**入门文献感想**

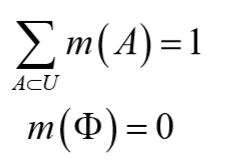
**邓吉祥**

**2019年8月4日**

### 1．DS证据理论

①证据理论主要用于不确定性推理（uncertainty reasoning），而这种不确定性来自于先验概率未知。

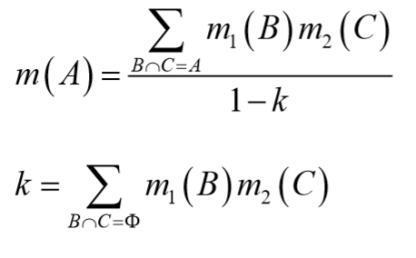
②辩框（frame of discernment）可以理解为对所有可能出现情况（即命题proposition）的罗列，幂集（power set）是对辩框内的所有命题进行组合，所有组合情况的总数是2n（n为命题的个数）。比如{A,B}这个组合，就是说A和B这两个命题都有可能，不确定是哪一个，但我知道就是他们中的一个，所以把信任度分给他们的组合，这也是不确定性的一种体现。

③基本概率分配函数BPA： 

其作用是对辩框的幂集内的命题（比如A）进行信任程度的分配（m(A)），工程中BPA可以由传感器获得并产生。

④对比概率论，证据理论的focal element 可以是多元素集合，也就是不确定的信息，这也是证据理论能区分不知道和不确定的原因。

⑤Bel(A)要考虑到A的所有subset，所以表示对A的总信任程度。而在Pl(A)中，只要与A有交集的集合，就需要考虑，所以表示不否定A的程度。显然Pl(A)考虑的范围更广，所以Pl(A) ≥Bel(A)。

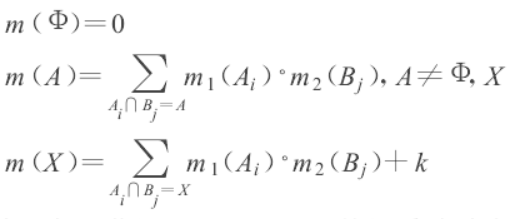
⑥证据的组合（combination rule）：

是归一化（normalization）的系数，k是组合后归入null set 中的信度。DS证据理论认为外界是close world，认为辩框内把所有情况都考虑到了，所以对于组合后产生的，需要归一化使。

### 2．孙全\_一种新的基于证据理论的合成公式

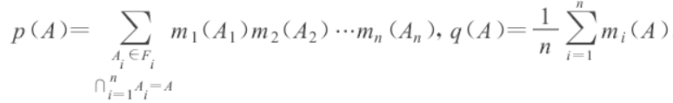
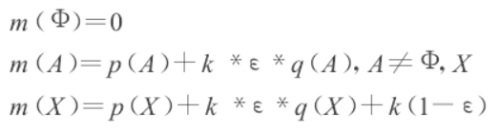
①DS证据理论在融合高度冲突的证据时（k≈1时），会得出有悖常理的结论（counter-intuitive），怎样解决这种悖论是很多论文关心的话题。

②Yager认为这种悖论是因为k≈1时，归一化系数接近无穷，从而无法正常合成。他认为证据存在冲突就无法使用，所以将系数去掉，把k放在m(X)中，即把信度放在了未知领域。

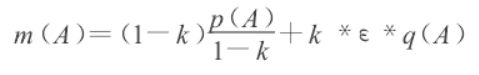


但Yager的方法存在以下问题：一旦算法将多数信度归入未知领域m(X)，即使之后再出现很多支持特定命题（比如A）的信息源，m(X)的信度依然变化不大，这在之后的论文中被称为：disappear of major opinion。所以Yager的方法在信息源较多的情况下，融合的结果不好。

③孙全对Yager的不足进行改进，引入证据的可信度ε。



其中 m(A)可改写为：

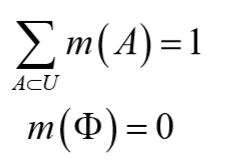


这是一种加权的形式（加权系数是1-k和k），前一项的为经典的DS证据理论合成公式，冲突低时这项占主导（退化为DS证据理论）；后一项的q(A)表示对A的平均支持度，当冲突高的时候，这项起作用。

孙全这么做因为他认为：即使证据之间有冲突，并不是完全不能用，需要用证据可信度ε衡量可用的程度。

在实验环节中：虽然有冲突，但随着支持A的证据源增加，对A的平均支持度q(A)也增加了，并伴随证据可信度ε的增加，所以m(A)也增加。在冲突较高时，这种思路有点类似Murphy C K中提到的Averaging方法，用平均值的方法保留major opinion（多个证据源共同支持的命题）。

### 3．E.lefevre\_Belief function combination and conflict management

①证据在融合的时候，有三种原因会产生冲突，从而产生（conflict mass）：

a. 传感器的异常；

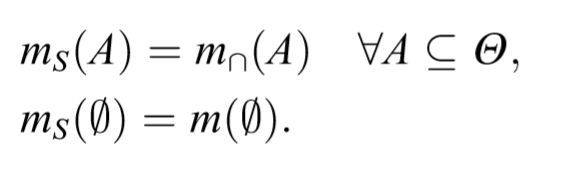
b. 信任函数Bel等模型建立不精确；

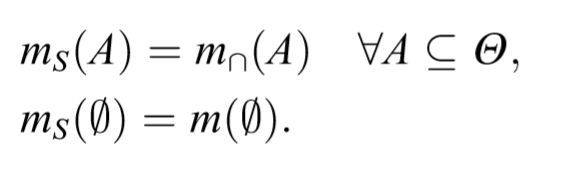
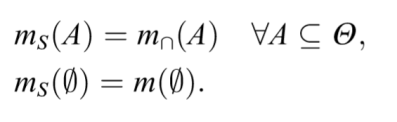
c. 证据源融合过多。

有以下几种解决方法：

②Combination of reliable sources（证据源可靠）

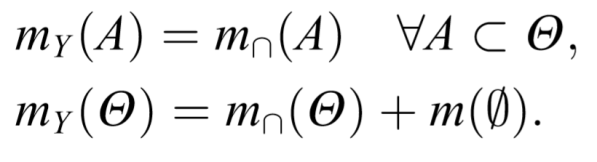
Smets的方法：

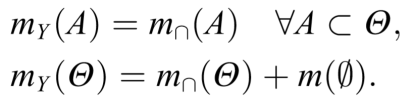
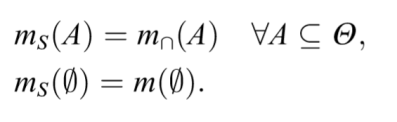
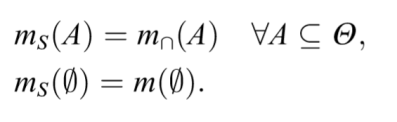
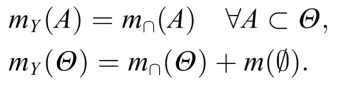


其中就是经典DS证据理论中还未normalization的融合结果。Smets认为冲突仅仅来自于辨框定义的不完整，所以他不做normalization这一步，而是将单独列开。

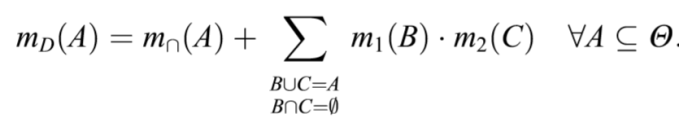
③Combination of non-reliable sources（证据源不可靠）

a. Yager的方法：



表示对情况一无所知，而（有冲突）也是对情况一无所知的表现，所以Yager的做法就是把分给了base set:，即把冲突分给整个集合。

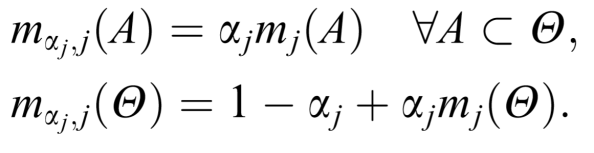
b. Dubois and Prade的方法：



当且时，满足minimum specificity principle，将这种冲突的conflicting mass分给A这个集合。

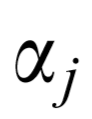
我的理解是（纯属个人理解）：这有点类似于基本的概率论中的相互独立事件【B与C是两事件，若满足等式P(B∩C)=P(BC)=P(B)P(C)，则称事件A,B相互独立】，而B与C的并集为A，所以把这种冲突分给A。

c. 利用折扣系数（discounting coefficients）：

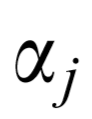


利用折扣系数表示对证据源的信任程度。

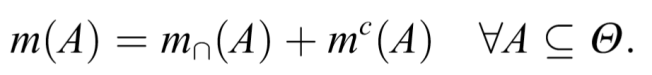
≈1时，表示非常信任这个证据源，该算法变为经典DS证据理论的BPA。

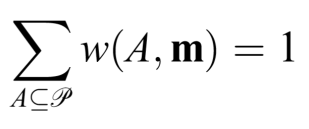
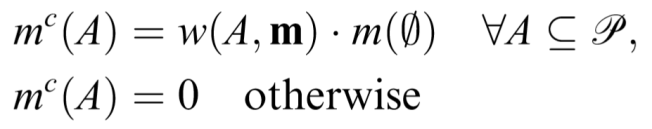


≈0时，表示完全不信任这个证据源，接近于1，这表示对情况一无所知。



④ Lefevre的方法：Generic framework





经典DS证据理论对于的分配，其做法是用统一的归一化系数进行normalization（相当于等比例放大），而Lefevre的方法是：对于不同的焦元，分配不同的冲突比重，这种比重就由权重因子（weighing factors）表示。

⑤求权重因子的方法分3类，共5种（前4种与文中提到的方法相对应）：

a. 用Smets的方法求权重因子（fixing the values of the weighting factors）：

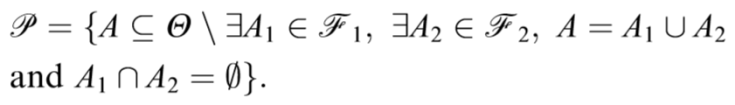
因为Smets的方法把冲突全给了空集，即，所以。

b. 用Yager的方法求权重因子（fixing the values of the weighting factors）：

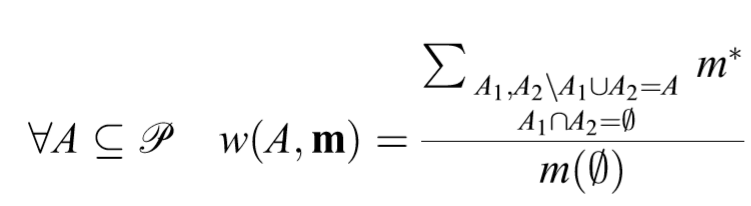
Yager的方法把冲突全给了base set，即，所以=1

c. 用Dubois and Prade的方法求权重因子（computing the weighting factors）：

在Dubois and Prade的方法中，冲突分配的集合为：



所以权重因子为：



其中，这与Dubois and Prade的方法相对应。

d. 用折扣系数求权重因子（computing the weighting factors）：

原文并未介绍折扣系数的求法，只是说折扣系数与权重因子有对应关系（原文附录中），我只是做了大概的了解，知道可以用折扣系数求权重因子。

e. 用自动学习求权重因子（learning the weighting factors）

### 4．Haenni R\_Are alternatives to Dempster's rule real alternatives?

①Haenni R认为用其他的方法对combination进行改进不合适，应该在DS证据理论的基础上寻求更高效的计算方法。

从实践角度出发，他认为其他方法计算量过大，而且有些算法没有交换性和结合性（communitive and associative），这在工程中应用并不实际。

从哲学角度出发，他认为与其寻求新的组合方法，不如把模型（BPA等）进行改进与修正；他同时支持close world的说法，并认为normalization很有必要。

从数学角度出发，他认为算法的复杂度很重要，实践中一般不用信任函数Bel。

②我的感想是：Haenni R大多从实际角度出发，很看重方法的simplicity和practical，这在工程应用（engineering）中是可取的。但是我认为：在理论研究领域（science），可以指出新算法的不足，但不应该无视新算法的优势。有些算法看似复杂难以应用，实际上它能提供新的思路，毕竟历史是螺旋上升的，没有完美的算法。再者，现在随着硬件的计算能力飞速发展，以前因为复杂度很高而没有实际应用的算法，现在也慢慢在实际中得到应用，要以动态发展的眼光看问题。

### 5．Murphy C K\_Combining belief functions when evidence conflicts

①DS证据理论主要的问题：

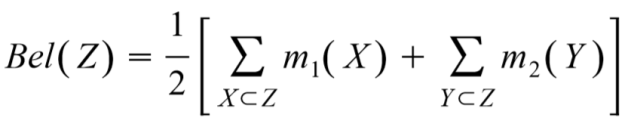
a. 在融合证据时，可能会把100%的信任度给融合前信任度较小的证据。

b. 当一个证据源把所有信任度都给了一个特定的命题时，ignorance interval会消失，也就是说m(Θ)会为0（Θ表示的是一无所知，所以是“ignorance”）。实际上，人们希望算法对m(Θ)做一定的保留（unassigned belief）。

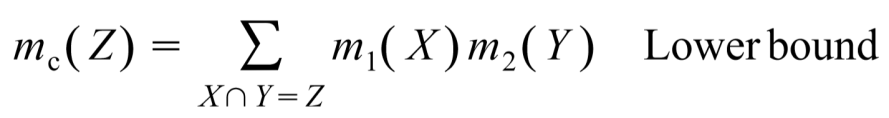
c. 基数大的集合（composed set）会获得的不成比例信任度。

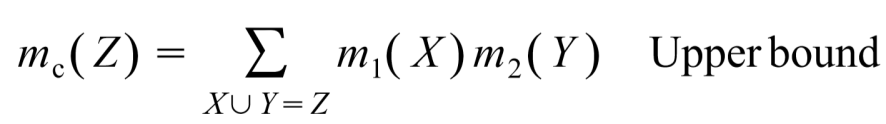
②可行的方法：

a. 对两个证据源中Z子集的mass进行平均，之后再进行融合。



b. 用交算子确定融合后的上限，用并算子确定融合后的下界。





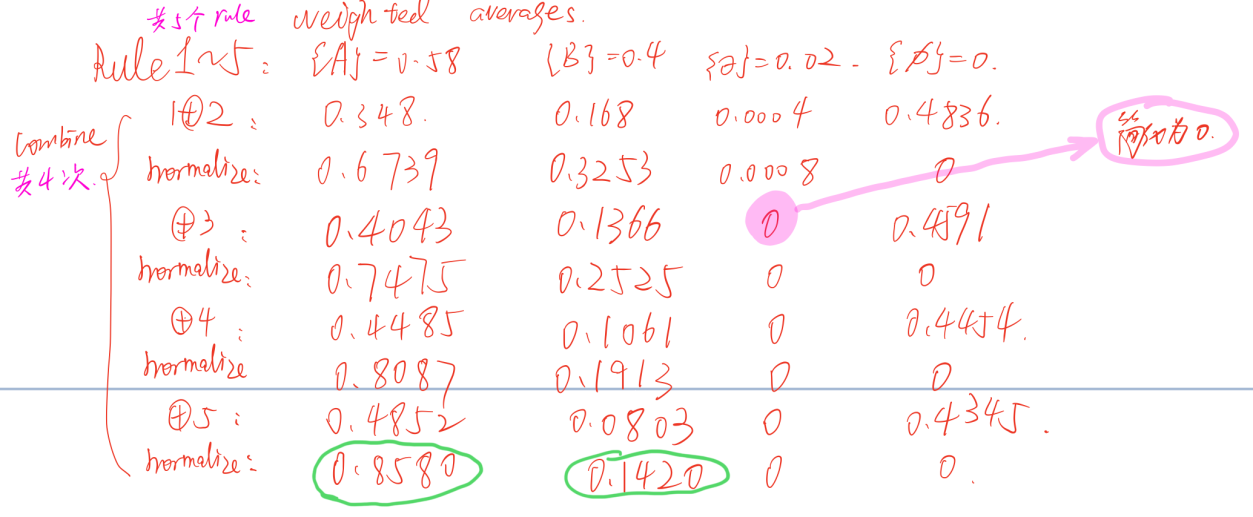
③Murphy C K的方法：

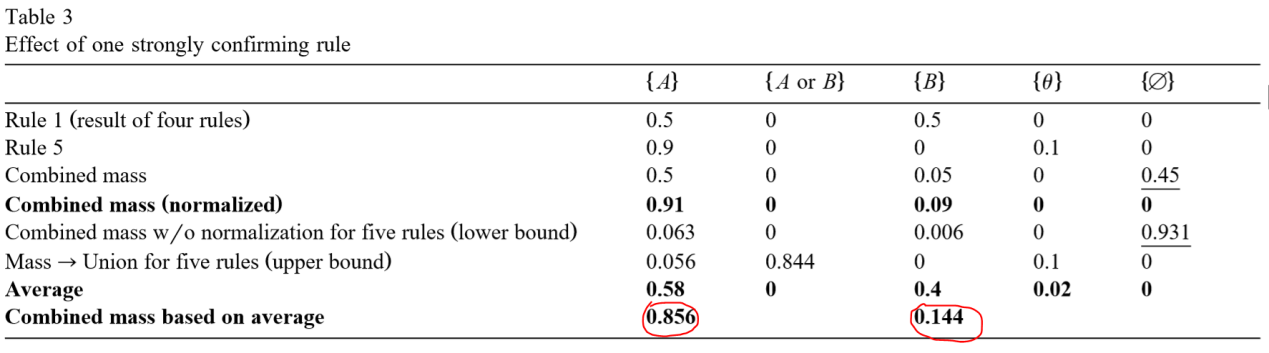
Averaging方法的收敛性不好（lack convergence），所以Murphy C K将Averaging与经典证据理论的融合方法结合。具体步骤如下：

**第一步：**假设有n个证据源，将所有证据源的mass进行平均（焦元要一一对应），得到n个相同的BPA（原文中写的是weighted averages of the masses）

**第二步：**使用经典证据理论的融合方法，对n个BPA进行两两融合，即：每次融合两个。

**注：这是我自己的理解，原文中写得有些含糊，我用这两个步骤验证了Table 3中的结果，发现有些差异，但是大体一致。以下是我的计算过程：**





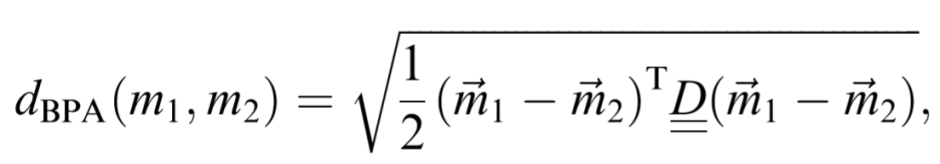
### 6．Anne-Laure Jousselme\_A new distance between two BOE

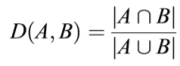
①仅仅使用欧式距离表示证据距离存在局限性：



欧式距离只能表示BPA分配的mass的相似程度，但无法比较焦元（focal element）的相似程度。

②证据距离公式：



是大小为的正定矩阵，矩阵的元素为，其中|·|表示基数（cardinality），即该集合内命题的个数。

m1和m2为两个不同证据源的BPA。

③用证据距离可以用两个角度衡量不同证据源间的相似程度：

a. BPA分配的mass的相似程度

b. 焦元的相似程度

当两个证据源完全相同(identical)时证据距离为0，当两个证据源完全冲突时证据距离为1，这与距离的近、远相对应。

### 7．邓勇\_Combining belief function based on distance function

①Murphy C K的融合方法（averageing 与经典DS融合方法结合）使用的是平均法，这意味着他认为每个证据源同等重要，但实际中每个证据源的重要程度不一定一样。

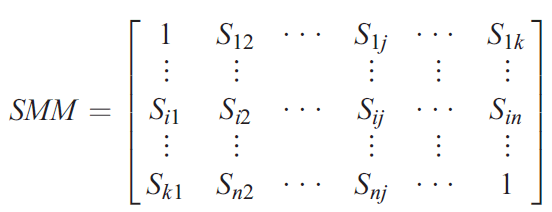
为此，邓老师的方法具体思路是：引入证据源的信任程度作为权重，当某个证据源与其他证据源相似时，这意味着证据源受到了支持，支持程度越大，这个证据源的信任程度（权重）就越大；反之，当某个证据源与其他证据源冲突越严重，这个证据源的就越小。

使用这个权重，用加权平均的方式（而非Murphy C K的直接平均），得到新的BPA（文中为MAE）。使用经典证据理论的融合方法，对MAE进行两两融合，得到最后的融合结果。

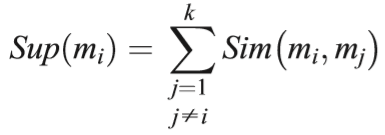
②具体步骤为：

对于不同的证据源，依据证据源之间的相似度（由证据距离得出），得出该，

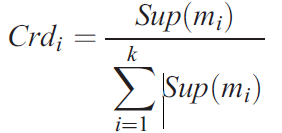
**第一步：**计算各个证据源之间的证据距离，得到相似度，由各个证据源间的相似度可得到相似矩阵：（我认为原文中有笔误，n应该都改成k）



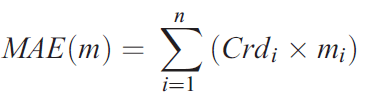
**第二步：**由相似度，得到证据源的支持度：



并由支持度得到可信度，这个可信度就是权重。



**第三步：**根据权重，对各个证据源进行加权平均，得到MAE（其实就是新的BPA）



注意：这一步不是在融合，假如加权平均前有n个BPA，加权平均后得到n个一样的MAE，之后才进行融合。

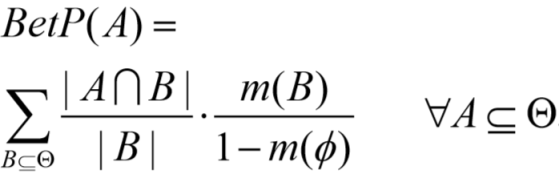
**第四步：**使用经典证据理论融合方法对MAE进行两两融合，若有n个MAE，则融合n-1次，这与Murphy C K的处理方式一样。

### 8．Liu W.R.\_Analyzing the degree of conflict among belief functions

①之前的很多论文都关注：对于特定冲突下的特定改进方案，然而很少关注对于冲突的识别。Liu的冲突识别方法是：只有当与赌博信度距离都大于某个特定的门槛时（如都大于0.75），才能称证据源出现了冲突。单一的或很大，并不能说明冲突（这也是之前论文的不足）。

②是经典证据理论融合时，还未进行normalization前，分配给空集的信度（conflicting mass）。

③pignistic概率变换（Pignistic Probability Transform, PPT）：



该变换可以把BPA变为概率分布，而这个概率分布在之后被成为：赌博信度。

④赌博信度距离：



对于两个证据源内的焦元，找的是赌博信度差距最大的值，用于表示赌博信度距离。

当两个BPA完全相同时，

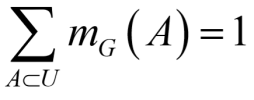
⑤用二维指标来衡量两个证据源是否冲突，当且仅当与同时满足时，才可判定两个证据源冲突。为门限值，可以人为设定，所以存在一定的主观性。

### 9．广义证据理论

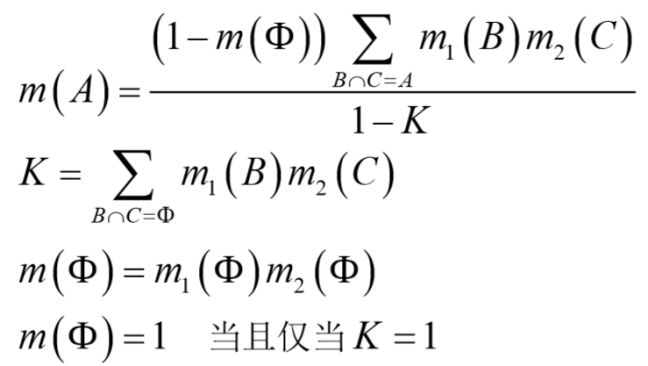
### 9.1 基本理论

（1）基本定义：

针对辨识框架不完整的情况下的信息融合、不确定信息处理，广义证据理论定义了广义基本概率指派函数（GBPA），其中m(∅)不需要强制为0，m(∅)的大小表明了支持辨识框架不完整这个命题的程度。



广义证据理论还定义了能够融合GBPA的广义组合规则（GCR）：



这里m(∅)≠K，从公式中可以看出此时的m(∅)并不是反映出证据间的冲突程度，而反映的是辩框中不存在的命题，最终合成后m(∅)表示对辩框中不存在的命题的支持程度。当m(∅)=0时，辨识框架完整，GBPA退化为经典的BPA，GCR退化为经典的组合规则。

（2）广义证据理论对比经典证据理论，最大的区别是：

广义证据理论面向开放世界，辨识框架可以不完整，m(∅)不要求强制为零，m(∅)衡量辨识框架不完整的程度。广义证据理论可以处理辨识框架不完整时的不确定信息。

经典证据理论面向封闭世界，辨识框架需完整，m(∅)强制为0，在信息融合时需要归一化使m(∅)=0。经典证据理论在处理辨识框架不完整的不确定信息时，结果不够理想。

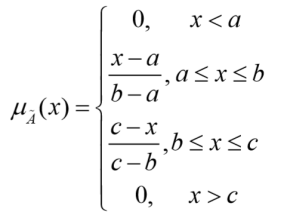
（3）提出的意义：

广义证据理论是经典证据理论的推广，面向开放世界（辨识框架不完整的情况），在开放世界的场景下，有着广阔的应用场景，而在封闭世界的情况下（辨识框架完整），广义证据理论退化为经典证据理论，可以说是向下兼容的。

### 9.2 GBPA生成方法

（1）多子集命题的模糊数表示模型

利用已知先验样本，建立模型：



其中a为先验样本的最小值，b为先验样本的平均值，c为先验样本的最大值。一个三角模糊数可以表示一个单子集命题。

从示意图角度看，（a,0）（b,ω）（c,0）三点连线与 x 轴围成一个三角形，两个三角模糊数交叠会产生一个较小的三角形，用推广模糊数表示，记为：。ω为隶属度，当ω=1 时为正则三角模糊数， 0＜ω＜1 时为推广三角模糊数。推广模糊数可用于表示多子集命题，从而更好地度量不确定度。

（2）强约束方法生成BPA

当不确定辩框是否完整时，可以采用强约束法生成GBPA。

具体方法是：设某待验证样本的一个属性值为 K，用直线 x=K表示；用先验样本建立三角模糊数表示模型（三角形）。

若x=K与对应属性的表示模型（三角形）没有交点，此时 m(∅)=1，强烈支持辩框不完整；

若x=K与三角形有交点，则该交点的纵坐标就对应着该样本支持命题的 GBPA；

若x=K与三角形有多个交点，则必然有交点处于三角模糊数的重叠部分（即推广三角模糊数），此交点为纵坐标的低点，低点的纵坐标表示该样本支持多子集命题的GBPA。交点的高点为该样本支持命题的GBPA。

最终的 GBPA 函数值之和如果大于1，则需要归一化，此时 m(∅)=0；若GBPA函数值之和小于 1，则将1减去GBPA 函数值之和赋值给m(∅)。

（3）弱约束方法生成BPA

在确认辨识框架是完整的情况下，可以采用弱约束方法生成 GBPA。弱约束方法可以样本与表示模型之间没有交点时也可以生成GBPA。

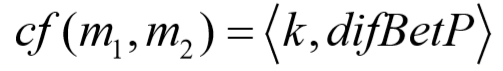
具体方法是：先使用相似度公式，计算待验证样本与各个三角模糊数（先验样本）的相似度，在计算待验证样本和三角模糊数的相似度时，可以将待验证样本看成一个特殊的三角模糊数，设待验证样本的一个属性值为 K，则它的三角模糊数参数a=b=c=K、ω=1。

然后将所有命题的相似度（包括多子集命题）归一化，即可得到对应的GBPA。注意：①样本与空集命题的相似度为0；②因为辨识框架完整，所以m(∅)=0。

### 9.3 新的冲突表示模型

（1）Liu的冲突表示模型

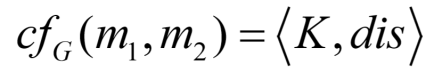
Liu的冲突表示模型中，使用“经典冲突系数 k”和“Pignistic概率距离dip”组成二维矩阵，用于表示冲突，当时，认为m1和m2两个BPA互相冲突。



但Pignistic证据距离存在的缺陷：无法区分“等可能”和“一无所知”两种情况。

（2）广义证据理论下的冲突表示模型

经过对比，使用k和 Jousselme 证据距离 dis表示冲突更为合理，将其推广到广义证据理论中，使用“广义冲突系数K”和“广义证据距离dis”组成二维矩阵表示冲突：



对于冲突系数和证据距离的区别：冲突系数会把空集（辨识框架不完整）看作是冲突来源的一部分，因为空集与任何命题的交集均为空集；而证据距离反映的是两个GBPA之间的差异程度。

针对以上特点，在辨识框架不完整的情况下（m(∅)≠0），二维矩阵应以广义冲突系数K为主度量冲突。在辨识框架完整的情况下（m(∅)=0），广义冲突系数K退化为经典冲突系数k，k存在很多局限性，所以此时二维矩阵应以证据距离为主度量冲突。

（3）冲突模型的使用策略

①处于封闭世界（m(∅)=0）：

②处于开放世界（m(∅)≠0）：不可使用经典DS组合规则→需修改组合规则

### 9.4 不完整辨识框架的识别方法

（1）总体思路

无论是 GBPA 的生成还是对冲突的表示，系列论文都对辨识框架是否完整进行了分类讨论，由于广义证据距离更多反应地是证据之间的差异性程度，所以选用m(∅)作为识别辨识框架完整与否的表征参数。

在此基础上，针对不同情况下表征参数m(∅)的变化规律进行了研究，并考虑了传感器受到干扰的情况，从证据内、证据间、样本间三种情况判断辨识框架是否完整。

（2）表征参数分析

选用m(∅)作为表征参数，针对不同情况，从证据内、证据间、样本间三个角度分析m(∅)的数据特征（无干扰）：

①辨识框架不完整时：证据内m(∅)大、证据间m(∅)数目超过半数、样本间m(∅)大。

②辨识框架完整时：证据内m(∅)小、证据间m(∅)数目少于半数、样本间m(∅)小。

（注：**证据内**是指一个证据源内的m(∅)；**证据间**是指对于同一个属性，若干证据源中，支持开放世界这个命题的证据源数目；**样本间**是指，对若干属性融合过后的m(∅)，之后求得的平均值）

当传感器受到干扰时，在辨识框架完整的条件下，数据特征如下：

①辨识框架完整、干扰小：与辨识框架完整、无干扰的数据特征一样。

②辨识框架完整、干扰大：证据内m(∅)大、证据间m(∅)数目**少于**半数、样本间m(∅)大。证据间m(∅)数目较少是这种情况最明显的特点，可以与辨识框架不完整、无干扰相区别。这种情况下，GCR不适用。

（3）辨识框架完整与否的识别策略

①用先验样本建立三角模糊数表示模型。

②强约束方法生成GBPA（用待验证样本(直线)与表示模型(三角)相交）。

③用GCR对各个样本属性的GBPA进行融合（一个样本有多个属性）。

④根据表征参数的特征，判断框架是否完整，或是存在干扰，具体方法见（2）。

### 9.5 封闭世界中冲突的处理

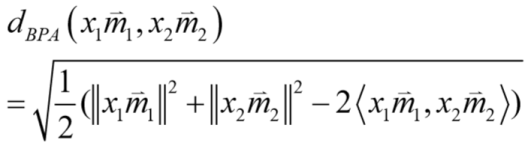
（1）总体思路

对于处理受到干扰的证据源，可以根据不同证据源的可信度，可采用“加权平均的思想”或是“打折系数的思想”，对证据源预处理，之后可采用经典的证据组合规则融合证据源。受到干扰的证据源和其余未被干扰证据源差别较大，可以给其分配一个较小的可信度。对于如何分配可信度，该文献采用的是“打折系数的思想”，用证据距离度量冲突程度，由此得出可信度作为权重，将权重归一化作为打折系数对证据源进行打折（预处理），之后采用经典的证据组合规则。

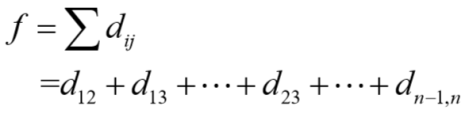
打折过程中，被折算掉的GBPA会被分配到多子集中（包括全集），也就是说，存在干扰时，这种打折处理方法可以使不确定度提高，具有一定柔性，也符合软计算的思想。相比之下，“加权平均的思想”比较刚性。

（2）结合遗传算法解决经典证据理论中的冲突

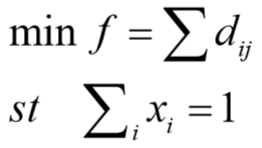
基于证据距离函数，定义加权证据距离，𝑥1、𝑥2为权重，其反映的是两个证据间的支持程度，证据距离小，相似度越高，支持度越高：



定义全局证据距离，其反映的是各个证据两两之间的加权证据距离：



冲突较小的证据之间，加权证据距离也应该小，所以定义最优化目标函数如下，其中𝑥𝑖由遗传算法在确定最值的过程中取得。



算例中首先给出了单个传感器受到干扰的情况。使用弱约束方法生成GBPA，一个传感器受干扰失效，另外三个传感都正常，所以会给失效传感器分配一个非常小的权重，最后的实际合成结果与无干扰一样，比较理想；然而，由于强约束方法会让被干扰传感器生成很多m(∅)=1，提示辨识框架不完整，导致最终合成后的 m(∅)也比较大，所以我们无法从中得出合理的结论。

对于随机干扰的情况，文献中给出了单传感器受到影响的算例，使用弱约束方法生成GBPA，可以看到被干扰证据的折扣系数远小于其他证据的折扣系数，最终同样得到了理想的实验结果。

**结合以上冲突处理的思路，以及实验结果，我认为存在的问题是：当被干扰的传感器过多时，遗传算法确定的权重已经不能很好地反映出不同证据的可信度，需要对其进行改进。**

**我的改进思路是：对于随机干扰，如果多数传感器都受到很大的随机干扰，最小全局证据距离（由遗传算法求得）会比少数传感器受干扰大，因此可以设定一个最小全局证据距离的阈值，若大于阈值，即使最终的结果看起来合理，也不可采用，应重新采集数据。**

（3）辨识框架完整、高度冲突情况下的证据融合策略

①当证据高度冲突时，先进行冲突分析，若判定系统辨识框架完整，则冲突是因为传感器受干扰，此时需利用弱约束方法生成GBPA。（用弱约束方法可以从BPA生成角度降低冲突）

②由最优化目标函数，确定各个证据的权重。

③将权重归一化（选出权重的最大值，各个证据的权重都与最大值相除），生成折扣系数，0≤折扣系数≤1。

④用折扣系数对相应的证据进行打折，之后把被折算掉的GBPA分配到全集中，使各个焦元的GBPA满足总和为1的条件。注意：因为是弱约束方法，所以m(∅)=0。

⑤基于经典的证据组合规则融合经过打折后的证据。