

一种快速分层递阶 DSmT 近似推理融合方法(A)

李新德¹, Jean Dezert², 黄心汉³, 孟正大¹, 吴雪建¹

(1. 东南大学自动化学院复杂工程系统测量与控制教育部重点实验室, 江苏南京 210096;

2. ONERA(The French Aerospace Lab), 29 Av. de la Division, Leclerc, Chatillon 92320, France; 3. 华中科技大学控制系, 湖北武汉 430074)

摘 要: 本文提出了一种分层递阶的 DSmT 快速近似推理融合方法, 该方法针对超幂集空间中仅单子焦元具有信度赋值的情况, 利用二叉树或三叉树分组技术对其刚性分组, 与此同时, 对每个信息源对应的各个分组焦元进行信度赋值求和, 以便实现细粒度超幂集空间向粗粒度超幂集空间映射. 然后运用 DSmT 组合规则和比例冲突分配规则对粗化超幂集空间的两个信息源进行融合, 保存该融合结果作为父子之间节点连接权值, 然后对每个分组焦元信度赋值归一化处理, 通过设定树的深度, 来确定分层递阶的次数. 最后通过从多个角度比较新、老方法, 从而充分地验证了新方法的优越性.

关键词: 近似推理; 信息融合; 分层递阶; Dezert-Smarandache Theory

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2010) 11-0-07

A Fast Approximate Reasoning Method in Hierarchical DSmT(A)

LI Xin-de¹, DEZERT Jean², HUANG Xin-han³, MENG Zheng-da¹, WU Xue-jian¹

(1. Key Laboratory of Measurement and Control of CSE (School of Automation, Southeast University), Ministry of Education, Nanjing,

Jiangsu 210096, China; 2. ONERA (The French Aerospace Lab), 29 Av. de la Division Leclerc, Chatillon 92320, France;

3. Department of Control Science and Control Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China)

Abstract: In this paper, a kind of fast approximate reasoning method in hierarchical DSmT is proposed. This method is only fit for the case that there are only singleton focal elements with assignment in hyper-power set. These focal elements in hyper-power set are forced to group through bintree or tritree technologies. At the same time, the assignments of focal elements in these different groups corresponding to each source are added up respectively, in order to realize the mapping from the refined hyper-power set to the coarsened one. And then, two sources with the coarsened hyper-power set are combined together according to DSmC (Classical DSm combination rule) and PCR5 (Proportional conflict redistribution No. 5). The fused results from different groups will be saved as the connecting weights between father and children nodes. And then, all assignments of focal elements in different groups will be normalized respectively. Tree depth is set, in order to decide the iterative times in hierarchical system. Finally, by comparing new method with old one from different views, the superiority of new one over old one is testified well.

Key words: approximate reasoning; information fusion; hierarchical; Dezert-Smarandache Theory (DSmT)

1 引言

随着计算机科学的发展, 越来越多的信息获取、融合和管理系统要求智能有效地处理复杂的不完善信息 (包括不确定信息、不完全信息、不一致信息和不精确信息, 以及定量和定性信息)^[1], 于是对信息融合的理论方法提出了更高的要求, 传统的方法很难适应这种高要求. Dezert-Smarandache Theory (DSmT) 是由法国的资深科学家 Jean Dezert 博士和美国的著名数学家 Florentin

Smarandache 教授于 2003 年共同提出的一种新的推理理论^[2]. 它是从概率论和 D-S 证据推理理论的基础上发展起来的, 能够有效地解决不确定、不精确、模糊、矛盾或者高度冲突、甚至不完全信息的管理和融合问题. 它既能够解决底层融合问题, 又能解决高层融合问题; 既能处理静态融合问题, 又能处理动态融合问题; 既能处理定量融合问题, 又能处理定性融合问题^[3,4], 以及能够有效地解决不完善信息的融合问题. 目前该理论方法在图像处理、机器人环境感知、军事上的多目标跟踪

与识别、多目标决策、雷达目标分类、地理科学、故障诊断、经济金融、地理信息系统等领域得到了广泛的应用。但是,同 D-S 证据推理理论一样^[5,6],随着鉴别框架中焦元数目的增多,其组合推理运算成指数增长,已成为制约该理论广泛应用与发展的瓶颈问题。

为了解决计算瓶颈问题,很多专家学者在 D-S 框架下进行不少尝试,如 Jean Gordon 和 Edward H. Shortliffe 提出了一种证据组合近似推理方法,这种方法主要分三步来实现,尽管可以避免产生幂集空间的非单子子集和其组合运算的麻烦,但由于第三步仍然需要考虑不同约束情况下,逐步组合不一致信息,随着鉴别框架中的焦元数目的增加,其计算量仍然是比较大的^[7]。Shafer 和 Logan 改进了 Jean Gordon 和 Edward H. Shortliffe 的工作,这是因为当冲突比较高时,用 Jean Gordon 和 Edward H. Shortliffe 的方法,效果不是太好,但 Shafer 和 Logan 的算法不能处理证据 $C_{A_i} \cup \{A_i^c\}$,这里 A_i 表示幂集空间集元素, C_{A_i} 表示 A_i 的子集, A_i^c 表示 A_i 的补集^[8]。Shafer, Shenoy 和 Mellouli 为了改善文献^[8]的缺点,提出了一种定性 Markov 树算法,但同时他们也指出该算法通过减小鉴别框架来降低计算,却导致其最大拆分计算又成指数增长^[9]。Ulla Bergsten 和 Johan Schubert 提出了证据的无环直达图,但由于证据要求具有先后次序,而且必须具有完整的具体路径,约束太强^[10]。Tesseem B. 通过忽略比较小的信度赋值焦元的影响,尽可能地缩减鉴别框架中的焦元个数,但这种近似一方面信息损失严重^[11]。Thierry Denoeux 和 Amel Ben Yaghlane 通过给出不同粒度层级的鉴别焦元,目的是粗化鉴别框,然后通过利用快速的 Möbius 转化算法,产生信任函数的上下边界。这种方法通过粗化鉴别框,能够有效地降低其计算量,且能保证其组合的真实值在一个范围内,但由于不精确信息的进一步处理也非常麻烦,而且需要同时计算上下边界,其计算量也是很大的^[12]。作者没有发现最近几年相关的文献,看上去关于这个课题的研究,由于它的难度,似乎处于停滞状态。

因此,本论文针对上面存在的问题,提出了一种分层递阶的 DSmT 快速近似推理融合方法。该方法的创新之处在于利用二叉树和三叉树分组技术对超幂集空间中单子赋值焦元集合元素进行硬性分组,实现细粒度超幂集空间向粗粒度超幂集空间映射。然后实行递归高效的融合方式,能够快速获得非常可靠的近似融合结果。

2 焦元分组

这里针对仅单子焦元赋值的情况,假设两个信息源 S_1 和 S_2 (鉴别框架相同,即 $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$,其中

$\theta_1, \theta_2, \dots, \theta_n$ 表示鉴别框中的焦元),各个焦元互相排斥,即 $\theta_i \cap \theta_j = \emptyset$ ($i \neq j$),对其超幂集空间 (Hyper-Power Set) D^Θ 进行聚类分组,映射到新的超幂集空间 $\Omega = \{\Theta'_1, \Theta'_2, \Theta'_3, \dots, \Theta'_k\}$,即新、老超幂集空间元素之间存在映射关系 $\rho(\cdot)$,使得 $\rho(\Theta'_k) = \{X_i, X_i \in D^\Theta\}$ 。因此根据文献^[12]的定义 1,也存在一个映射函数 $\varphi(\cdot)$,使得 $\varphi(m(\Theta'_k)) = \sum m(\{X_i, X_i \in D^\Theta\})$ 。这里采用二叉/三叉树分组技术进行单子赋值焦元刚性分组,并对每个信息源的各个分组的单子焦元信度赋值分别求和。由于非单子焦元分组相对复杂,这里暂不予以考虑。

2.1 非零赋值单子焦元分组

假设仅超幂集空间中单子焦元有信度赋值,其它非单子焦元赋值为零,由此可以看出,这里超幂集空间中的单子赋值焦元集合 $S_c \subseteq \Theta$ 。

(1) 采用二叉树的方式分组

首先假设超幂集空间中单子赋值焦元集合为 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$,其中 n 表示集合中单子赋值焦元个数。针对 k 个证据源 S_1, S_2, \dots, S_k 对 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$ 中的单子焦元赋值信度值为矩阵 M ,

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ m_{k1} & m_{k2} & \cdots & m_{kn} \end{bmatrix}$$

若 n 为偶数,将超幂集空间中单子赋值焦元集合 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$ 中前面的 $n/2$ 个焦元聚为一组,后面的 $n/2$ 个聚为另一组,分别对其信度赋值求和得,

$$m_k(\rho_1') = \sum_{i=1}^{\frac{n}{2}} m_{ki}, l = 1 \quad (1)$$

$$m_k(\rho_2') = \sum_{i=(\frac{n}{2}+1)}^n m_{ki}, l = 1 \quad (2)$$

这里下标 k 表示第 k 个证据源,下标 i 表示第 i 个单子焦元, ρ_1', ρ_2' 分别表示第 l 级粗化单子焦元,这里 $\rho_1' = \{\theta_1, \theta_2, \dots, \theta_{\frac{n}{2}}\}$, $\rho_2' = \{\theta_{\frac{n}{2}+1}, \theta_{\frac{n}{2}+2}, \dots, \theta_n\}$;若 n 为奇数,将前面 $[n/2] + 1$ 个焦元聚为一组(这里函数 $[\cdot]$ 表示取最小整数),将后面的 $[n/2]$ 聚为另一组。然后,将归一化后的每个信息源中前后两组焦元的信度赋值分别求和得,

$$m_k(\rho_1') = \sum_{i=1}^{[\frac{n}{2}] + 1} m_{ki}, l = 1 \quad (3)$$

$$m_k(\rho_2') = \sum_{i=[\frac{n}{2}] + 2}^n m_{ki}, l = 1 \quad (4)$$

$$\begin{aligned} \rho_1' &= \{\theta_1, \theta_2, \dots, \theta_{[\frac{n}{2}] + 1}\}, \\ \rho_2' &= \{\theta_{[\frac{n}{2}] + 2}, \theta_{[\frac{n}{2}] + 3}, \dots, \theta_n\}. \end{aligned}$$

然后将第一级分组得到的 θ_1' 和 θ_2' 中的焦元,在分别归一化(详见本文第4节)之后,根据第一级分组原理,依次可得 l 级分组结果,树的深度(l 级)取决于初始超幂集空间中非零单子赋值焦元个数 n ,以及最终分组中焦元的最少保留个数(2或3)。

(2)采用三叉树的方式分组

假设超幂集空间中单子赋值焦元集合为 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$,其中 n 表示集合中单子赋值焦元个数。针对 k 个证据源 S_1, S_2, \dots, S_k ,将 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$ 中的单子焦元分别赋值如矩阵 M 所示。若 n 能被三整除,将超幂集空间中单子赋值焦元集合 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$ 中的焦元分为三组,每组的焦元个数为 $n/3$,于是得到第一级粗化结果 $\theta_1'' = \{\theta_1, \theta_2, \dots, \theta_{n/3}\}$, $\theta_2'' = \{\theta_{n/3+1}, \theta_{n/3+2}, \dots, \theta_{2n/3}\}$, $\theta_3'' = \{\theta_{2n/3+1}, \theta_{2n/3+2}, \dots, \theta_n\}$,然后分别对其信度赋值求和得,

$$m_k(\theta_1'') = \sum_{i=1}^{n/3} m_{ki}, l=1 \quad (5)$$

$$m_k(\theta_2'') = \sum_{i=n/3+1}^{2n/3} m_{ki}, l=1 \quad (6)$$

$$m_k(\theta_3'') = \sum_{i=2n/3+1}^n m_{ki}, l=1 \quad (7)$$

若 n 不能被三整除,首先将前面的 $\lfloor n/3 \rfloor + 1$ 个焦元作为第一组,即 $\theta_1'' = \{\theta_1, \theta_2, \dots, \theta_{\lfloor n/3 \rfloor + 1}\}$,然后把后面 $n - 1 - \lfloor n/3 \rfloor$ 个焦元再次划分,判断 $n - 1 - \lfloor n/3 \rfloor$ 是否是偶数,若是,就将 $(n - 1 - \lfloor n/3 \rfloor)/2$ 个聚为第二组,即 $\theta_2'' = \{\theta_{\lfloor n/3 \rfloor + 2}, \theta_{\lfloor n/3 \rfloor + 3}, \dots, \theta_{n - \lfloor n/3 \rfloor + 1}\}$,剩下的作为第三组,即 $\theta_3'' = \{\theta_{n - \lfloor n/3 \rfloor + 2}, \theta_{n - \lfloor n/3 \rfloor + 3}, \dots, \theta_n\}$;然后分别对其信度赋值求和得,

$$m_k(\theta_1'') = \sum_{i=1}^{\lfloor n/3 \rfloor + 1} m_{ki}, l=1 \quad (8)$$

$$m_k(\theta_2'') = \sum_{i=\lfloor n/3 \rfloor + 2}^{n - \lfloor n/3 \rfloor + 1} m_{ki}, l=1 \quad (9)$$

$$m_k(\theta_3'') = \sum_{i=n - \lfloor n/3 \rfloor + 2}^n m_{ki}, l=1 \quad (10)$$

若 $n - 1 - \lfloor n/3 \rfloor$ 为奇数,将前面 $(n - 1 - \lfloor n/3 \rfloor)/2 + 1$ 个焦元聚为第二组,即 $\theta_2'' = \{\theta_{\lfloor n/3 \rfloor + 2}, \theta_{\lfloor n/3 \rfloor + 3}, \dots, \theta_{\lfloor n/3 \rfloor + 2 + \lfloor \frac{n-1-\lfloor n/3 \rfloor}{2} \rfloor}\}$,将剩下的聚为第三组,即 $\theta_3'' = \{\theta_{\lfloor n/3 \rfloor + 3 + \lfloor \frac{n-1-\lfloor n/3 \rfloor}{2} \rfloor}, \theta_{\lfloor n/3 \rfloor + 4 + \lfloor \frac{n-1-\lfloor n/3 \rfloor}{2} \rfloor}, \dots, \theta_n\}$ 。然后分别对其信度赋值求和,其中 $m_k(\theta_1'')$ 同公式(8),其它见式(11)和(12)

$$m_k(\theta_2'') = \sum_{i=\lfloor n/3 \rfloor + 2}^{\lfloor n/3 \rfloor + 2 + \lfloor \frac{n-1-\lfloor n/3 \rfloor}{2} \rfloor} m_{ki}, l=1 \quad (11)$$

$$m_k(\theta_3'') = \sum_{i=\lfloor n/3 \rfloor + 3 + \lfloor \frac{n-1-\lfloor n/3 \rfloor}{2} \rfloor}^n m_{ki}, l=1 \quad (12)$$

然后将第一级分组得到的 θ_1' , θ_2' 和 θ_3' 中的焦元,在分别归一化(详见本文第4节)之后,根据第一级分组原理,依次可得 l 级分组结果,树的深度(l 级)取决于初始超幂集空间中非零单子赋值焦元个数 n ,以及最终分组中焦元的最少保留个数(2或3)。

2.2 部分零赋值的单子焦元分组

假设超幂集空间中单子赋值焦元集合为 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$,其中 n 表示集合中单子赋值焦元个数。 k 个证据源 S_1, S_2, \dots, S_k 将 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$ 中的单子焦元赋信度值为矩阵 M 。如果矩阵 M 中某一个或者几个行向量中有两个及两个以上单子焦元被赋值为零,但每个列向量中至少有一个元素不为零,那么,首先,将超幂集空间中(信度质量矩阵 M)某些被赋值为零的元素分别对应的列向量聚为一组 z_1 。具体算法如下:假设 M 中任意三行 i, j, h 包含两个及两个以上赋值为零的元素(可以扩展到 $k-1$ 行),其中第 i 行存在焦元为零赋值的为 $m_{i\xi_1}, m_{i\xi_2}, \dots, m_{i\xi_p}, \dots, m_{i\xi_q}, \xi_1 \neq \xi_2 \neq \dots \neq \xi_p \neq \dots \neq \xi_q \in n$,第 j 行存在焦元为零赋值的为 $m_{j\gamma_1}, m_{j\gamma_2}, \dots, m_{j\gamma_p}, \dots, m_{j\gamma_q}, \gamma_1 \neq \gamma_2 \neq \dots \neq \gamma_p \neq \dots \neq \gamma_q \in n$,第 h 行存在焦元为零赋值的为 $m_{h\omega_1}, m_{h\omega_2}, \dots, m_{h\omega_p}, \dots, m_{h\omega_q}, \omega_1 \neq \omega_2 \neq \dots \neq \omega_p \neq \dots \neq \omega_q \in n$,这里 $i \neq j \neq h \in k$,如果 $\xi_m \neq \gamma_n \neq \omega_e \in n$,那么把 ξ, ω, γ 分别对应的列向量聚为一组 z_1 ,得到该分组各列信度赋值的平均赋值为

$$m_\xi = \frac{\sum_{i=1}^k m_{i\xi}}{k}, \xi = \{\xi_1, \xi_2, \dots, \xi_p, \dots, \xi_q\} \quad (13)$$

$$m_\gamma = \frac{\sum_{i=1}^k m_{i\gamma}}{k}, \gamma = \{\gamma_1, \gamma_2, \dots, \gamma_p, \dots, \gamma_q\} \quad (14)$$

$$m_\omega = \frac{\sum_{i=1}^k m_{i\omega}}{k}, \omega = \{\omega_1, \omega_2, \dots, \omega_p, \dots, \omega_q\} \quad (15)$$

如果 ξ, ω, γ 中有重复的,合并相同的列。将 m_ξ ,

m_γ 和 m_ω 三者之和,即 $\sum_{i=1}^{z_1} m_i = m_\xi + m_\gamma + m_\omega$ 作为该分组 z_1 的总权重。然后,将其它各列的元素划为一组 z_2 ,其所占的总权重为 $1 - \sum_{i=1}^{z_1} m_i$,再利用二叉树或者三叉树对 z_2 进行分组处理,接着对该组数据 k 个证据源分别进行归一化处理(详见第4节),然后利用DSmt + PCR5融合规则(详见第3节)进行融合,把得到的融

合结果乘以 $(1 - \sum_{i=1}^n m_i)$.

3 DSmT + PCR5 融合

法国的 Jean Dezert 博士和美国的 Florentin Smarandache 教授在文献[2]提出了两个及两个以上信息源的组合规则和 PCR5(第 5 种比例冲突分配规则),这里简单介绍如下:

当在经典 DSmT 模型下处理信息融合问题时, $Bel_1(\cdot)$ 和 $Bel_2(\cdot)$ 分别为同一鉴别框 Θ 下两个(多个请参见文献[2])独立证据源 S_1, S_2 的信任函数,与之相关联的广义基本信度赋值分别为 $m_1(\cdot)$ 和 $m_2(\cdot)$,其组合规则为:

$$\forall A \in D^\Theta, m_{M(\Theta)}^f(A) \equiv m(A) = \sum_{\substack{X_i, X_j \in D^\Theta \\ X_i \cap X_j = A}} m_1(X_i) m_2(X_j) \quad (16)$$

由于超幂集 D^Θ 在 \cup 和 \cap 集算子下封闭,表达式(16)给出的经典组合规则能够保证融合后的信度赋值 $m(\cdot)$ 恰好是一个广义的基本信度赋值,也就是说: $m(\cdot): D^\Theta \mapsto [0, 1]$. 这里 $m_{M(\Theta)}^f(\emptyset)$ 在封闭空间都假设其恒为零,除非在开放空间可以规定其不为零.

PCR5 考虑到冲突的规范形式,把部分冲突质量分配到卷入冲突的所有元素上.从数学意义上讲,它是目前最精确的冲突质量重新分配规则. PCR5 也满足 VBA 的中立属性,其两源(多源请参见文献[2])的重新分配规则如下:

当 $k=2$ 时, $\forall X \in D^\Theta \setminus \{\emptyset\}$,

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in D^\Theta \setminus X \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)} \right] \quad (17)$$

式(16)~(17)均来自于文献[2],其中卷入的所有元素都是规范形式, $m_{12}(\cdot)$ 和 $m_{12 \dots k}(\cdot)$ 分别对应着两个和两个以上证据源合取一致组合结果.

4 归一化处理

由于初始超幂集空间中所有赋值单子焦元信度赋值之和为 1,通过二叉树或者三叉树焦元聚类分组,分组后的焦元信度赋值之和不为 1,因此为了分层递阶地运用 DSmT 组合规则和 PCR5 冲突重新分配规则,这里需要对分组后的焦元进行信度赋值归一化处理.

假设经过二叉树分组之后,获得粗化的焦元 $\theta_{1'}$ 和 $\theta_{2'}$,如第一级分组,即 $l=1$ 时,若 n 为偶数, ${}_1\theta_{1'} = \{\theta_1, \theta_2, \dots, \theta_{\frac{n}{2}}\}$, ${}_1\theta_{2'} = \{\theta_{\frac{n}{2}+1}, \theta_{\frac{n}{2}+2}, \dots, \theta_n\}$; 若 n

为奇数, ${}_1\theta_{1'} = \{\theta_1, \theta_2, \dots, \theta_{[\frac{n}{2}]+1}\}$, ${}_1\theta_{2'} = \{\theta_{[\frac{n}{2}]+2}, \theta_{[\frac{n}{2}]+3}, \dots, \theta_n\}$,然后根据公式(1)~(4),分别获得各级粗化焦元的信度赋值 $m_k({}_l\theta_{1'})$ 和 $m_k({}_l\theta_{2'})$,接着对粗化焦元 $\theta_{1'}$ 和 $\theta_{2'}$ 中包含单子的信度赋值归一化处理,其递推公式如下:

$$m_{\theta_{1,2,k}'}(\theta_i) = \frac{m_k(\theta_i)}{m_k({}_l\theta_{1,2})} \quad (18)$$

公式(18)中的下标 k 表示证据源的个数, $\theta_{1,2}'$ 表示二叉树第 l 级分组后得到的两个粗化焦元 $\theta_{1'}$ 和 $\theta_{2'}$ 的结果, $m_k(\theta_i) = m_k(\theta_i)$,其左下标 l 决定 $\theta_{1,2}'$ 中 θ_i 的个数.

对于三叉树各级分组,以及具有部分零赋值的单子焦元分组,其归一化处理的基本原理同二叉树是一致的.

5 程序实现

分层递阶 DSmT 近似推理融合程序流程图如图 1 所示,其主要步骤介绍如下:

(1)首先判断超幂集空间中的单子焦元个数 n 是否大于 3.若是,则转入第二步;否则,转入第四步.

(2)判断是否有超过两个的零赋值焦元,若是,将所有赋值为零的单子焦元归为一组,根据部分零赋值的单子焦元分组融合处理方法进行处理;否则,转入第三步.

(3)对焦元进行二叉树或者三叉树分组,并统计归一化的每个信息源中各个分组的焦元信度赋值之和,把该和作为粗粒度焦元的信度赋值,然后转入下一步.

(4)利用 DSmT 和 PCR5 进行粗粒度信息的融合,并

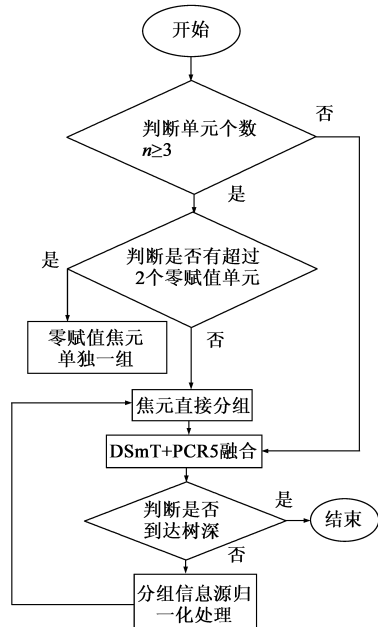


图1 分层递阶DSmT近似推理融合程序流程图

将融合结果作为父子之间节点的连接权值,然后转入下一步。

(5)判断是否到达树的深度,若是,计算超幂集空间中的每个单子焦元 $m(\theta_i)$,并结束程序.例如:若二叉树,如图 2 所示,焦元 θ_1 的信度赋值 $m(\theta_1) = m_{11} * m_{211} * m_{311} * m_{411}$,若二叉树,如图 3 所示, $m(\theta_1) = m_{11} * m_{211} * m_{311}$;否则,转入下一步。

(6)对每个信息源对应的各个分组焦元进行归一化处理,转入第一步。

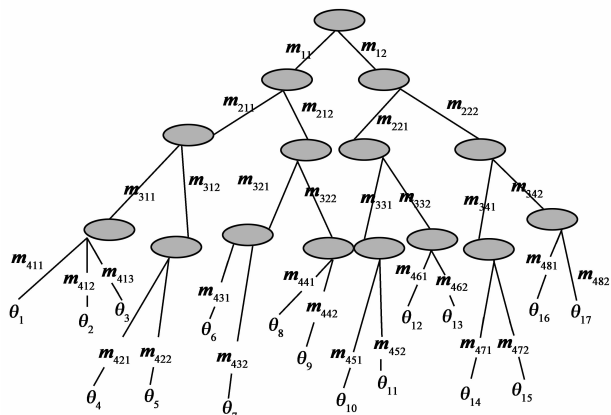


图2 二叉树分层推理融合示意图

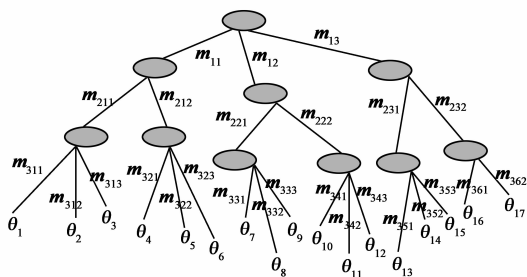


图3 三叉树分层推理融合示意图

6 计算复杂度对比分析

首先来研究一下 Jean Dezert 和 Florentin Smarandache 方法的计算复杂度,这里取两个证据源的情况,超幂集空间中单子赋值焦元集合为 $S_c = \{\theta_1, \theta_2, \dots, \theta_n\}$,包含 n 个单子焦元,通过式(17)计算,可以获得组合后 n 个单子焦元的信度赋值.假设一次乘积运算的复杂度用 K 表示,加运算用 Σ 表示,除运算用 Ψ 表示,总的计算复杂度为

$$O(n) = [K + (4K + 2\Psi + 4\Sigma)(n - 1)] \cdot n \quad (19)$$

显而易见,式(19)中的计算复杂度是关于 n 的线性关系与平方关系 n^2 两者之和。

接着给出本文提出方法的计算复杂度,这里采用算法递归分析的方法,为了描述问题简单,假设 n 是 2 的 q 次幂(如果 n 为大于 2 的 q 次幂,小于 2 的 $q + 1$ 次幂,那么其计算复杂度在 q 次幂和 $q + 1$ 次幂之间),那么树深

$l = q - 1$,忽略判断语句计算的复杂度(其对该算法计算复杂度影响很小).那么得 $l + 1$ 级计算复杂度为:

$$(2(\log_2^n - 2)n + 4)\Sigma + (n - 1)[10K + 4\Psi + 8\Sigma] + 2(\log_2^n - 1)n\Psi + n(\log_2^n - 1)K \quad (20)$$

从式(20)可知,本文提出方法的计算复杂度主要是关于 $n\log_2^n$ 的线性函数,因此同老方法相比,其计算效率得到了显著的提高。

7 融合结果的对比分析

为了说明新方法的优点,这里通过从四个方面进行与老方法进行对比分析,即融合结果的相似性,方法的高效性、冲突敏感性以及鲁棒性。

(1)相似性 假设 $S_c = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}\}$,对于两个信息源 S_1 和 S_2 分别进行信度赋值如表 1 所示,为了充分验证算法的相似性,在保证 S_2 中各个焦元及其信度赋值不变的情况下, S_1 中各个焦元位置不变,但其信度赋值依次向后移,即从表 1 中的初始焦元赋值情况可知, S_1 中 θ_1 的信度赋值为 0.1,对应 S_2 中的 0.05;后移一次之后, S_1 中 θ_1 的信度赋值为 0.04, S_2 中 θ_1 的信度赋值为 0.05, S_1 中 θ_2 的信度赋值为 0.1, S_2 中 θ_2 的信度赋值为 0.21,可以获得 10 种不同的焦元信度赋值情况,通过采用新、老方法进行融合,并且为了比较两个证据源之间新、老方法融合结果之间的相似性,这里通过 Euclidean 相似度函数 $N_E(m_1, m_2)$ (详见文献[13],其公式在(21))来描述。

$$N_E(m_1, m_2) = 1 - \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{D^0} (m_1(X_i) - m_2(X_i))^2} \quad (21)$$

通过计算,最低相似度为 0.9348,最高为 0.9758.因此可以得出高相似性的结论,即新方法的结果是完全可以信赖的。

表 1 两证据源对 S_c 的初始焦元赋值情况

焦元序号		1	2	3	4	5	6	7	8	9	10
信源	S1	0.1	0.3	0.03	0.07	0.2	0.14	0.06	0.02	0.04	0.04
	S2	0.05	0.21	0.04	0.06	0.1	0.24	0.13	0.07	0.06	0.04

(2)高效性 新方法是否能解决 DSmt 运算的瓶颈问题,在保证结果相似度很高的基础上,其高效性指标是至关重要的,尽管从理论分析的角度(详见第 6 节),得出新方法比老方法的效率高的结论,为了充分验证这个结论,下面通过融合两个证据源的具体算例,当超幂集空间中的焦元个数不同时,比较其加、乘、除运算次数以及整体运行时间。

从表 2 中的比较结果看,新方法的高计算效率是显

而易见的,尤其二叉树的效果更为明显,从表 2 中的结果可以进一步分析得到:在同一层其分叉越多,其计算量越大,就降低计算量而言,二叉树是最好的分层方法.

表 2 运行效率比较

超幂集空间中焦元个数	方法	加运算(+)次数	乘运算(×)次数	除运算(÷)次数	运行时间(ms)
10000	老方法	399,953,796	399,963,796	199,976,898	3688
	二叉树	335,344	166,532	82,958	15
	三叉树	260,724	155,583	83,882	62
20000	老方法	1,599,901,648	1,599,921,648	799,950,824	14,672
	二叉树	709,510	340,888	165,714	16
	三叉树	574,778	333,670	178,224	94
30000	老方法	3,599,846,872	3,599,876,872	1,799,923,436	33,625
	二叉树	1,085,018	520,204	244,394	31
	三叉树	843,766	480,924	252,234	469
50000	老方法	9,999,701,168	9,999,751,168	4,999,850,584	93,709
	二叉树	1,950,442	877,918	429,000	47
	三叉树	1,493,368	845,091	437,782	609

(3)冲突敏感性 假设 $S_c = \{a, b, c, d\}$,对于两个信息源 S_1 和 S_2 分别进行信度赋值如下:

$S_1: m_1(a) = x - \epsilon, m_1(b) = \epsilon, m_1(c) = 1 - x - \epsilon,$
 $m_1(d) = \epsilon;$
 $S_2: m_2(a) = \epsilon, m_2(b) = y - \epsilon, m_2(c) = \epsilon,$
 $m_2(d) = 1 - y - \epsilon.$

这里作者通过二叉树方式进行焦元聚类,假设 $\epsilon = 0.01$,为了保证每个焦元的信度赋值大于零,设 $x, y \in [0.02, 0.98]$,为了比较新老方法所得结果的相似性,根据 Euclidean 证据支持贴近度函数(21),当 x, y 分别在 $[0.02, 0.98]$ 变化时,其 Euclidean 相似度变化如图 4 所示(其中,绿色部分表示相似度小于 0.75,红色部分表示在 0.75~0.8,蓝色部分表示在 0.8~0.85,黑色部分表示在 0.85~0.9,黄色部分表示在 0.9~1).其中最小相似度是 0.7110,当信度质量矩阵

$M_1 = \begin{bmatrix} 0.1000 & 0.0100 & 0.8800 & 0.0100 \\ 0.0100 & 0.8800 & 0.0100 & 0.1000 \end{bmatrix}$ 或者
 $M_2 = \begin{bmatrix} 0.8800 & 0.0100 & 0.1000 & 0.0100 \\ 0.0100 & 0.1000 & 0.0100 & 0.8800 \end{bmatrix}$

时得到该最小相似度,对于 M_1 ,通过新、旧方法融合的

结果为

$M_{r_1} = \begin{bmatrix} 0.2204 & 0.2796 & 0.2796 & 0.2204 \\ 0.0161 & 0.4839 & 0.4839 & 0.0161 \end{bmatrix};$
对于 M_2 ,通过新、旧方法融合的结果为

$M_{r_2} = \begin{bmatrix} 0.2796 & 0.2204 & 0.2204 & 0.2796 \\ 0.4839 & 0.0161 & 0.0161 & 0.4839 \end{bmatrix},$
 $M_{r_i}, i = 1, 2$ 的第一行为新方法结果,第二行为旧方法

的结果.可见,信息源存在较高的冲突时,对新方法的融合结果具有一定的影响,但与老方法结果之间的相似度依然很高.当然,这种影响视具体情况也不容忽视,从当前作者的研究现状看,合理的排序能有效地降低新方法对冲突的敏感性.另外,通过计算,作者也发现一个重要的规律,即相似度越高,其超幂集空间中信度赋值比较大的焦元与老方法中对应的焦元越一致,即对应一致排序的焦元数目越多(比较容易证明).

(4)鲁棒性 为了验证新方法的鲁棒性,前面的例子都是两证据源的,这里给出多源同步融合的实例,在表 3 中列出五证据源对 $S_c = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}\}$ 的初始焦元赋值情况,当同步改变五个证据源焦元及其赋值的次序时,采用老方法,其结果是不变的;新方法得到的结果,当赋值次序变化时,对新方法的结果稍微有点影响,但仍然保持与老方法结果之间的高相似性(最低相似度为 0.9766,最高为 0.9867),从而体现

表 3 五证据源对 S_c 的初始焦元赋值情况

焦元序号		1	2	3	4	5	6	7	8	9	10
信源	S1	0.1	0.25	0.05	0.1	0.16	0.15	0.07	0.04	0.06	0.02
	S2	0.07	0.3	0.02	0.07	0.18	0.16	0.1	0.05	0.04	0.01
	S3	0.08	0.24	0.02	0.06	0.12	0.2	0.15	0.06	0.04	0.03
	S4	0.12	0.34	0.01	0.05	0.15	0.13	0.1	0.04	0.05	0.01
	S5	0.06	0.28	0.05	0.08	0.14	0.22	0.08	0.03	0.05	0.01

了新方法具有很好的鲁棒性.

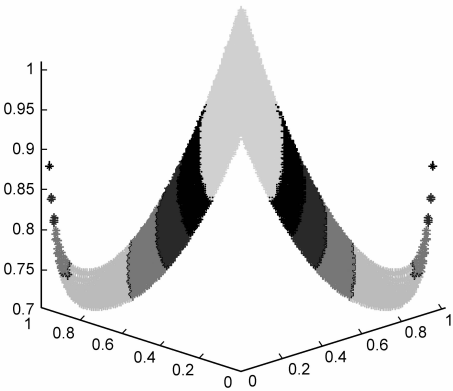


图4 新方法对冲突的敏感性

8 总结与展望

随着 DSMT 在不同领域的广泛应用,其鉴别框架中焦元数目的增多,将导致其运算出现组合爆炸问题.对该问题的解决具有重要的理论价值和应用价值.于是本文提出一种分层递阶的 DSMT 快速近似推理融合方法,该方法能够很好地解决 DSMT 的计算瓶颈问题.该方法目前仅适用超幂集空间中单子焦元具有信度赋值的情况.对于非单子焦元情况,作者也正在开展这方面的研究,请感兴趣的读者关注我们后期的报道.

参考文献:

- [1] 潘泉,于昕,程咏梅,等.信息融合理论的基本方法与进展[J].自动化学报,2003,29(4): 599 – 615.
Pan Quan, YU Xin, Cheng Yongmei, et al. Essential methods and progress of information fusion theory[J]. Acta Automatica Sinica, 2003, 29(4): 599 – 615. (in Chinese)
- [2] Smarandache F, Dezert J. Advances and Applications of DSMT for Information Fusion [M]. American Research Press, Rehoboth, USA, Vol. 1, Vol. 2 and Vol. 3, 2004/2006/2009.
- [3] Li X, Dezert J, Smarandache F, Dai X. Combination of qualitative information based on 2 – Tuple modelings in DSMT[J]. Journal of Computer Science and Technology. 2009, 24(4): 786 – 798.
- [4] Li X, Dai X, Dezert J Smarandache F. Fusion of imprecise qualitative information[J]. Applied Intelligence (In press), 2009, online available, DOI: 10. 1007/s10489 – 009 – 0170 – 2.
- [5] 胡昌华,司小胜,周志杰,王鹏.新的证据冲突衡量标准下的 D-S 改进算法[J].电子学报,2009,37(1): 1578 – 1583.
Hu Changhua, Xi Xiaosheng, Zhou Zhijie, Wang Peng. An improved D-S algorithm under the new measure criteria of evidence conflict[J]. Acta Electronica Sinica, 2009, 37(1): 1578 – 1583. (in Chinese)
- [6] 罗志增,蒋静坪.基于 D-S 理论的多信息融合方法及应用[J].电子学报,1999,27(9): 100 – 102.
Luo Zhizeng, Jiang Jingping. Multisensor data fusion based on dempster shafer evidential reasoning and its application[J]. Acta Electronica Sinica, 1999, 27(9): 100 – 102. (in Chinese)
- [7] Gordon J, Edward H, Shortliffe. A method for managing evidential reasoning in a hierarchical hypothesis space[J]. Artificial Intell., 1985, 26(3): 323 – 357.
- [8] Shafer G, Logan R. Implementing Dempster's rule for hierarchical evidence[J]. Artificial Intelligence., 1987, 33(3): 271 – 298.
- [9] Shafer G, Shenoy P P, Mellouli K. Propagating belief functions in qualitative Markov trees[J]. Int J Approx Reasoning, 1987, 1(4): 349 – 400.
- [10] Bergsten U, Schubert J. Dempster's rule for evidence ordered in a complete directed acyclic graph[J]. International Journal of Approximate Reasoning, 1993, 9: 37 – 73.
- [11] Tessem B. Approximations for efficient computation in the theory of evidence[J]. Artificial Intelligence, 1993, 61: 315 – 329.
- [12] Denux T, Yaghlane A B. Approximating the combination of belief functions using the fast mbius Transform in a coarsened frame[J]. International Journal of Approximate reasoning, 2002, 31: 77 – 101.
- [13] Li X, Dezert J, Smarandache F, Huang X. Evidence supporting measure of similarity for reducing the complexity in information fusion[J]. Information Sciences, 2009 (Accepted).

作者简介:

李新德 男,1975 生于山东烟台,博士,副教授,2007 年博士毕业于华中科技大学控制理论与控制工程专业,目前任教于东南大学自动化学院,主要研究方向是信息融合、不确定推理、机器感知与智能机器人等。 E-mail: xindeli@seu.edu.cn

Jean Dezert 男,1962 年出生于法国 l'Hay les Roses,目前为法国国家航空研究中心资深科学家.其主要研究兴趣包括信息融合、不确定推理、自动导航与多传感器多目标跟踪等.

