个数,同时又始终保持了 D-S 证据理论特有的表达“不确定”和“不知道”信息的优点。然而,截断型 D-S 算法在对 BPA 值较小的焦点的选择以及剩余的 BPA 值的分配上缺乏合理的理论依据。

本文在借鉴已有近似计算方法优缺点的基础上,对截断型 D-S 算法进行了改进。该方法从提高决策的准确度出发,用 $m(A)/|A|$ 作为选择 BPA 值较小的焦点的标准,同时给出了对 BPA 值和基数分别相等的多个焦点进行选择的准则。此外,还给出了剩余的 BPA 值的分配方法。该方法与截断型 D-S 算法不同,不是简单地将剩余的 BPA 值赋给“未知”项 Θ ,而是考虑剩余的焦点与被保留的焦点之间的包含关系。最后,通过仿真实验对两种算法进行了性能比较,证明了本文所提出的近似算法不仅能提高组合的速度,而且在提高基于基本概率分配函数的决策的准确度方面要明显优于截断型 D-S 算法。

¹ 2001-06-15 收到, 2002-01-28 定稿

2 截断型 D-S 方法

截断型 D-S 算法的基本思想^[9]是: 首先给定要保留的焦元的最大个数 MaxNum, BPA 值的上限 MaxBe 和下限 MinBe, 然后根据 BPA 值的大小来决定其对应的焦元是否应该保留。焦元的取舍遵循如下的“截断准则”:

(1) 若 BPA 值大于 MaxBe, 表明其对应的焦元需要更多的证据来支持, 必须保留该焦元;

(2) 而小于 MinBe 的 BPA 值可以忽略不计, 它所对应的焦元可以删除;

(3) 若在准则 (1) 中被保留的焦元个数少于 MaxNum, 则适当减小 MinBe 以保留那些包含单个元素且 BPA 值大于 MinBe 的焦元, 直到保留的焦元个数达到 MaxNum 为止; 如果保留的焦元个数还未达到 MaxNum, 则继续保留那些包含两个元素且 BPA 值大于 MinBe 的焦元, 直到保留的焦元个数达到 MaxNum 为止, 依次类推。

(4) 被去掉的焦元的 BPA 值将被分配给“未知”项, 而为了确保 $m(\Theta)$ 永远不会达到 0, 则采用如下公式来计算 $m(\Theta)$:

$$m(\Theta) = \text{Maximum} \left[1 - \sum_{i=1}^{\text{MaxNum}} m(A_i), \text{MinIgno} \right] \quad (1)$$

其中中括号中第一个关系式的求和仅计算了被保留的焦元的 BPA 值。如果被去掉的焦元的 BPA 值非常小, 那么 $m(\Theta)$ 将得到一个较小的值“MinIgno”。这是截断型 D-S 算法中最关键的近似, 使得该算法具有了自适应能力, 将不会剥夺随后到来的某些焦元的“生存权利”。若 $m(\Theta) = \text{MinIgno}$, 那么被保留的焦元的 BPA 值将根据下式重新归一化:

$$m'(A_i) = (1 - \text{MinIgno}) \cdot \left[\sum_{i=1}^{\text{MaxNum}} m(A_i) \right]^{-1} \cdot m(A_i) \quad (2)$$

截断型 D-S 算法将已经合成的中间结果和新到的证据进行“截断”后再进行合成。该算法有两个优点: 一是控制了被合成焦元个数, 从而减少了计算量; 二是在证据合成过程中, 始终保持“未知”项的 BPA 值不为 0, 即 $m(\Theta) > 0$ 。这样当新的传感器声明与先前保留的焦元发生冲突时, 算法具有自适应能力。在一个高度对抗的复杂环境下(例如欺骗和干扰等), 自适应能力是非常必要的。因为如果 $m(\Theta) = 0$, 那么组合规则将使新的焦元一律被截断, 也就是说新的证据无法用来更新已有的结果。该算法的最大缺点是不满足结合律, 因而证据合成的顺序对计算结果有影响。

3 截断型 D-S 算法的改进

设计近似 D-S 方法一般基于以下两点: 一是为了更快地计算信任函数、似然函数以及信度函数的组合; 二是为了更好地理解包含在证据中的信息。截断型 D-S 算法虽然减少了焦元个数, 但是它仍然保留了那些对数据融合的结果影响较大的包含多个元素的焦元, 而且始终保持了 D-S 证据理论特有的表达“不确定”和“不知道”信息的优点。从这个意义上说, 它更接近于传统 D-S 方法的“风格”。

然而当焦元较多时, 截断型 D-S 算法在对 BPA 值小于 MaxBe 的焦元的选择以及对剩余 BPA 值的分配缺乏合理的理论依据。我们认为应该以提高基于基本概率分配函数的决策的准确度为出发点来考虑。由于基于信度函数的数据融合进行决策所需要的是 pignistic 概率^[10], 因

此对 BPA 值小于 MaxBe 的焦元的选择和剩余 BPA 值的分配应以产生最小的 pignistic 概率偏差为原则。pignistic 概率分布公式^[11]为

$$P_0(\{x\}) = \sum_{x \in A \subseteq \Theta} \frac{m(A)}{|A|} \quad (3)$$

其中 $|\cdot|$ 表示焦元的基数。从 (3) 式不难看出, 基于 pignistic 概率的决策^[10] 不仅考虑了焦元的 BPA 值的大小, 而且考虑了焦元本身的大小。因此采用 $m(A)/|A|$ 作为对 BPA 值小于 MaxBe 的焦元的选择标准, 保证了基于 pignistic 概率的决策的偏差最小, 符合实际决策需要并为 BPA 值较小的焦元的选择提供了合理的理论依据。而对 BPA 值大于 MaxBe 的焦元仍以 BPA 值作为删除标准, 则又保持了 D-S 证据理论特有的表达“不确定”和“不知道”信息的优点。另外, 我们将被删除焦元的 BPA 值分配给包含它的焦元中基数最小的焦元, 或者与它存在交集的焦元中基数最小的焦元^[8], 同样也蕴含了以 $m(A)/|A|$ 作为决策标准思想, 避免了将剩余的 BPA 值任意分配给“未知”项。

此外当对 BPA 值和基数分别相等的多个焦元进行选择时, 截断型 D-S 算法并没有给出具体的准则, 而任意的选择将可能导致决策的偏差。为此, 我们以信息损失最少为原则来进行选择, 其基本思想是: 从 BPA 值和基数分别相等的焦元中选择那些使得信息距离测度 $I(C, B)$ 达到最小的焦元, 其中 C 和 B 分别为近似后的和原有的信度函数。各种各样的信息或熵测度已经被广泛应用于 D-S 证据理论。这里, 我们采用了一个简单但密切相关的信息距离测度^[12]:

$$I(C, B) = - \sum_{\theta_i \in \Theta} \text{Pl}(\theta_i|C) \ln \text{Pl}(\theta_i|B) \quad (4)$$

其中 Pl 为似然函数。

基于上述分析, 我们对截断型 D-S 进行了改进。改进的算法保留了原有的参数 MaxNum 和 MaxBe , 并根据 BPA 值, $m(A)/|A|$ 和 $I(C, B)$ 的大小来共同决定其对应的焦元是否应该保留。焦元的取舍遵循如下新的“截断准则”:

- (1) BPA 值大于 MaxBe 的焦元被保留;
- (2) 若在准则 (1) 中被保留的焦元个数少于 MaxNum , 则对于 BPA 值小于 MaxBe 的焦元, 按照 $m(A)/|A|$ 由大到小的顺序进行保留, 直至保留的焦元个数达到 MaxNum 为止;
- (3) 若在准则 (1) 或 (2) 中需要从 BPA 值和基数分别相等的 p 个焦元中保留 q 个焦元 ($p > q$), 则根据信息距离测度 $I(C, B)$ 的大小, 选择使得信息损失最小的 q 个焦元;
- (4) 从被保留的焦元中寻找包含被删除的焦元, 若存在, 则将被删除的焦元的 BPA 值均匀地分配给这些焦元中基数最小的; 否则, 从被保留的焦元中寻找属于集合 M_A 的焦元, 集合 M_A 定义为

$$M_A = \{B | |B| \geq |A|, B \cup A \neq \emptyset\} \quad (5)$$

其中 B 为被保留的焦元, A 为被删除的焦元。 $m(A)$ 将在集合 M_A 中基数最小的焦元上进行分配, 分配给这些焦元的大小为

$$D_i = (|B_i \cup A|/n) \cdot q \cdot m(A) \quad (6)$$

其中 $B_i (i = 1, \dots, l)$ 为集合 M_A 中基数最小的焦元, 加权因子 q 和 n 定义如下:

$$q = \left| \bigcup_{i=1}^l B_i \cap A \right| / |A| \quad (7)$$

$$n = \sum_{i=1}^l |B_i \cap A| \quad (8)$$

若 $q < 1$, 即 $m(A)$ 没有完全分配, 则用 $A_{\text{rest}} = A / \bigcup_{i=1}^l B_i$ (即由 A 中不属于 $B_i (i = 1, \dots, l)$ 的元素构成的集合) 代替焦点 A , $(1 - q) \cdot m(A)$ 代替 $m(A)$, $|B_1|$ 代替 $|A|$, 重复进行准则 (4), 直至 $m(A)$ 完全分配给被保留的焦点或者 $M_A = \emptyset$, 此时将剩余的 $m(A)$ 分配给“未知”项 Θ .

(5) 若 $m(\Theta)$ 小于 $\text{MinI}gno$, 则 $m(\Theta) = \text{MinI}gno$, 且被保留的焦点的 BPA 值将根据 (2) 式重新归一化.

4 仿真实验

4.1 测试环境

对所有测试, 识别框架 Θ 的基数都为 32, 即 $|\Theta| = 32$. 对上述两种近似算法, 每次模拟产生的信度函数 m 均包含 8 个焦点 (具体方法见文献 [6]), 并且利用 Dempster 组合规则进行 5 次组合计算, 而这一过程重复进行 1000 次. 每次组合结果都将被近似和评价.

参数 MaxNum , MaxBe , MinBe 和 $\text{MinI}gno$ 按照如下两种情况取值:

(1) $\text{MaxNum} = 8$, $\text{MaxBe} = 0.1$, $\text{MinBe} = 0.001$, $\text{MinI}gno = 0.01$; 此时截断型 D-S 方法简称为 $\text{Trunct}8$, 而改进的截断型 D-S 方法简称为 $\text{Entropy}8$.

(2) $\text{MaxNum} = 30$, $\text{MaxBe} = 0.05$, $\text{MinBe} = 0.0005$, $\text{MinI}gno = 0.005$; 此时截断型 D-S 方法简称为 $\text{Trunct}30$, 而改进的截断型 D-S 方法简称为 $\text{Entropy}30$.

由于是为了比较两种近似算法对决策制定的准确度的影响, 因此误差测量参数都是基于由信度函数 m 推导出来的 pignistic 概率 P_0 (见 (3) 式) 的. 第一个误差测量参数是用来刻画近似算法产生的最大 pignistic 概率偏差的. 设 P_0 为原始 D-S 方法产生的信度函数 m 推导出的 pignistic 概率, 而 P_{app} 为近似算法产生的信度函数 m' 推导出的 pignistic 概率, 则误差测量参数 $\text{Error}1^{[6]}$ 定义如下:

$$\text{Error}1(m') = \max_{A \subseteq \Theta} |P_0(A) - P_{\text{app}}(A)| \quad (9)$$

为了全面反映基于 P_{app} 的决策质量, 还需要引入其他误差测量参数. 设 $x_0, x_{\text{app}} \in \Theta$ 具有以下性质:

$$P_0(\{x_0\}) = \max_{x \in \Theta} P_0(\{x\}) \quad (10)$$

$$P_{\text{app}}(\{x_{\text{app}}\}) = \max_{x \in \Theta} P_{\text{app}}(\{x\}) \quad (11)$$

这意味着 x_0 是基于原始信度函数 m 的决策结果, 即理想的决策结果; 而 x_{app} 是基于近似后的信度函数 m' 的决策结果, 即实际的决策结果. 因而误差测量参数 $\text{Error}2$ 和 $\text{Error}3^{[8]}$ 分别定义如下:

$$\text{Error2}(m') = |\{x | P_{\text{app}}(\{x\}) > P_{\text{app}}(\{x\})\}| \quad (12)$$

$$\text{Error3}(m') = |\{x | P_0(\{x\}) > P_0(\{x_{\text{app}})\})\}| \quad (13)$$

Error2 反映了 x_0 在 P_{app} 中的地位, 而 Error3 反映了 x_{app} 在 P_0 中的地位. Error3 对于从决策制定的角度来评价近似算法尤为重要, 因为它直接反映了决策 x_{app} 的质量.

4.2 测试结果

为了公平比较上述两种近似算法, 仅考虑基于 pignistic 概率的误差测量参数是不够的. 我们从实际应用出发, 用组合结果的焦点个数来作为对近似算法运算时间的估计. 表 1 给出了原始 D-S 方法和参与测试的近似算法的第 5 次组合结果的平均焦点个数、最大焦点个数和最小焦点个数 (我们采用的信度函数模拟产生方法^[6], 保证了每种算法使用的测试数据都是相同的). 从表 1 可以看出, 两种近似算法的组合结果的焦点个数都远远小于原始 D-S 方法, 而改进的截断型 D-S 方法的组合结果的焦点个数略小于截断型 D-S 方法. 由于改进的截断型 D-S 方法在剩余的 BPA 值的分配上要耗费一定时间, 因此总的来说, 两种近似算法在计算复杂度上基本相同.

表 1 近似算法的第 5 次组合结果的焦点个数比较

焦点个数	原始 D-S	Trunct8	Trunct30	Entropy8	Entropy30
平均值	2176.61	24.938	74.743	23.751	60.912
最大值	3859	48	128	40	93
最小值	1105	15	37	15	37

图 1, 图 2 和图 3 分别给出了不同近似算法 5 次融合结果的平均误差测量值 Error1, Error2 和 Error3. 从中不难发现, 在算法参数值相同的情况下, 改进的截断型 D-S 方法的平均误差测量值明显小于截断型 D-S 方法. 作为对误差测量参数的补充, 图 4 给出了不同近似算法 5 次融合结果的正确决策率, 即由近似算法导致的且与原始 D-S 方法相同的决策结果比例. 同样在算法参数值相同的情况下, 改进的截断型 D-S 方法的正确决策率明显高于截断型 D-S 方法. 从图 1 至图 4 中, 可以看出保留的焦点个数越多, 决策误差就越小而正确决策率就越大, 即通过牺牲计算量来提高决策质量. 此外, 随着组合次数的增加, 两种近似算法的决策误差都有所增大, 正确决策率也有所下降.

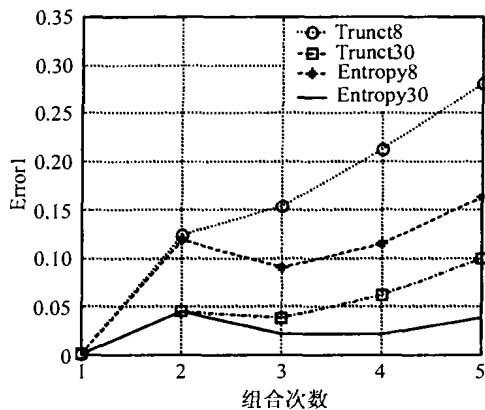


图 1 近似算法的 Error1 值

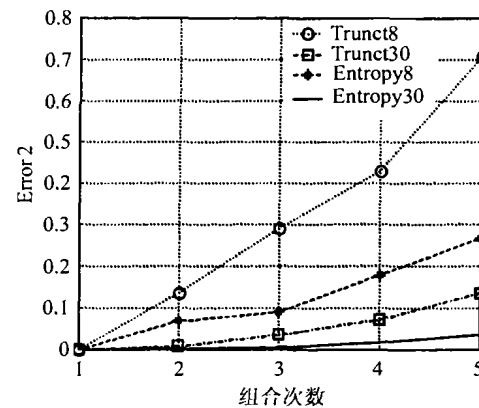


图 2 近似算法的 Error2 值

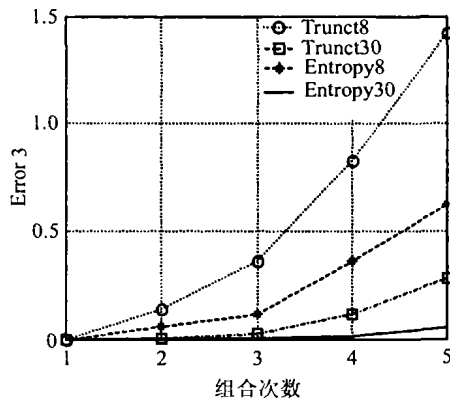


图3 近似算法的 Error3 值

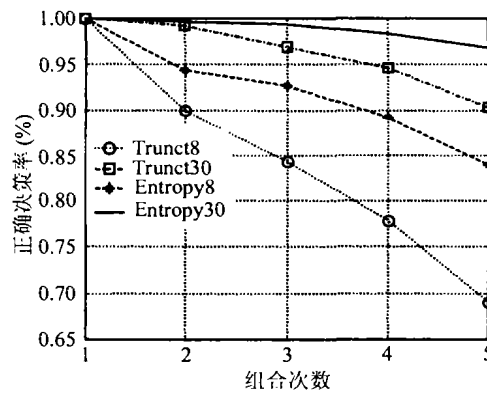


图4 近似算法的正确决策率

5 结论

本文从提高决策的准确度出发,对截断型 D-S 算法进行了改进,用 $m(A)/|A|$ 作为选择 BPA 值较小的焦元的标准,并给出了对 BPA 值和基数分别相等的多个焦元进行选择的标准;此外,新的近似计算方法还给出了剩余的 BPA 值的分配方法.实验结果表明,本算法不仅能提高组合的速度,而且在提高基于基本概率分配函数的决策的准确度方面要明显优于截断型 D-S 算法.当然,对参数 MaxNum, MaxBe, Minlgn0, N 以及信息距离测度 $I(C, B)$ 的选择还需要结合具体的问题来进一步考虑;本算法还需与其他近似计算方法进行全面深入的比较.

参 考 文 献

- [1] J. A. Barnett, Computational method for a mathematical theory of evidence, In: Proc. of Int. Joint Conf. on Artificial Intelligence, Vancouver, BC, Canada, 1981, 868-875.
- [2] J. Gordon, E. H. Shortliffe, A method of managing evidential reasoning in a hierarchical hypothesis space, *Artificial Intelligence*, 1985, 26(3), 323-357.
- [3] G. Shafer, R. Logan, Implementing Dempster's rule for hierarchical evidence, *Artificial Intelligence*, 1987, 33(3), 271-298.
- [4] F. Voorbraak, A computationally efficient approximation of Dempster-Shafer theory, *Int. J. Man-Machine Studies*, 1989, 30(3), 525-536.
- [5] D. Dubois, H. Prade, Consonant approximations of belief functions, *Int. J. Approx. Reasoning*, 1990, 4(2), 419-449.
- [6] B. Tessem, Approximations for efficient computation in the theory of evidence, *Artificial Intelligence*, 1993, 61(2), 315-329.
- [7] 李岳峰, 刘大有, 证据理论中的近似计算方法, *吉林大学学报*, 1995, 31(1), 28-32.
- [8] M. Bauer, Approximation algorithms and decision making in the Dempster-Shafer theory of evidence—an empirical study, *Int. J. Approximate Reasoning*, 1997, 17(2), 217-237.
- [9] P. Valin, *et al.*, Fusion of imaging and non-imaging sensor information for airborne surveillance, *SPIE*, 3719, 1999, 126-138.
- [10] P. Smets, Belief functions versus probability functions, *Uncertainty and Intelligent Systems, Lecture Notes in Comput. Sci.*, 1988, 313, 17-24.
- [11] P. Smets, Constructing the pignistic probability functions in a context of uncertainty, *Uncertainty in Artificial Intelligences*, M. Henrion, R. D. Shachter, L. N. Kanal, and J. F. Lemmer(Editors), Elsevier Science Publishers B. V. (North-Holland), 1990, 29-40.

[12] D. F. Kerridge, Inaccuracy and inference, Proc. Royal Statist. Soc., ser., 1961, 23, 284-294.

A FAST EVIDENTIAL COMBINATION METHOD BASED ON TRUNCATED DEMPSTER-SHAFER

Wang Zhuang Hu Weidong Yu Wenxian Zhuang Zhaowen

(ATR State Key Lab., National Univ. of Defense Technology, Changsha 410073, China)

Abstract How to solve the computation complexity of the Dempster rule is a hot topic at present. In order to improve the accuracy of decision, this paper presents an approximation method which improves truncated D-S algorithm. Simulation results show that this method is superior to truncated D-S algorithm in terms of the accuracy of decision based on basic probability assignment function.

Key words Evidential combination method, Approximation computation, Truncated Dempster-Shafer

王 壮: 男, 1973 年生, 博士生, 研究方向为目标识别与数据融合.

胡卫东: 男, 1967 年生, 博士, 副教授, 主要研究领域为雷达目标识别与数据融合.

郁文贤: 男, 1964 年生, 教授, 博士生导师, 主要研究领域为智能信号处理, 目标识别与数据融合.

庄钊文: 男, 1958 年生, 教授, 博士生导师, 主要研究领域为模糊信息处理, 目标识别与复合制导.