

文章编号: 1000 8055(2008)09 1633-04

基于 Bayes-模糊逻辑算子的小子样可靠性信息融合方法

冯 静, 周经伦

(国防科技大学 信息系统与管理学院系统工程系, 长沙 410073)

摘 要: 为了解决航空航天复杂系统可靠性评定中样本量小而导致评估结果可信度不高的问题, 引入模糊逻辑算子这一非线性模型对多源可靠性验前信息进行融合, 并给出了其参数估计的第二类极大似然(ML-II)估计方法, 通过仿真示例说明了融合方法的有效性。

关 键 词: 可靠性评定; 信息融合; 模糊逻辑融合算子; Bayes 方法; ML-II 方法

中图分类号: O213.2; TB114

文献标识码: A

Small sample reliability information fusion approach based on bayes-fuzzy logistic operator

FENG Jing, ZHOU Jing-lun

(Information System and Management School,
National University of Defense Technology, Changsha 410073, China)

Abstract: In order to solve lower confidence level of reliability assessment results for complex system in aero and aircraft fields caused by little test data, a fusion method based on fuzzy aggregation operator (FAO) was presented to fully embody the degree of redundancy and complementarity of information from multiple sources. The unknown parameters of the fusion operator were estimated by the second maximum likelihood estimation method (ML-II). The simulation example shows that the proposed method is effective.

Key words: reliability assessment; information fusion; fuzzy aggregation operator; bayesian analysis; the second maximum likelihood estimation method (ML-II)

对于由许多不同单元组成的复杂产品, 例如卫星、火箭等, 一方面, 由于人力、资金、时间等限制, 系统现场可靠性试验数量很少, 其可靠性评定和寿命预测属于小子样问题, 依赖传统的基于大样本的数理统计方法将难以获得客观结论; 另一方面, 在产品研制过程中, 积累了各种来源的验前信息, 包括相似型号产品试验和使用信息、单元在不同环境下的试验信息、专家经验信息等, Bayes

方法可以把多源验前信息和少量的现场试验信息合理融合, 从信息论的角度考虑, 这种扩充信息量的途径必然会使可靠性评定结论更为客观和可信。决定 Bayes 方法使用成败的关键, 是如何根据多源验前信息确定验前分布的问题, 验前分布选择不慎将直接影响到评定结论的客观性。

多源验前信息融合方法已成为国内外学者研究的热点问题。现有的融合方法包括加权融合、环

收稿日期: 2007-08-21; 修订日期: 2007-12-17

基金项目: 国家自然科学基金(60701006)

作者简介: 冯静(1975-), 女, 安徽芜湖人, 讲师, 博士, 主要从事小子样试验分析和信息融合技术的研究。

境因子融合、多阶段变总体融合等. 其中加权融合应用极为广泛, 如可信度融合^[1]、相关函数融合^[2]、专家设定权重^[3]等, 但现有的加权融合方法大多是基于线性加权技术进行的, 是在各信源相互独立这一假设的基础上建立的. 随着系统的日益复杂化和检测方式的多样化, 人们越来越注意到, 在复杂系统的可靠性试验中, 一方面, 各信息源之间存在着许多不可忽略的相互作用; 另一方面, 各信息源所提供的信息具有一定程度的不确定性^[4], 如信息本身的不完全性, 表达信息的不严密性以及各种信息的矛盾性等等, 因此, 很多场合下不再适用于用简单的线性加权模型来描述.

考虑到在复杂系统可靠性信息融合过程中, 不同来源的验前信息可能是冗余的, 或者是互补的. 当多个来源的信息对同一目标进行描述时, 所得到的信息是冗余的, 对冗余信息进行融合可以降低不可靠性. 当多个来源的信息对目标的不同特征进行描述时, 处理的信息是互补的, 对互补信息进行处理可以降低不确定性.

本文将采用模糊逻辑融合算子来描述多源验前信息之间的冗余性和互补性, 确定融合验前分布的非线性加权形式, 并给出融合验前分布超参数的 ML-II 估计方法. 在文中通过一个实例说明了该方法的有效性.

1 模糊逻辑算子用于信息融合的原理

1.1 信息融合的物理基础

1) KSC 准则: 从一个传感器获得一些信号时, 由于种种原因, 往往不足以反映目标或对象的全部实际情况, 因此, 需要一些其他信息来补充说明这一信息. 这种补充信息不是来自同一信息源, 而是来自另一个传感器. 这些传感器信息的补充称之为知识补充 KSC^[57] (Knowledge source corroboration). 由 KSC 构成的检测原理叫做知识补充原理.

2) BEW 准则: 当只使用一个传感器时, 由于测量环境不可避免地存在不同程度的噪声, 所得到的信息需要进一步得到证实, 否则这种信息将不能充分的证明事件是否一定发生或没有发生. 这种用一个或一个以上的附加传感器来证实另一个传感器信息的可靠程度, 称之为信任程度的增强或削弱 BEW^[58] (Belief enhancement/withdraw). 由 BEW 构成的检测原理叫做信任程度的增强或削弱原理.

一个测量系统中可能存在一个以上的传感

器, 这些传感器集中在一起, 用以产生一定程度的信息冗余和互补, 来判断、证实、确定所测目标和事件的准确性. 这种利用多个传感器反映同一目标或同一事件的过程叫做数据融合过程 (Data fusion process). 利用上面所提的两个原则, 可以对传感器的集合提供比较科学的融合过程.

1.2 模糊逻辑算子的理论基础

假设有两个传感器 S_1 和 S_2 , 它们分别从不同角度去描述同一目标 O 和反映同一事件 E , 不同的传感器分别描述目标子集 O_1 和 O_2 或反映事件子集 E_1 和 E_2 . 对于 KSC 准则, $O = O_1 \cup O_2$, $\{O_1, O_2 \in O\}$; 对于 BEW 准则, $E = E_1 \cap E_2$, $\{E_1, E_2 \in E\}$. 在实际应用中, 还有一种情形, 既属于 KSC 原则, 又属于 BEW 原则, 此时需要利用介于交和并之间的模糊逻辑算子来表示. 文献[9]中 Djumovic 提出的广义期望算子, 能很好地确定处于交和并之间的模糊关系. 在这里, 我们将这个数学模型用于可靠性信息的融合中, 因为各来源验前信息之间, 既存在互补性, 也存在冗余性.

广义期望算子 (Generalized mean operator) 的定义为

$$g(x_1, \dots, x_n; p, w_1, \dots, w_n) = \left(\sum_{i=1}^n w_i x_i^p \right)^{1/p}$$

其中 w_i 为传感器 x_i 的权值, 并且 $\sum_{i=1}^n w_i = 1$; p 表示模糊度, $p \in (-\infty, +\infty)$, 表现了各来源信息之间的互补和冗余度.

考虑两个元素 x_1, x_2 的广义期望算子 $g(x_1, x_2; p, w_1, w_2)$, 显然,

$$\min(x_1, x_2) \leq g(x_1, x_2; p, w_1, w_2) \leq \max(x_1, x_2)$$

由此可见, 广义期望算子是介于并和交之间的操作. 现在的关键问题是, 针对不同的融合系统, 确定参数 p 和 w_i 的值. 如果关于传感器的可靠性信息已知, 那么权值 w_i 可以确定, 如果关于冗余度或互补程度的信息知道, 那么可确定参数 p . 然而在通常情况下, 这些信息都不知道, 因此, 需要针对具体问题, 结合物理原理和数据特点合理确定它们的取值.

2 基于模糊逻辑算子的信息融合

2.1 基于模糊逻辑算子的融合验前分布

复杂系统的寿命 X (即系统无故障工作时间) 为随机变量, 服从密度函数为 $f(x|\theta)$ 的分布, 其中 θ 为未知参数. 假设有 m 个来源的验前信息

$\{x_1^{(i)}, \dots, x_{n_i}^{(i)}\}, i = 1, \dots, m$; 有 n 个现场样本数据 $x_1, \dots, x_n$.

$\forall i$, 设 $\pi_i(\theta)$ 为通过验前数据 $x_1^{(i)}, \dots, x_{n_i}^{(i)}$ 求得的验前分布^[10], $i = 1, 2, \dots, m$. 融合后的验前分布记为 $\pi(\theta)$

$$\pi(\theta; p, w_1, \dots, w_m) = \left[\sum_{i=1}^m w_i \pi_i^p(\theta) \right]^{1/p}$$

其中 p 为模糊度, $p \in (-\infty, +\infty)$, 表示信息源之间的冗余和互补程度. w_i 为 $\pi_i(\theta)$ 的权值, $\sum_{i=1}^m w_i = 1$, 表示信息源 $\pi_i(\theta)$ 的可信度.

下面以 $m = 2$ 为例进行说明, 即

$$\pi(\theta) = [w_1 \pi_1(\theta) + (1 - w_1) \pi_2(\theta)]^{1/p} \quad (1)$$

2.2 Bayes 公式计算验后分布

关于参数 θ 的任何统计推断必须依据未知参数 θ 的验后分布来进行, 因为, 在获得样本之后, 验后分布反映了对参数 θ 的全部了解. 设 $X = (X_1, \dots, X_n)$ 为独立同分布的现场样本, 未知分布参数 θ 的验前密度函数为 $\pi(\theta)$, 当获得 X 之后, θ 的验后密度函数 $\pi(\theta|X)$ 由 Bayes 公式给出

$$\pi(\theta|X) = \frac{f(X|\theta)\pi(\theta)}{\int_{\Theta} f(X|\theta)\pi(\theta)d\theta} \quad (2)$$

2.3 ML-II 方法估计参数

基于 Robbins 经验 Bayes 方法的第二类极大似然估计方法^[10]的基本思想是: 试验中子样的出现看作是由子样 X 的边缘分布 $m(X)$ 产生的, 而边缘分布中含有未知参数 θ 的验前分布 $\pi(\theta)$, $m(X)$ 可表示为

$$m(X) = \int_{\Theta} f(X|\theta)\pi(\theta)d\theta$$

取 $\pi(\theta)$ 为某种特定形式的分布, 其分布参数由 $m(X)$ 取极大值的方法获得, 这就是第二类极大似然估计方法(The second maximum likelihood estimation method, ML-II). 边缘分布密度从频率学派的观点来看, 它描述了样本 X 在验前分布密度 $\pi(\theta)$ 之下的平均行为, 样本将按此分布而发生. 于是, 现场样本 $X = (x_1, \dots, x_n)$ 的似然函数为

$$m(X|w_1, p) = \prod_{i=1}^n m(x_i|w_1, p)$$

其中 $m(x_i|w_1, p) = \int_{\Theta} f(x_i|\theta)\pi(\theta)d\theta$. 求 $m(X|w_1, p)$ 关于 w_1, p 的极大值, 即令

$$\frac{\partial m(X|w_1, p)}{\partial w_1} = 0, \frac{\partial m(X|w_1, p)}{\partial p} = 0$$

从而解得 w_1, p 的取值 $\hat{w}_1, \hat{p}$.

将 $\hat{w}_1, \hat{p}$ 代入(1), 得到融合验前分布密度, 进而得到融合验后分布, 再以验后分布密度为基础进行未知参数的点估计、置信区间估计、假设检验等. 从信息论的角度看, 这种融合方法充分合理的利用了多源验前信息, 可以增强统计决策的稳健性, 合理地减少现场试验的次数.

3 仿真示例

设有一组现场可靠性试验数据 $y_1, \dots, y_5$ 服从 $N(\theta, 1)$ 分布, θ 为未知参数, 表示某项可靠性指标, 代表系统的可靠性水平, 是需要通过试验确定的值. 仿真时, 令 $\theta = 4.5$, 得到一组随机数 $Y = (y_1, \dots, y_5) = (4.238\ 21, 6.527, 5.247\ 8, 3.838\ 74, 2.959\ 84)$. 根据经验可知, θ 的验前分布为正态分布, 但正态分布的超参数未知, 关于未知参数 θ 有两组验前信息, 得到的验前分布分别为 $\pi_1(\theta) = N(3.9, 1.2)$ 和 $\pi_2(\theta) = N(4.8, 1)$ 由 Bayes 公式, 分别计算得到相应的后验分布为 $N(4.467\ 7, 0.171\ 429)$ 和 $N(4.601\ 93, 0.166\ 67)$ 设融合验前分布为

$$\pi(\theta) = [w_1 \pi_1(\theta) + (1 - w_1) \pi_2(\theta)]^{1/p}$$

于是

$$m(y_i) \triangleq \int_{\Theta} N(y_i|\theta)\pi(\theta)d\theta, \quad i = 1, \dots, 5$$

则 $Y = (y_1, \dots, y_5)$ 的似然函数为

$$L(Y) \triangleq L(y_1, \dots, y_5|w_1, p) = \prod_{i=1}^5 m(y_i)$$

由 ML-II 方法, 得到参数估计结果为 $w_1 = 0.4$, $p = 1.4$.

图 1 给出了单个验前分布 $\pi_1(\theta)$, $\pi_2(\theta)$ 和融合验前 $\pi(\theta)$ 的比较, 从图中可以看出, 在 $\pi(\theta)$ 中, 被两个验前 $\pi_1(\theta)$ 和 $\pi_2(\theta)$ 都支持的部分的函数值得到了加强, 而其余部分则被减弱了.

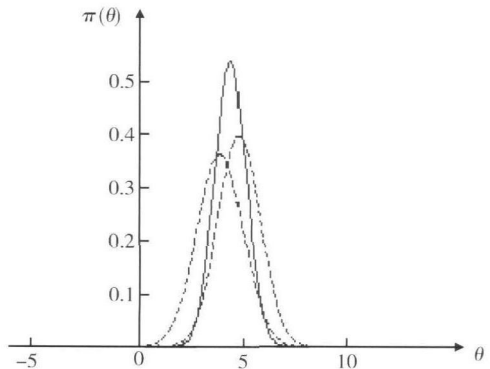


图 1 验前分布比较

Fig. 1 Compare of prior distributions

融合验前分布为 $N(4.390\ 91, 0.738\ 549)$, 进一步计算得到融合验后分布为 $N(4.516\ 33, 0.146\ 341)$. 融合验后的均值比单个验后 $N(4.467\ 7, 0.171\ 429)$ 和 $(4.601\ 93, 0.166\ 67)$ 的均值更加接近为 θ 的真值 4.5, 说明该融合方法是合理的、有效的.

4 结 论

本文介绍的基于非线性加权的模糊逻辑算子可适用于多源验前信息的冗余度、互补程度都不知道的情形. 融合过程中, 通过模糊逻辑算子及 ML-II 方法确定信息的冗余度或互补程度, 减少了评定过程中的主观性, 增强了评定结论的可信性.

参考文献:

- [1] 张金槐. 多源信息的 Bayes 融合精度鉴定方法[J]. 国防科技大学学报, 2001, 23(3): 93-97.
ZHANG Jinhuai. Accuracy detection method using Bayesian multi-sensor data fusion technique[J]. Journal of National University of Defense Technology, 2001, 23(3): 93-97. (in Chinese)
- [2] 冯静, 刘琦, 周经伦, 等. 相关函数融合法及其在可靠性分析中的应用[J]. 系统工程与电子技术, 2003, 25(5): 4-6.
FENG Jing, LIU Qi, ZHOU Jinlun, et al. Correlation information fusion method and application in reliability[J]. System Engineering and Electronics, 2003, 25(5): 4-6. (in Chinese)
- [3] 张士峰, 蔡洪. Bayes 分析中的多源信息融合问题[J]. 系统仿真学报, 2000, 12(1): 54-56.
ZHANG Shifeng, CAI Hong. Fusion of information of multiple sources in Bayesian analysis[J]. Journal of System Simulation, 2000, 12(1): 54-56. (in Chinese)
- [4] Hall D L, Llinas J. An introduction to multisensor data fusion[J]. Proceedings of the IEEE, 1997, 85(1): 6-23.
- [5] Zimmermann H J, Zysno P. Latent connectives in human decision making.[J]. Fuzzy Sets and System, 1980, 4: 37-51.
- [6] 童树鸿, 沈毅, 刘志言. 数据融合的模糊模型和算法研究[J]. 哈尔滨工业大学学报, 2002, 34(1): 7-11.
TONG Shuhong, SHE Yi, LIU Zhiyan. Fuzzy model and algorithms for data fusion[J]. Journal of Harbin Institute of Technology, 2002, 34(1): 7-11. (in Chinese)
- [7] 张军英. 信息融合中重要性测度的区间估计[J]. 西安电子科技大学学报, 1999, 26(3): 332-336.
ZHANG Junying. Interval estimation of importance measure in information fusion[J]. Journal of Xidian University, 1999, 26(3): 332-336. (in Chinese)
- [8] 游林儒, 张晋格, 王炎. 一种多源信息融合方法及其应用[J]. 哈尔滨工业大学学报, 2000, 32(4): 101-103.
YOU Linru, ZHANG Jing, WANG Yan. Multi-source information fusion and its application[J]. Journal of Harbin Institute of Technology, 2000, 32(4): 101-103. (in Chinese)
- [9] 刘莹, 刘三阳, 马建峰. 基于遗传算法的多传感器数据融合[J]. 微电子学与计算机, 1999, (2): 22-27.
LIU Ying, LIU Sanyang, MA Jianfeng. Genetic algorithm for multisensor fusion [J]. Microelectronics and Computer, 1999, (2): 22-27. (in Chinese)
- [10] 张金槐, 唐雪梅. Bayes 方法[M]. 修订版. 长沙: 国防科技大学出版社, 1992.