

一种基本概率赋值转换为概率的最优化方法

许培达¹, 韩德强², 邓 勇^{1,3}

(1. 上海交通大学电子信息与电气工程学院, 上海 200240; 2. 西安交通大学综合自动化研究所, 陕西西安 710049;
3. 西南大学计算机与信息科学学院, 重庆 400715)

摘 要: 将基本概率赋值转换为概率在不确定性决策中有重要作用, 本文提出了一种基于证据关联系数的最优化转换方法. 主要思想是: 转换后所得到的概率分布可以视为一个特殊的基本概率赋值函数, 从命题的上下概率之间寻优使得被转换的基本概率赋值与转换所得概率之间的证据关联系数最大. 将提出的最优化转换方法与常用的 Pignistic 概率转换方法通过算例进行比较, 表明了所提出方法的有效性.

关键词: 基本概率赋值; 证据关联系数; 最优化; 决策

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2011) 3A-121-05

An Optimal Transformation of Basic Probability Assignment to Probability

XU Pei-da¹, HAN De-qiang², DENG Yong^{1,3}

(1. School of Electronic Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China;
2. Institute of Integrated Automation, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China;
3. College of Computer and Information Sciences, Southwest University, Chongqing 400715, China)

Abstract: One of the key issues in the application of Dempster-Shafer evidence theory is how to make decision based on basic probability assignment (BPA). The usual method is to transform the BPA to probability. In this paper, an optimal transformation method based on a new presented evidence correlation coefficient is proposed. The obtained probability has the maximum correlation coefficient to the original BPA. Numerical examples are used to illustrate the efficiency of the proposed method.

Key words: basic probability assignment; evidence correlation coefficient; optimization; decision-making

1 引言

由于基本概率赋值 (Basic Probability Assignment, BPA) 在命题的幂集上分配, 这个特点使得 BPA 能够比只在单子集命题上分配的概率更好地表示不确定信息, 进而在诸如信息融合、智能信息处理等涉及不确定信息处理的领域得到广泛应用^[1~5]. 一个实际应用存在的问题是: 如何基于 BPA 函数进行决策? 以某机载多传感器数据融合目标识别系统为例, 雷达、ESM 和 IFF 传感器的报告用 BPA 表示和建模, 通过 Dempster 组合规则进行融合, 其融合结果仍然是一个关于目标身份的 BPA 函数, 这就使得系统必须根据 BPA 做出最终的决策以识别目标身份. 受限于相关数学理论研究, 目前直接基于 BPA 进行决策的方法还比较少, 某些系统是根据专家经验设定阈值作为决策依据, 难以推广到其他应用领

域, 同时这种方法主观性较强.

工程上迫切需要研究人员提供一个具有较好数学基础的决策模型, 针对这个问题, 常用的方法是: 将 BPA 转换为概率后再基于概率决策的方法给出最终决策. 这一思路是简单且有效的: 概率论的研究较为充分, 相关决策方法和模型都比较多, 可供工程设计人员参考. 这一思路最为典型的方法是 Smets 提出的 Pignistic 概率转换^[6] (Pignistic Probability Transformation, PPT). PPT 是将多子集命题上的 BPA 根据多子集命题的势均分到单子集命题上, 这是一种类似熵最大化的转换方法, 造成的信息损失较大, 本文提出一种最优化的概率转换方法 (Optimal Probability Transformation, OPT), 在引入本文提出的证据关联系数的基础上, 新方法的基本思想是使得转换后的概率与转换前的 BPA 之间的证据关联系数最大.

收稿日期: 2010-09-28; 修回日期: 2011-03-19

基金项目: 国家自然科学基金 (No. 60874105, No. 60904099); 国家重点基础研究发展计划资助项目 (No. 2007CB311006); 教育部新世纪优秀人才支持计划 (No. NCET-08-0345); 上海市青年科技启明星计划 (No. 09QA1402900); 重庆市自然科学基金 (No. CSCT, 2010BA2003); 航空科学基金 (No. 20090557004, No. 20095153022); 上海交通大学“晨星学者计划”资助 (No. T241460612); 上海交通大学国防预研基金 (No. 11GFF-17); 西南大学博士基金 (No. SWU110021)

2 理论基础

本节将简要回顾 D-S 证据理论^[7]和可传递信度模型 (Transferable Belief Model, TBM) 的一些基本概念^[6].

2.1 D-S 证据理论

定义 1 $\Theta = \{H_1, H_2, \dots, H_N\}$ 是一个 N 个元素的有限集合, 用 $P(\Theta)$ 表示 Θ 中 2^N 个元素的幂集.

$$P(\Theta) = \{\Phi, \{H_1\}, \{H_2\}, \dots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\}, \dots, \Theta\} \quad (1)$$

定义 2 BPA 函数是幂集 $P(\Theta)$ 的映射, 它的大小在 0 到 1 之间.

$$m: P(\Theta) \rightarrow [0, 1] \quad (2)$$

它满足以下条件:

$$m(\Phi) = 0, \sum_{A \subseteq \Theta} m(A) = 1 \quad (3)$$

其中 A 表示幂集 $P(\Theta)$ 中的一个命题, 可以是单子集命题, 也可以为多子集命题, $m(A)$ 表示证据 (BPA) 支持命题 A 的程度, 又称为 mass 函数. 对一个命题, 若其 mass 函数非 0, 则将这个命题称为焦点.

定义 3 信度函数 Bel 定义如下:

$$Bel: P(\Theta) \rightarrow [0, 1] \quad (4)$$

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

定义 4 似真函数 Pl 定义如下:

$$Pl: P(\Theta) \rightarrow [0, 1] \quad (5)$$

$$Pl(A) = 1 - Bel(\bar{A}) = \sum_{B \cap A \neq \Phi} m(B)$$

$Bel(A)$ 和 $Pl(A)$ 分别是命题 A 的概率下限和上限.

定义 5 经典证据理论的 Dempster 组合规则可以将两个 BPA 进行融合, 该规则定义为:

$$m(A) = \frac{\sum m_1(B) m_2(C)}{1 - k} \quad (6)$$

其中

$$k = \sum_{B \cap C = \Phi} m_1(B) m_2(C) \quad (7)$$

k 表征 BPA 之间的冲突程度.

2.2 可传递信度模型

可传递信度模型^[6] (Transferable Belief Model, TBM) 是一个双层结构, 分别是 Credal 层和 Pignistic 层. 信度在 Credal 层进行传递, 当要做出决策时使用 Pignistic Probability Transformation 得到 Pignistic 概率, 将其概率分布记为 $BetP$, 在 TBM 中, Pignistic 层首先由 Credal 层的信度函数构造出一个合理的概率分布, 然后以此概率分布为依据做出最终决策.

Pignistic 概率转换 (PPT) 根据下式进行:

$$BetP(x) = \sum_{x \in A \in P(\Theta)} \frac{m(A)}{|A|} \quad (8)$$

其中, x 为 Θ 中的单个元素, 又称为原子, 仅由一个原子构成的命题称为单子集命题. $|A|$ 表示 A 中所含的原子个数, 如: 若 $A = \{H_1 \cup H_2\}$, 则 $|A| = 2$. 通过上式可将 BPA 转换到单子集命题的概率分布上. $BetP$ 是一个概率分布函数, 称它为 Pignistic 概率分布函数. 通过 Pignistic 变换, 可方便地对不确定情形做出决策. 从 $BetP$ 的计算公式可知, PPT 是将多子集命题对应的信度依据均分的思想平均分配到各单子集命题上.

3 证据关联系数

熵作为事物不确定性的一种度量, 文献[8]将其推广后引入随机变量的偏熵和关联熵, 进而定义了关联系数.

定义 6 设随机变量 X 与 Y 的分布为

$$X: \begin{Bmatrix} a_1 & \cdots & a_K \\ p_1 & \cdots & p_K \end{Bmatrix} \quad Y: \begin{Bmatrix} b_1 & \cdots & b_K \\ q_1 & \cdots & q_K \end{Bmatrix}$$

随机变量 X 关于随机变量 Y 的偏熵定义如下:

$$H_Y(X) = - \sum_{k=1}^K q_k \log p_k = - \log \prod_{k=1}^K p_k^{q_k} \quad (9)$$

将 $H_Y(Y)$ 简记为 $H(Y)$, 偏熵满足性质: $H_Y(X) \geq H(Y)$.

定义 7 随机变量 X 与 Y 之间的关联熵定义为它们的偏熵之和, 即:

$$H(X; Y) = H_Y(X) + H_X(Y) = - \log \prod_{k=1}^K p_k^{q_k} \cdot q_k^{p_k} \quad (10)$$

定义 8 随机变量的偏关联系数与关联系数分别定义为:

$$r_Y(X) = \frac{H(Y)}{H_Y(X)} \quad (11)$$

$$r_X(Y) = \frac{H(X)}{H_X(Y)} \quad (12)$$

$$r(X; Y) = \frac{H(X \otimes Y)}{H(X; Y)} = \frac{H(X) + H(Y)}{H_X(Y) + H_Y(X)} \quad (13)$$

关联系数具有以下性质:

$$0 \leq r_X(Y), r_Y(X), r(X; Y) \leq 1,$$

当且仅当 X, Y 为同分布时,

$$r_X(Y) = r_Y(X) = r(X, Y) = 1.$$

此性质表明关联系数 $r(X; Y)$ 是随机变量 X 与 Y 的分布之同一性, 一致性的特征度量. 式(13)综合考虑了随机变量 X 与 Y 各自的偏关联系数对总的关联系数的影响.

以上述定义为基础, 我们将关联系数应用到信息融合中, 并根据 BPA 的特点, 引入一种新的关联系数——证据关联系数.

用 $a_1, \dots, a_K$ 表示证据中的命题, 用 $p_1, \dots, p_K$ 表示命题所对应的基本概率赋值, 可以将 BPA 表示为:

$$X: \begin{Bmatrix} a_1 & \cdots & a_K \\ p_1 & \cdots & p_K \end{Bmatrix}$$

式(9)中定义的偏熵在 $p_K = 0$ 时会出现结果无穷大的问题, 文献[9]提出使用无穷小正数替代 0 的方法来解决, 但我们发现: 经过如此修改, 在实际应用中仍然会出现问题. 为此我们提出一种类偏熵:

定义 9 随机变量 X 关于随机变量 Y 的类偏熵定义为

$$H_Y(X) = \sum_{k=1}^K q_k \cdot e^{-5p_k} \quad (14)$$

我们使用指数形式来拟合原偏熵中的对数变化, 并通过实验确定了 e^{-5p_k} 的形式. 类偏熵可有效克服无穷小正数参数设置的缺陷, 改进后关联系数计算更符合直观性. 在类偏熵基础上, 利用式(11)~(13), 我们定义其得到的新关联系数为指数型关联系数.

BPA 中各命题并非相互独立, 多子集命题表现出对各单子集命题的支持, 为了体现这种内联关系, 使得计算结果更加合理, 我们先引入文献[10]提出的矩阵 D , 然后再对原 BPA 进行预处理.

将两证据中出现的所有命题集构成命题集合 W , 记其包含的命题个数为 n .

D 为 n 阶方阵, 由 W 中命题两两之间的 $D(\cdot, \cdot)$ 构成. $D(\cdot, \cdot)$ 按下式计算:

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}, A, B \in W \quad (15)$$

其中 A, B 表示命题, $|A \cap B|$ 表示 A 与 B 交集的势, 也就是交集中原子的个数, 同理 $|A \cup B|$ 表示 A 与 B 并集的势. 矩阵 D 用以体现 W 中各命题的相互关系. D 为对称阵, 对角元均为 1.

预处理过程为: 将两证据 BPA 按照 W 的格式构成向量 v_1 和 v_2 , 然后用矩阵 D 分别左乘向量 v_1 和 v_2 , 可得到新的向量如下:

$$v'_1 = D \times v_1, \quad v'_2 = D \times v_2 \quad (16)$$

接下来使用前面的方法就计算预处理之后向量的指数型关联系数, 我们将此时得到的指数型关联系数定义为证据关联系数. 证据关联系数不仅计算稳定, 且其对高度冲突证据能够准确表征.

4 最优化转换方法

现在对于从 BPA 到概率的转换方法还没有统一的评价标准, 目前的研究中有采用信息熵来作为评价方法的, 文献[11]对几种转换方法通过算例对其转换结果进行了信息熵的比较, 我们认为: 信息熵表征的是证据的不确定程度, 是证据携带信息并予以表现的一种形

式, 对于决策者来说, 如果仅用于排序决策, 那么转换结果的信息熵自然越低越好, 因为这样的信息确定性越强, 但在其他非排序决策的情况下, 比如要根据转换结果来对资源进行分配, 达到系统的预定目标, 此时就并非如此. 本文旨在通过转换使得原 BPA 与转换后的概率之间关联性最强, 使结果能较好地重现证据的原始信息.

在此我们提出一种应用最优化思想的概率转换方法. 通过定义 2 可以看出, 证据理论中的 BPA 是在辨识框架命题的幂集上进行分配, 如果分配到单子集, 则 BPA 就退化为一个概率分布. 从这个角度来看, 概率可以看作是一种特殊的 BPA. 因此本文所提出的最优转换是基于这样一个思想: 转换后的概率与转换前的 BPA 之间的冲突应该最小, 也即两者之间的关联系数应该最大, 以求更完整地保存 BPA 本身信息. 设转换前的 BPA 为 m_{input} , 转换后的 BPA 为 m_{output} , 则最优转换的目标函数为:

$$\max r(m_{\text{output}}, m_{\text{input}}) \quad (17)$$

约束条件为: 转换后的 BPA (也即概率结果, 在此视其为一种特殊的 BPA) 中单子集命题得到的概率应该在该命题的上下概率区间 $[Bel(\cdot), Pl(\cdot)]$ 之中. 其中, 下概率为定义 3 中的信度函数, 上概率为定义 4 中的似真函数. 另外, 转换后的 BPA 各命题 mass 值之和应为 1.

实现寻优的过程为: 先初始化, 即设定初始目标 BPA (即转换后的概率结果, 在此将其视为一种特殊的 BPA), 然后将其输入到证据关联系数子函数中处理, 求得证据关联系数, 判断证据关联系数是否达到最大, 若不是最大, 则根据约束条件产生新的目标 BPA, 再将其输入到证据关联系数子函数中处理, 直到得到的证据关联系数最大, 此时的目标 BPA 即为最优解. 以上过程可以通过相关非线性最优化的方法来实现, 我们通过 MATLAB 软件中的 `fmincon` 函数来实现这一寻优.

具体实现过程如下:

(1) 初始化

由于目标 BPA (m_{output}) 仅在单子集命题上有概率赋值, 根据单子集命题的上下概率区间约束, 我们可以求出各单子集命题的信度函数 (即概率下限), 并由它们构成初始目标 BPA, 其余多子集命题在寻优过程中均赋值为 0. 同时目标 BPA 各命题 mass 值之和应为 1.

(2) 证据关联系数子函数

该子函数输入为目标 BPA, 输出为目标 BPA 和原 BPA 的证据关联系数 $r(m_{\text{output}}, m_{\text{input}})$, 具体过程如下:

(a) 将目标 BPA 和原 BPA 中出现的所有命题构成命题集合 W , 然后将原 BPA 和目标 BPA 分别按照 W 的格式写成向量 v_{input} 和 v_{output} ;

(b)根据式(15)生成相应的 D 矩阵;

(c)根据式(16)对 v_{input} 和 v_{output} 进行预处理,得到 v'_{input} 和 v'_{output} ;

(d)根据式(11)~(14)求出 v'_{input} 和 v'_{output} 的关联系数 $r(m_{output}, m_{input})$,具体步骤在第三节已举例介绍;

利用上述的证据关联系数函数和初始化过程产生的初始目标 BPA,再加上前面介绍的两个约束条件,就可以构成寻优问题的所有输入,再结合相应寻优算法或软件(如 Matlab 或 1stopt)即可得到最终的最优 BPA,也就是我们需要的概率转换结果。

5 实验算例

本节通过两个例子来分析本文提出的概率转换方法。

例 1 假设有一 BPA 如下:

$$m(a) = m(b) = m(c) = m(d) = 0.1$$

$$m(\{a, b\}) = m(\{b, c\}) = 0.3$$

使用 PPT 的转换结果为:

$$p(a) = p(c) = 0.2500$$

$$p(b) = 0.4000; p(d) = 0.1000$$

使用本文 OPT 方法的转换结果为:

$$p(a) = p(c) = 0.2277$$

$$p(b) = 0.4445; p(d) = 0.1000$$

本例中,使用 PPT 转换后的概率与原 BPA 之间的证据关联系数是 0.947557,使用 OPT 转换后的概率与原 BPA 之间的证据关联系数是 0.951319。从原 BPA 来看,命题 $\{b\}$ 转换后应该有最大的概率,由于命题 $\{d\}$ 只有 $m(d) = 0.1$ 支持,其概率上下限均为 0.1,故转换结果仍保留为 0.1,OPT 结果较之 PPT 来说,对 $\{b\}$ 的支持更高,对于决策者来说,确定性更强,更有利于决策。

例 2 假设有三套导弹防御系统,但是有四个地方 a, b, c, d 需要保卫。根据专家的评估,各个地方受攻击的可能性如下:

$$m(a) = m(d) = 0.2$$

$$m(b) = 0.1$$

$$m(\{a, c\}) = 0.2$$

$$m(\{a, b, c\}) = 0.3$$

使用 PPT 的转换结果为:

$$p(a) = 0.4000$$

$$p(b) = p(c) = p(d) = 0.2000$$

使用本文 OPT 方法的转换结果为:

$$p(a) = 0.4177$$

$$p(b) = 0.1725$$

$$p(c) = 0.2097$$

$$p(d) = 0.2000$$

本例中,使用 PPT 转换后的概率与原 BPA 之间的

证据关联系数是 0.905361,使用 OPT 转换后的概率与原 BPA 之间的关联系数是 0.907880。从这个例子来看,PPT 无法给出决策排序,而 OPT 给出的优先保卫顺序依次为: $a > c > d > b$,因此从受攻击的可能性角度来看,应该保卫的三个目标是: a, c, d ,这同时也符合我们的直观推测。

通过上述例子我们可以看到,本文方法较之 PPT,转换效果较好,是可行的。

6 结论

模糊数学中在决策阶段通常有去模糊化的过程。类似地,在基于证据理论的应用中,也需要将 BPA 转化为概率。本文基于证据之间的关联系数提出了一种基于最优化思想的转换方法,具有算法实现简单、物理意义清晰的优点,在多传感器数据融合的目标识别决策中应用良好。

参考文献

- [1] 韩崇昭,朱洪艳,等.多源信息融合[M].北京:清华大学出版社,2006.
Han Chong-zhao, Zhu Hong-yan, et al. Multi-Source Information Fusion[M]. Beijing: Tsinghua University Press, 2006. (in Chinese)
- [2] 何友,王国宏,等.传感器信息融合及应用[M].北京:电子工业出版社,2000.
He You, Wang Guo-hong, et al. Multisensor Information Fusion with Applications[M]. Beijing: Electronic Industry Press, 2000. (in Chinese)
- [3] 赖海光,许峰,等.基于 Dempster-Shafer 证据理论的端口扫描检测方法[J].电子学报,2006,34(11):1946-1950.
Lai Hai-guang, Xu Feng, et al. A portscan detection method based on Dempster-Shafer theory of evidence[J]. Acta Electronica Sinica, 2006, 34(11):1946-1950. (in Chinese)
- [4] 邓勇,施文康,等.一种有效处理冲突证据的组合规则[J].红外与毫米波学报,2004,29(1):27-32.
Deng Yong, Shi Wen-kang, et al. Efficient combination approach of conflict evidence[J]. Journal of Infrared and Millimeter Waves, 2004, 29(1):27-32. (in Chinese)
- [5] Guo H W, Shi W K, Deng Yong. Evaluating sensor reliability in classification problems based on evidence theory[J]. IEEE Transactions on System Man Cybernetics, B, 2006, 36(5):970-981.
- [6] Smets P, Kennes R. The transferable belief model[J]. Artificial Intelligence, 1994, 66(3):191-243.
- [7] Shafer G. A Mathematical Theory of Evidence[M]. New Jersey: Princeton University Press, 1976.
- [8] 吴敏金,白治江.关联熵及其应用[J].华东师范大学学报(自然科学版),1998,2(2):28-35.

Wu Min-jin, Bai Zhi-jiang. Relative entropy and its application [J]. Journal of East China Normal University (Natural Science), 1998, 2(2): 28 – 35. (in Chinese)

[9] 郭效芝, 辛小龙. Fuzzy 集的偏熵与关联熵[J]. 模糊系统与数学, 2005, 19(2): 24 – 28.

Guo Xiao-zhi, Xin Xiao-long. Partial entropy and relative entropy of Fuzzy Set[J]. Fuzzy System and Mathematics, 2005, 19(2): 24 – 28. (in Chinese)

[10] Jousselme A L, Grenier D, et al. A new distance between two bodies of evidence[J]. Information Fusion, 2001, 2(2): 91 – 101.

[11] Pan W, Yang H. New methods of transforming belief functions to pignistic probability functions in evidence theory[A]. 2009

International IEEE Workshop on Intelligent Systems and Applications[C]. Wuhan, China: IEEE, 2009. 1 – 5.

作者简介

许培达 男, 1987 年出生, 博士生, 主要研究方向: 信息融合.

韩德强 男, 1980 年出生, 博士, 讲师, 主要研究方向: 信息融合、模式识别、图像处理.

邓 勇 男, 1975 年出生, 博士, 教授, 博士生导师, 上海交通大学晨星奖励计划 SMC 优秀学者, 入选上海市青年科技启明星计划和教育部新世纪优秀人才支持计划, 目前研究方向: 信息融合、智能信息处理. E-mail: doctordengyong@yahoo.com.cn; dengyong@sjtu.edu.cn.