

M

odeling Approach on Cognizing Specific Parameters in Thermal Process: Extensions and Applications of Evidence Theory

热工过程特殊参量的认知建模方法

——证据理论的拓展与应用

王培红 苏志刚 / 著



东南大学出版社
SOUTHEAST UNIVERSITY PRESS

国家自然科学基金(51106025)资助
东南大学科技出版基金资助

热工过程特殊参量的认知建模方法

——证据理论的拓展与应用

王培红 苏志刚/著



 东南大学出版社
SOUTHEAST UNIVERSITY PRESS

内容提要

热工过程是一个受传热、传质等诸多因素影响且具有非线性、强耦合的复杂物理过程。该过程中涉及大量的、未知的不精确、不确定性信息。本书旨在通过证据理论的学习和研究,提出新的理论和方法,(主要)用于解决热工过程特殊参量(无样本参量)认知问题,为实现热工对象的在线可控、在线优化以及提高热工对象的生产能力并降低能耗奠定基础。

本书可作为热能工程、动力机械及工程、能源信息技术、人工智能、应用数学与工程等专业研究人员的学术文献,也可以作为研究生的教材或参考读物。同时对相关专业的工程技术人员和管理人员也具有参考价值。

图书在版编目(CIP)数据

热工过程特殊参量的认知建模方法:证据理论的拓展与应用 / 王培红,苏志刚著. — 南京:东南大学出版社, 2012. 8

ISBN 978-7-5641-3145-6

I. ①热… II. ①王… ②苏… III. ①证据—理论研究
IV. ①D915.130.1

中国版本图书馆 CIP 数据核字(2011)第 247553 号

热工过程特殊参量的认知建模方法——证据理论的拓展与应用

出版发行 东南大学出版社

出版人 江建中

责任编辑 马 伟

电 话 (025)83791797

社 址 南京市四牌楼 2 号

邮 编 210096

经 销 全国各地新华书店

印 刷 兴化印刷有限责任公司

开 本 700mm×1000mm 1/16

印 张 18.5

字 数 388 千字

版 次 2012 年 8 月第 1 版

印 次 2012 年 8 月第 1 次印刷

书 号 ISBN 978-7-5641-3145-6

定 价 59.00 元

(凡有印装质量问题,请与我社营销部联系。电话:025-83791830)

前言

1976年, G. Shafer 出版了 *A Mathematical Theory of Evidence* 一书, 标志着证据理论的诞生。证据理论是一种广泛用于处理不精确、不确定性信息的理论工具, 经过近 50 年的发展, 在决策、模式识别和数据挖掘等众多领域取得重要应用成果, 得到了国内外学术界的广泛认同和关注。

热力设备(过程)是一个受传热、传质等诸多因素影响且具有非线性、强耦合特点的复杂物理对象。在热力设备的优化与控制过程中, 往往会涉及不精确、不确定性信息的处理问题, 属于证据理论的应用范畴。本书针对热力设备(过程)特殊参量认知的需求, 基于证据理论及其拓展研究, 提出热力设备(过程)特殊参量(无样本参量)认知的理论和方法, 为实现热力设备(过程)的性能评价、控制优化提供软测量信息保障。

全书共 10 章。第 1 章绪论, 对热力设备(过程)特殊参量认知的背景及需求进行了描述, 介绍了证据理论的概念、框架与认知建模的基本思想。第 2 章~第 6 章主要介绍证据理论的拓展研究成果, 包括: 拓展权函数(第 2 章)、证据融合独立性假设及广义范数融合法则(第 3 章)、证据 k -NN 分类算法及其在模式识别中的应用(第 4 章)、证据邻域粗糙集模型及其在属性约简中的应用(第 5 章)、证据回归多模型建模方法及其在过程预测中的应用(第 6 章)。第 7 章~第 10 章主要介绍证据理论拓展研究成果在热力设备(过程)特殊参量认知中的应用, 其中, 第 7 章介绍电站钢球磨煤机料位(无样本)特殊参量的认知建模、第 8 章介绍电站汽轮机排汽焓(数值模拟样本)特殊参量的认知建模、第 9 章介绍了电站锅炉 NO_x 排放(有样本)特殊参量的认知建模。在第 10 章对区间证据理论中最大置信区间的存在性进行了初探, 为无法获得数据情况下的无样本参量认知问题奠定了基础。

本书可作为热能工程、动力机械及工程、能源信息技术、人工智能、应用数学与工程等专业研究人员的学术文献, 也可以作为研究生的教材或参考读物。同时对相关专业的工程技术人员和管理人员也具有参考价值。

本书是国家自然科学基金(51106025)项目的主要研究成果, 获得东南大学科技出版基金的资助。全书由王培红教授和苏志刚博士合著, 王培红教授撰写第 1 章及第 8、9 章, 并负责统稿; 苏志刚博士撰写完成其余章节。刘莎、唐菲菲、胡恩俊、葛举生、赵春等研究生参与了本书的文字编辑工作。

目 录

第1章 绪论	1
1.1 研究背景及问题描述	1
1.2 证据理论:基本概念及相关计算	8
1.2.1 证据表征	8
1.2.2 证据融合	11
1.2.3 广义贝叶斯定理	17
1.2.4 证据决策	18
1.2.5 模糊证据理论	18
1.2.6 证据理论的矩阵计算	20
1.3 证据理论发展现状及存在的问题	23
1.3.1 证据(信度函数)的解释	23
1.3.2 证据的构建	24
1.3.3 证据融合悖论和独立性假设	25
1.3.4 证据融合计算复杂度的简化	28
1.3.5 证据理论存在的问题和发展趋势	29
1.4 主要研究内容及安排	30
1.5 主要贡献	34
第2章 拓展权函数	36
2.1 证据正则分解	36
2.1.1 与权函数	36
2.1.2 并权函数	37
2.1.3 基于权函数的证据融合法则	38
2.2 权函数拓展研究	40
2.3 拓展权函数的应用	42
2.3.1 基于拓展权函数的信度函数间的转换	42
2.3.2 基于拓展权函数的改进型融合法则	44
2.4 本章小节	46
第3章 证据融合独立性假设及广义范数融合法则	47
3.1 引言	47

3.2	广义范数的定义及证明	47
3.3	基于广义范数的证据融合法则研究	57
3.3.1	广义 T 范数与法则	58
3.3.2	广义 U 范数与法则	59
3.3.3	广义 T 范数并法则和广义 U 范数并法则	62
3.3.4	广义范数融合法则的讨论	64
3.3.5	广义范数融合法则的简化	65
3.3.6	广义范数融合法则应用示例及分析	68
3.4	模糊证据融合法则研究	74
3.4.1	经典模糊证据融合法则分析	74
3.4.2	模糊证据的离散化策略研究	77
3.4.3	基于广义 T 范数与法则及离散化策略的模糊证据融合	84
3.5	本章小结	86
第4章	证据 k-NN 分类算法及其在模式识别中的应用	87
4.1	引言	87
4.2	经典证据 k -NN 分类器	88
4.3	经典证据 k -NN 分类器存在的局限性分析	89
4.4	三种变体证据 k -NN 分类器	92
4.4.1	基于自适应度量空间及参数优化的证据 k -NN 分类器	92
4.4.2	基于广义 T 范数与法则的证据 k -NN 分类器	94
4.4.3	鲁棒自适应证据 k -NN 分类器	97
4.4.4	实验分析	101
4.5	本章小结	106
第5章	证据邻域粗糙集模型及其在属性约简中的应用	107
5.1	引言	107
5.2	邻域粗糙集模型及属性约简	108
5.3	基于邻域证据决策误差率的属性约简算法研究	113
5.4	证据邻域粗糙集模型及属性约简算法研究	120
5.5	实验分析	125
5.5.1	属性约简算法敏感性分析及分类精度验证	126
5.5.2	证据邻域决策系统的属性约简模拟分析	132
5.6	本章小结	134
第6章	证据回归多模型建模方法及其在过程预测中的应用	135
6.1	引言	135

6.2	不精确和不确定性知识表征及简化策略	136
6.3	证据回归多模型	140
6.3.1	全局模型建模	141
6.3.2	局部模型建模	141
6.3.3	证据回归多模型参数辨识策略	144
6.3.4	证据回归多模型校正策略	146
6.4	证据回归多模型应用举例	147
6.4.1	基于证据回归多模型的过程预测分析	147
6.4.2	传感器可靠性及测量精度时变过程的预测分析	155
6.5	证据回归多模型讨论	159
6.5.1	样本冲突性处理	159
6.5.2	样本不确定性反映	162
6.5.3	证据回归多模型与 EVREG 模型的比较	164
6.6	本章小结	165
第7章	基于证据回归多模型的钢球磨煤机料位认知建模	167
7.1	引言	167
7.2	制粉系统钢球磨煤机机理分析	168
7.2.1	钢球磨煤机数学模型及特性	168
7.2.2	料位影响因素分析	173
7.3	实验	177
7.3.1	专家知识库构建	178
7.3.2	钢球磨煤机现场实验及实验结果	179
7.4	无样本参量证据样本构建策略	185
7.5	无样本参量料位在线认知模型	189
7.5.1	基于证据回归多模型的料位认知	189
7.5.2	认知结果分析与验证	196
7.6	本章小结	206
第8章	基于证据回归多模型的汽轮机排汽焓认知建模	207
8.1	引言	207
8.2	汽轮机全工况特性重构模型	208
8.2.1	基准工况的选择原则	209
8.2.2	汽轮机本体变工况特性	209
8.2.3	凝汽器特性	211
8.2.4	回热加热器特性	212
8.2.5	给水泵及给水泵汽轮机特性	213

8.2.6	汽轮机工况重构方法步骤	214
8.2.7	基于汽轮机工况特性重构模型的全日负荷数据样本	215
8.3	排汽焓证据样本构建	219
8.4	无样本参量排汽焓在线认知模型	222
8.5	本章小节	229
第9章	证据回归多模型简化及其用于热工过程有样本参量的监测	230
9.1	引言	230
9.2	证据回归多模型简化研究及算例分析	231
9.2.1	简化证据回归模型	231
9.2.2	简化证据回归模型预测精度验证分析	233
9.3	NO _x 排放响应特性建模	236
9.3.1	600 MW 机组(12 试验样本)的实例分析	236
9.3.2	300 MW 机组(279 试验样本)的实例分析	240
9.4	本章小节	246
第10章	区间证据理论及其在决策中的应用	247
10.1	引言	247
10.2	区间证据理论及区间证据融合	248
10.2.1	区间证据的定义	248
10.2.2	区间证据的融合法则及其分析	249
10.3	区间证据的最大置信区间	253
10.3.1	数值模拟实验及实验结果:最大置信区间的存在性 演示	253
10.3.2	最大置信区间的存在性分析及理论证明	257
10.3.3	不确定情况分析	263
10.3.4	最大置信区间的推广性分析	264
10.4	区间证据最大置信区间在决策中的应用	264
10.5	本章小结	267
参考文献		268

第1章 绪论

1.1 研究背景及问题描述

在流程工业领域,热力设备得到了极其广泛的应用。由于设备结构的复杂性、设备性能的时变性以及设备工作环境的恶劣性,一些重要过程参量的测量问题尚未得到有效的解决。

目前热力设备过程参数的测量问题可以分为三类,一是直接测量参量类(如常规热工参量压力、温度等);二是在线测量受到测量精度、测量成本等因素的影响,不宜直接测量参量类(如锅炉排烟飞灰含碳量等);三是热力设备特性监测需要但无法直接测量参量类(如钢球磨煤机的料位)。

对于第一类参量,现有仪表设备已解决其测量问题;针对第二类过程参量的测量问题,产生了软测量建模的技术需求,软测量是指利用可测辅助变量,通过数学模型及软件来实现不便测量的主导变量的间接测量。软测量技术最初产生于化工领域,主要用于解决可以获得离线测量样本的测量问题。由于第三类过程参量无法测量,也就无法利用在线或离线测量的手段获得该参量的测量样本。因此,产生了对无样本过程参量认知的现实需求。无样本参量认知是指对真实存在、具有明确物理意义、但无法在线实时监测且(在一般条件下)无法获得离线采样样本的过程参量当前状态(值)的一种预测或评估。实现无样本参量认知建模,有利于实现复杂对象的在线优化、在线可控,从而提高该对象的生产能力并降低能耗。

本书重点研究第三类参量的测量问题,即无样本过程参量的认知建模,并探讨无样本参量认知建模方法用于实现第二类过程参量的测量问题。

根据获得的数据信息的不同形式,无样本参量的认知问题可以细分为三种,即:(1)输入为连续型变量,无样本参量亦为连续型变量的无样本参量认知问题;(2)输入为连续型变量,无样本参量取值为离散型变量的无样本参量认知问题;(3)输入为语义变量(如用模糊数表征的专家知识),输出为决策值的无样本参量认知问题。针对第(1)种无样本参量认知问题,可将其看作回归预测问题来研究;针对第(2)种无样本参量认知问题,可将其看作特征分类问题来研究;针对第(3)种,可将其作为决策问题来研究。本书重点研究第(1)种无样本参量的认知问题,

初步研究第(2)种和第(3)种无样本参量的认知问题(分别见第4章、第10章)。下文中,在不混淆的情况下,无样本参量的认知即指第(1)种无样本参量的认知问题。

下面对无样本参量的认知建模问题进行数学描述。

假设具有无样本参量的物理设备为 O 。该设备的状态通过一组过程参量来表征,包括可测过程参量和无样本参量(Unmeasured Parameter/Variable)。记无样本参量为 y 。记可测过程参量为 $x_1, x_2, \dots, x_n$, 或记为向量形式 $x = [x_1, x_2, \dots, x_n]^T$, 如图 1-1 所示。那么,在任意 τ 时刻,各可测过程参量的观测值可记为: $x(\tau) = [x_1(\tau), x_2(\tau), \dots, x_n(\tau)]^T$ 。当进行 N 次采样时,则可测过程参量样本集可表示为如下形式: $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T, i=1, 2, \dots, N$, 其中 N 被称为样本容量。

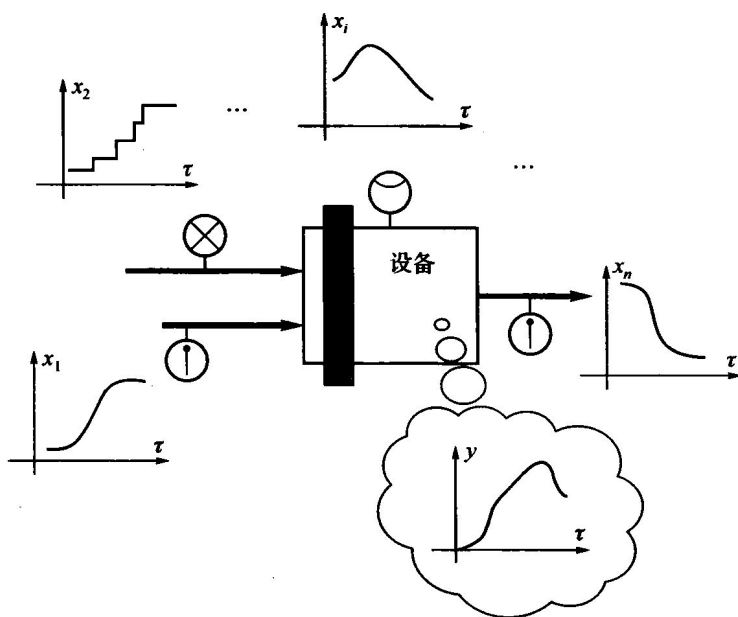


图 1-1 反映设备对象特性的过程参量及无样本特性参量

如图 1-1 所示,过程参量 x 与无样本参量 y 从不同方面反映设备的物理状态。当设备处于某一工作状态时,存在一组观测值 (x_i, y_i) 与之对应,如图 1-2(a) 所示;反之,当观测到一组观测值 (x_i, y_i) , 则可确定该设备处于某一工作状态,如图 1-2(b)。那么,过程参量 x 与无样本参量 y 通过“设备状态”而具有内在联系,如图 1-2(c) 所示,即:当无样本参量 y 为某一数据取值时,设备处于某一工作状态,同时,对应于该设备状态,可观测一组或几组观测值 x , 那么也就存在相应的一组或几组观测值 x 与 y 相对应;反之,当参量 x 的取值一定时,亦存在着 y 与之相对应。

以上过程可以通过以下数学形式进行描述。当物理设备 O 处于任意时刻 τ 时, $x(\tau) = [x_1(\tau), x_2(\tau), \dots, x_n(\tau)]^T$ 及 $y(\tau)$ 之间存在着对应关系 f (各变量均标

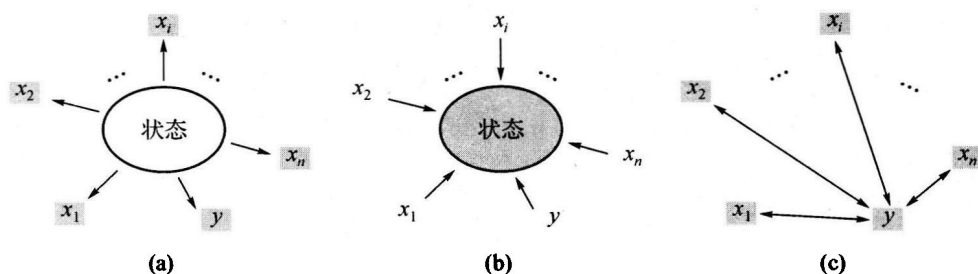


图 1-2 无样本参量与其关联可测过程参量之间内在关系示意图

准化为 $[0,1]$ 无量纲量),即:

$$f: \underbrace{[0,1] \times [0,1] \times \cdots \times [0,1]}_n \rightarrow [0,1], \text{ 或 } f: [0,1]^n \rightarrow [0,1] \quad (1-1)$$

$$y = f(x) = f(x_1, x_2, \dots, x_n) \quad (1-2)$$

式(1-1)和式(1-2)中的对应关系 f 可能为一对一、多对一或一对多的非线性关系。在现实中,可测量参量 x 和无样本参量 y 之间具体的对应关系无法获得,而往往只能通过专家知识描述该对应关系,如: x 增大, y 减小;或“ $\text{If } x \text{ 是 } \tilde{A}, \text{ Then } y \text{ 可能是 } \tilde{B}$ ”,其中, $\tilde{A}$ 和 $\tilde{B}$ 为模糊数。实现无样本参量 y 的认知,意味着建立对应关系 f ,并利用建立的对应关系式 f 对任意输入 x 的输出 y 进行预测或状态评估。

通过以上描述可以发现,在无样本参量认知问题中,可获得的信息为:与无样本参量相关联的相同时刻(与设备状态一致)的其他可测过程参量的数据信息及相同时刻下无样本参量的不精确取值(如模糊数表征的可能取值),以及根据专家知识确定的这些过程参量与无样本参量之间具有的不精确、不确定对应关系。显然,无样本参量认知问题就是利用这些包含着不精确性和不确定性的信息,构建对无样本参量的认知模型问题。

我们可以从参量的取值和置信度两个方面来理解信息的不精确性和不确定性:所谓不精确性是针对参量取值是否精确来说的,而不确定性是针对参量取某数值的可信程度来说的。

通常意义下的不确定性称为随机性。为了研究随机性,人们提出了随机和概率等相关概念,由此诞生了概率论^[1]。在很长时间内,概率理论几乎成为处理不确定性问题的唯一工具。直到 20 世纪下半叶,随着对众多不确定性问题的深入研究,人们逐渐发现不确定性与概率理论不能完全等同,主要体现在以下两个方面:一是不确定性具有不同的内容和表现形式,具有多维特性,需要与概率论完全不同的理论来表达和处理^[2-4];二是概率理论中采用的推理方法不能适用于所有的不确定性问题。比如,在概率理论中,概率测度采用相加性原理^[5,6]。根据该原理,如果一个命题为真的程度为 p ,那么,我们必须以 $(1-p)$ 的程度去相信其反命题。在许多情况下,这是不合理的。

为了更好地解决各类不确定性问题, Dempster 于 1967 年首先提出了证据理论, 随后经 Shafer 系统完善, 故又称为 Dempster-Shafer 理论, 简称 DS 理论^[7]。由于证据理论把点值函数形式推广到集合函数形式, 其研究的基本对象由点值变为集合和区间数, 因而证据理论具有直接表达“不确定”和“不知道”的能力。证据理论与概率理论具有较好的兼容性: 当集合和区间数退化为点值时, 证据理论也就相应的退化为概率理论。举例来说, 在概率论中, 将随机变量 x 的可能取值定义为基本事件集 $\Omega_x = \{x_1, x_2, \dots, x_N\}$, 其中, N 为基本事件个数; 那么, Ω_x 上的一个概率分布 $P: \Omega_x \rightarrow [0, 1]$ 表征了一条用于反映随机变量的可能取值的不确定性信息, 如: $P(x_i) = 1/N (i=1, 2, \dots, N)$ 表示随机变量以等可能的概率取值。而在证据理论中, 基本事件集 Ω_x 称为变量 x 的识别框架, 证据是定义在识别框架上的一个集函数: $m: 2^\Omega \rightarrow [0, 1]$ (值得注意的是, 如果识别框架为某一实数区间, 则 2^Ω 为该实数区间的所有子区间的集合)。显然, 证据 m 是概率分布的拓展。当我们对变量 x 的取值完全“不知道”(即无知)时, 可以用证据 $m(\Omega_x) = 1$ 来表征, 而概率论则不能有效表征这种完全无知的情况。此外, 由于将 Ω_x 拓展为 2^Ω , 证据不但能够表征不确定性, 还能有效表征不精确性。现实中, 由于所获信息通常为(或包含)不确定信息(Uncertainty)和不精确信息(Imprecise)。那么, 就需要一种理论框架来支撑这些不精确和不确定信息的表征及合成。而证据理论正是这样的理论框架。

在证据理论中, 可通过证据来表征不精确和不确定性信息(即证据表征)。下面的示例 1-1 说明如何用证据理论对不精确和不确定性信息进行描述。

[示例 1-1] 不精确、不确定信息的描述和表征

以 Ω_x 表征天气温度的可能取值, 假设我们可以获得以下信息用于描述今天的天气温度:

- (1) “今天温度可能为 25℃”;
- (2) “今天温度在 20℃~25℃之间”;
- (3) “今天温度可能在 20℃~25℃之间”。

显然, 第(1)条信息为精确的、不确定性信息; 第(2)条为不精确、确定性信息; 第(3)条为不精确且不确定信息。在证据理论中, 这种不精确、不确定信息均可以用证据进行有效表征。上述三条信息用证据可以分别表示如下:

$$m_1(\{25\}) = p_1, m_1(\Omega_x) = 1 - p_1;$$

$$m_2([20, 25]) = 1;$$

$$m_3([20, 25]) = p_2, m_3(\Omega_x) = 1 - p_2;$$

其中, p_1, p_2 表示了不确定程度, Ω_x 及其子区间表示不精确程度。

为了更直观地对证据理论进行进一步的了解, 下面我们再通过一个生活中的实例来形象说明如何用证据来表征不精确和不确定性信息, 以及如何对这些用于表征不精确和不确定性信息的证据进行证据融合的, 具体的有关证据的定义及证

据融合的相关计算将在随后 1.2 小节讨论。

[示例 1-2] 证据表征及证据融合举例

某夜发生了一次刑事案件,经警方初步排查,认为嫌疑人可能为: $\Omega_x = \{\text{张三, 李四, 王五}\}$ 。经警方初步调查,这三个嫌疑人在案发当天具有如下外部特征:张三,男,短发;李四,男,长发;王五,女,长发。一个目击证人到警察局举证,称其看见谋杀者遁入黑暗中,但只能确定谋杀者是男性。经警方取证发现,该目击证人每五天中有一天在酗酒,也就是说,该目击证人提供的证词只有 $1-1/5=0.8$ 的可信度。那么,可以构建一条证据用于反映嫌疑人的可能性(值得注意的是,我们不能将 0.2 的可信度分配给{王五},因为证词没有提到有关王五的任何信息):

$$m_1: m_1(\{\text{张三, 李四}\}) = 0.8, m_1(\Omega_x) = 0.2$$

警方经进一步侦查获得另一条独立的证据,认为嫌疑人有大约 0.6 的可能性蓄有长发。那么,可以构建第二条证据:

$$m_2: m_2(\{\text{李四, 王五}\}) = 0.6, m_2(\Omega_x) = 0.4$$

可以发现,此时构建的两条证据既表征了不精确信息也表征了是不确定信息。为了便于警方作出判断,需要将这两条证据进行融合得到一条最终的证据。根据经典证据融合法则(随后小节将会详细讨论证据融合法则,在此只需将其理解为“正交和”即可,证据融合过程见表 1-1 所示),可以获得以下融合证据 m_{12} :

$$m_{12}: m_{12}(\{\text{李四}\}) = 0.48, m_{12}(\{\text{李四, 王五}\}) = 0.12, m_{12}(\{\text{张三, 李四}\}) = 0.32, m_{12}(\Omega_x) = 0.08$$

表 1-1 基于经典证据融合法则的证据融合计算示意过程^{*1}

		m_1			
ϕ		{张}	{李}	{张, 李} 0.8	{王} {张, 王} {李, 王} Ω_x 0.2
m_2	ϕ				
	{张}				
	{李}				
	{张, 李}				
	{王}				
	{张, 王}				
	{李, 王} 0.6			{李} 0.48	{李, 王} 0.12
	Ω_x 0.4			{张, 李} 0.32	Ω_x 0.08

*1 表中空格处分配给其他交集的信度均为 0。

一旦获得融合证据 m_{12} , 就可以根据某一规则进行决策了(即证据决策)。在此,对于 m_{12} 中的 $m_{12}(\{\text{李四, 王五}\})$ 、 $m_{12}(\{\text{张三, 李四}\})$ 、 $m_{12}(\Omega_x)$ 三种情况,我们假设将信度平分给个人并将所有情况下个人所得的信度相加作为个人是嫌疑人的最终信度(该决策策略即后续小节中的 BetPm 概率分布的定义,常用于证据决策),那么有以下结论:

- 张三是嫌疑人的可能性为: $\text{BetPm}(\{\text{张三}\}) = 0.32/2 + 0.08/3 = 0.187$;
- 李四是嫌疑人的可能性为: $\text{BetPm}(\{\text{李四}\}) = 0.48 + 0.12/2 + 0.32/2 + 0.08/3 = 0.727$;
- 王五是嫌疑人的可能性为: $\text{BetPm}(\{\text{王五}\}) = 0.12/2 + 0.08/3 = 0.086$ 。

根据以上计算分析发现,分配给李四的信度最大(即 $\text{BetPm}(\{\text{李四}\})$ 概率最大),所以,我们有理由认定李四是真凶。

示例 1-1 和 1-2 让我们初步了解了证据理论如何表征不精确、不确定性信息,以及基于这些信息的证据融合过程(多证据的融合可以通过证据的两两融合来实现)。证据理论在处理不精确不确定信息时主要涉及证据表征、证据构建、证据融合以及证据决策(即基于 BetPm 概率分布进行决策)等过程。在证据理论范畴内研究无样本参量认知问题,同样涉及到这些概念,分析如下示例 1-3 所示。

[示例 1-3] 无样本参量认知建模

尽管无法直接获取待认知变量 y 的取值样本,但基于专家经验或机理知识(以下简称专家知识),以及证据融合法则,可以估计无样本参量“可能”取某一“大致范围”的数值,亦即可以基于专家知识和证据理论的方法,获取待认知变量的不确定性(如“可能”)、不精确性(如“大致范围”)信息的证据表征,包括无样本参量的可能取值 y_i 、不精确程度 σ_i 和其可靠性程度 p_i (具体方法详见第 7 章~第 9 章)。

举例来说,通过设计实验,使设备由低负荷运行到全负荷,并获得反映设备全负荷过程的可观测运行数据样本(x_i)。基于专家知识,假设无样本参量 y 与其输入 x 之间存在正相关关系,即当 x 由小变大时, y 亦由小变大。

首先采用模糊聚类的方法对输入数据 x_i 进行分析。假设输入数据样本空间被划分为 $1, 2, \dots, c$ 个类,我们以其中对应于低负荷和最高负荷的那两个类作为参考类,并对其类心所对应的无样本参量进行赋值,称为无样本参量的初始化。比如,对于归一化后值域为 $Y=[0, 1]$ 的无样本参量 y ,我们认为最低负荷参考类类心对应的 $y^1=0.05$,而最高负荷参考类类心对应的 $y^c=0.95$ 。一旦确定了最低负荷和最高负荷所对应类的无样本参量 y 的可能取值,其他类类心对应的无样本参量的可能取值 $y^k (k=2, 3, \dots, c-1)$ 可以根据其与参考类的类心距离确定其可能取值,例如距离最低负荷参考类类心越远,则其对应的 y^k 的可能取值越大。

其次,对于分属于各类的样本 x_i ,可以分别计算其隶属度 u_{ik} 。其中样本 x_i 隶属于各个类的隶属度 u_{ik} 定义为样本与某类类心距离相关的函数,距离某类类心的距离越近,其与该类的隶属度就越大。进而,将其作为权重系数,通过加权聚类中心所对应的 y^k 取值计算对应于 x_i 的可能取值 y_i ,即: $y_i = \sum_{k=1}^c u_{ik} y^k$ 。这样,我们就获得了所有输入样本 x_i 对应的可能取值 y_i 。

第三,不精确性 σ_i 的确定。我们可以计算以 y_i 为中心某个邻域内的样本方差将其作为 σ_i ,并用于构建区间 $[y_i - \sigma_i, y_i + \sigma_i]$ 来描述“无样本参量取值大致在 y_i 附近”这种不精确性信息。

最后,无样本参量的可靠性程度 p_i 可以基于仪表(如用于测量 x 的仪表)的可靠性系数或基于专家知识获得。在没有专家知识的情况下,我们一般认为我们获得 y_i 是可靠的,即有 $p_i=1$ 。

当我们采用上述方法获得了无样本参量的可能取值 y_i 、不精确程度 σ_i 和其可靠性程度 p_i ,可以构建证据。比如: $m_i([y_i - \sigma_i, y_i + \sigma_i]) = p_i, m_i(Y) = 1 - p_i$ (详见第6章~第9章)。

图1-3中给出了一维空间中证据构建的示例。图中每个小圆圈的圆心当前位置为 (x_i, y_i) , 其中, y_i 为无样本参量的可能性取值; 平行于纵坐标的竖直线段表示无样本参量可能取值的不精确程度, 其长度为 $2\sigma_i$, 线段长度越短, 则其对应的无样本参量可能取值越精确; 圆圈直径大小表示当前无样本参量可能取值的可信程度, 直径越大可信度越高, 即不确定性越小。

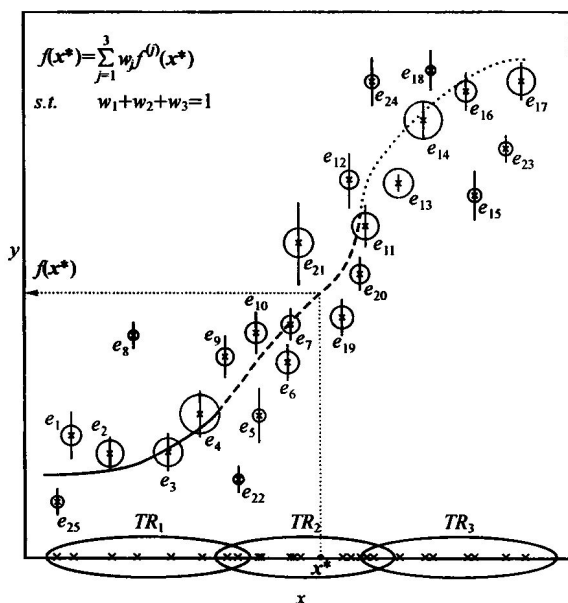


图1-3 基于证据理论的无样本参量认知示意图

一旦获得 m_i , 即可获得用证据形式表征的证据样本集 $TR = \{e_i = (x_i, m_i) \mid i = 1, 2, \dots, n\}$ 。图1-3中给出了容量为25的证据样本集 $TR = \{e_i = (x_i, m_i) \mid i = 1, 2, \dots, 25\}$ 。基于构建的证据样本集, 通过证据融合(类似于例1-2)可建立无样本参量认知模型 $f(x)$, 利用该模型可以计算或者预测其他证据, 亦即对于任意的 x , 可以估算其可能取值 y 、以及相应的不精确性 σ_i 和不确定性 p_i 。

针对工业过程的非线性和工况范围广的特点, 一般可采用多模型策略建立认知模型, 故可将样本集 TR 划分为多个子区域 $TR_j (j = 1, 2, \dots, c)$; 在每个子区域 TR_j 内建立局部认知模型 $f^{(j)}$, 通过加权(或融合)局部认知模型 $f^{(j)}$ 建立全局认知模型 $f(x)$, 而这种子区域的划分和确定与前面所说的子过程的获取是一致的。

图 1-3 中样本集被划分为三个子区域(可以为模糊划分),获得的全局认知模型 $f(x)$ 如实线所示。该全局模型有效反应样本的不精确性和不确定性:不精确性越小、可靠性越大的样本越靠近模型。

由于本书的研究范畴涉及证据理论,包括证据理论的相关理论研究和应用研究,在随后章节,将对证据理论的发展现状和存在的问题进行综述。

1.2 证据理论:基本概念及相关计算

1967 年, Dempster 给出了上下概率的概念,第一次明确给出了不满足可加性的概率^[11]。随后,他于 1968 年探讨了统计推理的一般化问题^[12],并提出了证据融合法则用于融合来自两个独立信息源的证据。继 Dempster 之后, Shafer 将 Dempster 的证据融合法则推广到一般情形,于 1976 年出版了名为《证据的数学理论》一书,该书的出版标志着证据理论的诞生^[7]。由于证据理论的独特优点,证据理论得到了广泛应用,如:信息融合^[13-16],专家系统^[17-21],模式识别及分类^[22-30],诊断和推理^[32-39],知识约简^[40-42],决策分析^[20,43-54],测试数据处理^[55],智能控制^[56],过程监测与预测^[57-60],风险评估^[61-68],环境影响评价^[69],图像处理^[70-75],地理信息系统^[76],安全分析^[77,78],回归分析^[97,80],等。

本小节的主要目的是为后续章节中的研究做理论准备。主要介绍证据理论中与本书研究相关的基本概念和定理。这些基本概念、定义、定理及符号主要参考于文献[7,81,82,83]。在这些文献的基础上,作者基于自己对证据理论的理解组织了本小节内容,并基于自己的理解给出了示意图及计算示例。为了表述简洁,相关定义和定理没有具体标明出处。在后续章节中,没有标明出处的定义、定理、命题均为专著提出的内容。

1.2.1 证据表征

对于一个问题,人们所能够认识到的可能结果可以用集合 Ω 表示,称 Ω 为识别框架(Frame of Discernment)。人们所关心的任一命题都对应于 Ω 的某一个子集。识别框架 Ω 的选取取决于人们的认识水平。在 TBM 中,设 $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$ 表示变量 ω 的有限的可能取值,一个当事人 Ag 对变量 ω 取得实际值 ω_0 的信度被称为基本信度分配(Basic Belief Assignment, BBA)。基本信度分配定义如下。

[定义 1-1] 设 Ω 为识别框架,如果集函数 $m: 2^\Omega \rightarrow [0,1]$ (2^Ω 为 Ω 的幂集),满足:

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (1-3)$$

则称 m 为识别框架 Ω 上的基本信度分配,也称为质量函数(Mass Function)。

通常,将人们用来描述某个问题的一个质量函数(或随后定义的与质量函数一

一对应的其他函数,如:信度函数,等)称为一条或一个证据。

对于 Ω 的任意子集 A , $m(A)$ 表示命题 A 为真的基本可信度,或分配到 A 上的质量。如果 $m(A)>0$,称集合 A 为 m 的焦元。 $m(A)$ 反映了对 A 本身的信度大小,即 $\omega_0 \in A$ 的信度。由于证据不充分,不能进一步赋予 A 的任意真子集信任程度。根据定义1-1,“完全无知”则可以表示为 $m_\Omega = m(\Omega) = 1$,称 m_Ω 为空质量函数,或无知证据。根据焦元类型不同,质量函数一般有几类:

- 标准质量函数(标准证据): 焦元不包括空集 Φ ;
- 非标准质量函数(非标准证据): 焦元包括空集 Φ ;
- 教条质量函数(教条证据): 焦元不包括 Ω ;
- 非教条质量函数(非教条证据): 焦元包括 Ω ;
- 类质量函数(类证据): 有且仅有一个焦元;
- 贝叶斯质量函数(贝叶斯证据): 所有焦元的基数均为1;
- 嵌套质量函数(嵌套证据): 所有焦元相互嵌套;
- 空质量函数(无知证据): 仅有一个焦元为 Ω ;
- 全冲突质量函数(全冲突证据): 仅有一个焦元为 Φ ;
- 简单质量函数(单证据): 最多两个焦元,如有两个,则 Ω 为其中之一。

图1-4给出了典型质量函数的结构示意图。

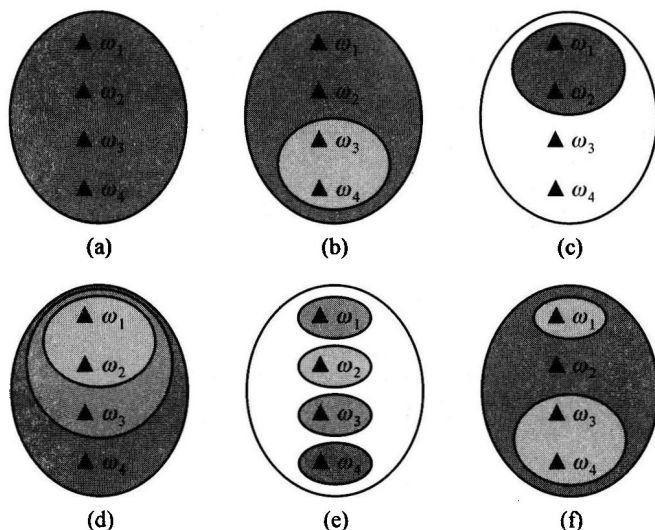


图1-4 不同类型质量函数示意图

(a) 空质量函数, (b) 简单质量函数, (c) 类质量函数, (d) 嵌套质量函数, (e) 贝叶斯质量函数, (f) 一般质量函数, (a)(b)(d)(f) 非教条质量函数, (c)(e) 教条质量函数。

(注: 图中空白为没有分配质量, 阴影为有分配质量, 颜色深浅代表质量大小)

在TBM中, 条件 $m(\Phi)=0$ 是不被定义1-1要求的。实质上, 在开放边界假