



石油企业供应商选择方法 ——范例推理技术及应用

赵 凯 编著

中国石化出版社

[HTTP://WWW.SINOPEC-PRESS.COM](http://www.sinopec-press.com)

内 容 提 要

本书基于模仿人类思维能力的范例推理系统,结合相关数据挖掘方法,解决客观确定属性权重、准确评价目标范例与源范例之间的相似度,提取选择结果中潜在的规则等难点问题,从而提高石油企业供应商选择决策的准确性和效率。

本书可供石油生产企业管理人员参考,也可作为大专院校相关专业教学参考书。

图书在版编目(CIP)数据

石油企业供应商选择方法:范例推理技术与应用/
赵凯编著. —北京:中国石化出版社, 2010. 3
ISBN 978 - 7 - 5114 - 0329 - 2

I. ①石… II. ①赵… III. ①石油工业 - 工业企业管理: 供销管理 IV. ①F407. 226. 5

中国版本图书馆 CIP 数据核字 (2010) 第 038155 号

未经本社书面授权,本书任何部分不得被复制、抄袭,或者以任何形式或任何方式传播。版权所有,侵权必究。

中国石化出版社出版发行

地址:北京市东城区安定门外大街 58 号

邮编:100011 电话:(010)84271850

读者服务部电话:(010)84289974

<http://www.sinopec-press.com>

E-mail: press@sinopec.com.cn

北京宏伟双华印刷有限公司印刷

全国各地新华书店经销

*

787 × 1092 毫米 16 开本 9.5 印张 122 千字

2010 年 4 月第 1 版 2010 年 4 月第 1 次印刷

定价:26.00 元

前言

供应商的选择关系到石油生产和供给的安全,能够保障石油企业及时、敏捷、高效地为国家建设和国民生活提供能源。随着全球化浪潮和信息技术的发展,供应商选择方法的优劣对提高石油企业的竞争力起到了非常重要的作用。本书基于模仿人类思维能力的范例推理系统,结合相关数据挖掘方法,解决客观确定属性权重、准确评价目标范例与源范例之间的相似度,提取选择结果中潜在的规则等难点问题,从而提高石油企业供应商选择决策的准确性和效率。

本书的主要内容:第一章为绪论部分,介绍本书的背景、研究目标与意义、研究方法和创新点等。第二章对现有供应商选择的主要方法和国内外研究成果进行综述,并且对本书研究的主要方法——范例推理(CBR)的基本思想和重要理论进行了总结。第三章至第五章是本书的主干部分,围绕基于范例推理的供应商选择决策过程中存在的三个重要问题提出相应的数据挖掘解决方法,以提高范例推理系统的决策效率和准确率,这也是本书的创新性成果之所在。探索建立基于数据挖掘的范例推理系统,提出综合三个创新点的实例,将研究成果实际应用于石油企业的供应商选择决策之中。最后对研究结论进行总结。

本书可供石油生产企业管理人员参考,也可作为大专院校相关专业教学参考书。

在本书的编写过程中,得到了中国石油大庆油田有限责任公司及其他油田的领导和同事们的鼓励与支持,特向他们表示感谢!

本书得到了清华大学博士生导师杨炘教授、北京科技大学经济管

理学院副院长博士生导师高学东教授、博士生导师王道平教授以及土木与环境工程学院博士后于昕的大力支持，特在此致以衷心感谢！

由于作者研究水平有限，书中难免有疏漏与不妥之处，恳请有关专家和读者批评、指正。

目 录

第一章 绪论	(1)
第一节 研究背景	(1)
第二节 本书结构	(4)
第二章 相关研究综述	(9)
第一节 供应商选择的主要方法	(9)
一、线性权重法	(9)
二、成本方法	(10)
三、数学规划方法	(11)
四、其它方法	(13)
第二节 范例推理的理论综述	(15)
一、CBR 研究的历史与现状	(15)
二、CBR 中范例的表示方法和工作过程	(16)
三、CBR 的特点及发展方向	(18)
第三节 基于数据挖掘的范例推理系统	(20)
一、范例推理中基于数据库的知识发现技术	(21)
二、范例库上数据挖掘的主要步骤与目标	(22)
三、范例库上数据挖掘的重要方法与技术	(24)
四、范例推理各阶段的关键难点	(26)
五、本章小结	(27)
第三章 基于信息熵的范例特征项权重确定方法	(28)
第一节 特征项赋权的一般方法	(28)
一、主观赋权方法	(29)
二、主客观综合赋权方法	(31)

三、客观赋权方法	(33)
第二节 决策树与信息熵	(36)
一、决策树分类算法	(36)
二、信息增益和熵	(38)
第三节 范例特征项赋权的熵方法	(41)
第四节 应用实例	(43)
一、背景与数据	(43)
二、实验步骤	(54)
三、结果分析	(58)
第五节 本章小结	(59)
第四章 基于 k -prototypes 聚类的范例相似度评价模型	(60)
第一节 范例检索的相似性度量	(60)
一、范例相似性度量	(60)
二、常用的相似性度量函数	(63)
三、范例检索的策略和方法	(66)
第二节 聚类差异度计算和 k -均值算法	(67)
第三节 基于 k -prototypes 聚类的多属性相似度综合评价	(71)
第四节 应用实例	(74)
一、实例背景	(74)
二、实验步骤	(75)
三、分析结果	(81)
第五节 本章小结	(85)
第五章 基于 BP 神经网络的供应商选择 CBR 系统	(86)
第一节 基于神经网络的范例推理系统	(86)
一、范例推理与神经网络	(86)
二、基于神经网络的范例推理应用	(87)
三、BP 算法	(89)
第二节 神经网络规则提取综述	(93)

一、基于结构分析的方法	(94)
二、基于性能分析的方法	(98)
三、规则抽取研究中存在的问题	(102)
第三节 构建基于 BP-CBR 的供应商选择系统	(103)
一、基于数据挖掘技术的 CBR 系统框架结构	(103)
二、基于神经网络构建 CBR 系统	(105)
三、基于 BP 算法的石油企业供应商选择 CBR 模型	(108)
第四节 应用实例	(114)
一、实例背景	(114)
二、实验步骤	(115)
三、结果分析	(122)
第五节 本章小结	(123)
第六章 结论	(125)
附 录 A 公司年度合格建筑承包商评估表	(129)
参考文献	(131)

第一章 绪 论

第一节 研究背景

石油是目前使用最普遍的不可再生能源，也是关系国民经济命脉和国家经济安全的重要战略物资。目前，世界各国对石油的争夺和控制日趋激烈，使全球石油资源形势呈现出许多新趋势。一是国际石油资源争夺政治色彩日益浓厚；二是大型石油公司主导国际石油市场的格局已经形成，形成了一批如 ExxonMobil、BP 等全球化超级石油集团，对世界经济产生日益重大的影响；三是石油资源争夺已成为国际军事斗争的重要目标；四是亚洲日益增长的能源需求将产生深远的政治影响；五是国际油价大幅波动，严重影响世界经济发展^[1]。2008 年 7 月 11 日，国际原油价格暴涨，欧美期货盘中突破每桶 147 美元，创历史新高。从总体上看，当前已进入高油价时代，低油价的行情已经很难再出现，这对世界经济，特别是发展中国家的经济发展将产生较大影响。预计今后 5 至 15 年，由于美元贬值、石油需求旺盛、富余产能减少、OPEC 期望值提高等因素，国际原油价格仍将在高位运行。

近年来，我国石油供需矛盾加剧，且对外依存度不断攀高。“十五”期间，随着我国经济的快速增长，能源供需的矛盾迅速升级。预计 2020 年以后，国内原油产量将由顶峰期向下滑动，而国内需求量仍将继续增加，因而，石油供需矛盾将更加突出。此外，中国石油战略储

备体系尚不完善, 应对石油风险的能力仍然不足。

在现有体制下, 国家对石油产品实行统购统销, 这就决定了其供应链不存在“销售”环节这一重要特点^[3]。因此, 石油企业的供应商所承担的责任和产生的作用更加重要。显而易见, 供应商, 如建筑承包商、物资供应商、技术承包商等, 作为最重要的合作伙伴, 在企业资质、技术人员比例、工程优良率、产品质量、供货、产品研究与开发能力等方面都能对石油企业降低生产成本、提高生产效率及实现可持续发展等产生重大影响, 同时对供应链中各个节点企业的核心竞争力也会产生重要的影响。

如今, 全球化的竞争环境和信息化浪潮促进了企业动态联盟的产生, 同时也改变了竞争的方式, 使竞争不再是单个企业的竞争而是整个供应链的竞争。这促使了传统的管理模式向供应链管理模式的转变^[4]。在供应链管理模式下, 联盟成员与核心企业之间不再是建立在物质交换基础上的简单的货物买卖关系。供应链管理强调的是一种协调、集成的思想, 是一种建立在各自企业业务核心之上的强强联合。供应链中的各企业是一个利益共同体, 通过优势互补和协同效应产生企业独立时所不能产生的优势, 也就是通过建立供应商关系双方都能从提高生产率、节约资源、降低成本中获益, 同时创造更大的顾客价值^[5]。在全球贸易和敏捷供应链环境^[6]下, 对供应商的选择是企业的一个重要决策, 如何进行供应商的评价、筛选, 对企业的健康成长有重要的意义。

根据 1982 年 11 月美国《采购世界月刊》对 1280 个采购部门所作的调查显示, 采购部门的职责按其重要性来排列, 其前九种重要跟踪的顺序为: ①评估现有的供应商; ②选择及开发新的供应商; ③安排采购及交货日期; ④谈判采购合约; ⑤从事价值分析工作; ⑥自制或采购(外包)的决策; ⑦制定运输方式; ⑧控制存货; ⑨租赁或买断的决策。由此可见, 选择合格的供应商是采购部门最为重要的职责。

本书的研究基础来源于国家自然科学基金项目“面向敏捷供应链的

知识服务体系及管理策略研究”(70872010)对供应商选择的研究成果,研究背景来源于我国大型石油企业的供应商管理实际。通过应用一系列数据挖掘和机器学习的算法,如基于范例的推理,模糊神经网络等等,对决策规则的学习获得成功的决策规则,解决供应商选择决策过程中的关键问题。对供应商的资质进行审核,绩效进行评价,以获得具有预测性的规则,协助供应商后续进行改进,以求获得能够应用于实际商业活动的供应商选择的解决方案,最终依据方法的研究成果,完成基于规则挖掘的专家系统设计。

供应商的选择关系到石油生产和供给的安全,对保障石油企业及时、敏捷、高效地为国家建设和国民生活提供能源具有重要意义。石油企业的供应商,由于其所承担的角色和重要责任,越来越受到企业管理层的重视^[7]。因此,供应商选择方法的优劣,对提高石油企业的竞争力可以起到非常重要的作用。随着全球化浪潮和信息技术的发展,传统的打分式的选择方法由于存在对不确定性信息的处理能力不足,结果不易于解释等缺陷,已经难以适应新竞争环境的需要。本书的研究目标就是通过应用模仿人类思维的范例推理系统,基于数据挖掘方法,提取重要程度高的特征向量和客观确定权重,准确评价目标范例与源范例之间的相似度,并且提取选择结果中潜在的规则,以提高决策的可解释性和预测能力。本书的最终目的是将研究成果应用于企业的实际供应商选择决策中,从而增强石油企业的核心能力和可持续发展能力。

本书研究的意义包括以下几个方面:

(1) 从理论研究角度。国家自然科学基金项目“面向敏捷供应链的知识服务体系及管理策略研究”(70872010)为本书提供了资助。本书基于数据挖掘方法,分别对范例推理的关键问题提出新的解决方法,包括:①提取特征向量及赋权,以解决属性特征重要性的客观确定问题;②建立能够综合处理多种性质属性的相似度评价模型,以解决属性多种语义表达的综合评价问题;③基于神经网络方法,对范例库进

行修正,通过分析石油企业供应商的历史信息,挖掘数据中包含的模式信息,最终建立基于范例推理的石油企业供应商选择系统。

(2) 从应用创新角度。本书的目标是提出基于范例推理的石油企业供应商选择系统。与传统的多属性打分评价,或者运筹学方法,例如数学规划等选择决策方法最大的不同是:决策信息的表达形式多样,覆盖面广,决策过程的客观性强,决策方法具备学习能力和处理不确定信息的能力,决策结论基于历史数据分析,并且表现为一组规则描述的形式,而不是一个简单的分数,因此也具备更好的可解释性和预测性。根据这些规则,企业能够对供应商的行为进行实时监控,并协助供应商对其薄弱环节进行改进。

(3) 从实际应用角度。本书的研究获得了国内多家大型石油生产企业管理人员的鼎力相助,并获得了大量的第一手资料。针对石油企业建筑承包商资质审核问题的背景,对本书研究的基于规则挖掘的供应商选择方法进行应用,并结合实际进行案例分析,验证新方法的可行性和可操作性。最后,提出石油企业供应商选择的整体解决方案,建立基于范例推理的决策系统,从而为石油企业供应商管理提供有效支持。

第二节 本书结构

本书的研究内容主要包括以下几点:

(1) 提出问题。对石油企业供应商管理,以及传统供应商选择方法的国内外研究成果进行综述。分析在全球贸易和电子商务环境下,石油企业供应商选择所面临的关键问题,论述基于数据挖掘的范例推理系统解决这些问题的必要性。

(2) 范例库的组织及特征赋权。基于决策树和信息熵方法,合理组织范例库,并从中提取特征向量及确定权重,以解决范例多属性特

征项的重要性的客观评价问题。

(3) 建立相似度评估模型。基于 k -prototypes 聚类算法, 以距离为度量标准, 建立多种性质特征属性的相似度评价模型, 以解决属性多种语义表达的综合评价问题。

(4) 构建供应商选择系统模型。基于 BP 神经网络方法, 对范例库进行修正, 通过分析石油企业供应商的历史信息, 挖掘数据中包含的模式信息, 最终建立基于范例推理的石油企业供应商选择系统。

(5) 综合本书的三个创新点, 探索设计集成化的基于数据挖掘的范例推理系统, 用于石油企业供应商选择决策, 并以建筑承包商的选择为实例, 开发基于 BP-CBR 的供应商选择系统。

综上所述, 本书的组织结构见图 1.1 所示。其中, 第一章为绪论部分, 介绍本书的背景、研究目标与意义、书的研究方法和创新点等。第二章对现有供应商选择的主要方法和国内外研究成果进行综述, 并且对本书研究的主要方法—范例推理(CBR)的基本思想和重要理论进行了总结。第三章至第五章是本书的主干部分, 围绕基于范例推理的供应商选择决策过程中存在的三个重要问题提出相应的数据挖掘解决方法, 以提高范例推理系统的决策效率和准确率, 这也是本书的创新性成果之所在。探索建立基于数据挖掘的范例推理系统, 提出综合三个创新点的实例, 将研究成果实际应用于石油企业的供应商选择决策

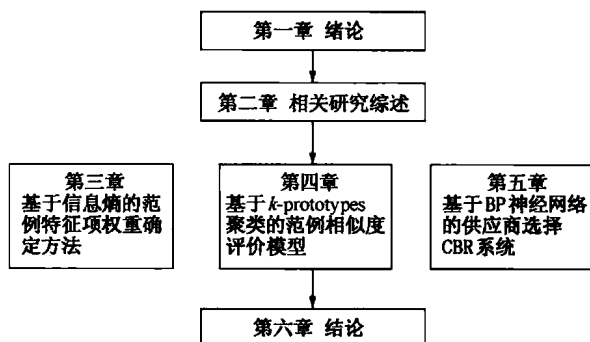


图 1.1 本书组织结构图

之中。最后对本书得到的结论进行总结。

由于石油企业供应商选择决策的重要性和复杂性,本书选择智能化的范例推理系统解决决策中存在的难点问题。范例推理(Case - Based Reasoning, CBR)是由目标范例的提示而得到历史记忆中的源范例,并由源范例来指导目标范例求解的一种策略,它是一种重要的机器学习方法。CBR 是区别于基于规则推理的推理和学习模式,它是指借用旧的事例或经验来解决问题、评价解决方案、解释异常情况或理解新情况。CBR 兴起的主要原因是传统的基于规则的系统存在诸多缺点,如知识获取的瓶颈问题,对于处理过的问题没有记忆,导致推理效率低下,处理领域中例外事件格外困难,整体性能较为脆弱等等,而 CBR 恰好能很好地解决以上问题。

虽然范例推理系统比传统专家系统较为容易地获取知识,但范例推理系统也存在一定程度的知识获取瓶颈问题,如范例、修正知识、相似性评估等知识的获取问题,无疑 CBR 中的知识获取瓶颈问题也影响了 CBR 系统的性能。通过在范例推理中引入数据挖掘技术,针对 CBR 系统中的多种数据源进行数据挖掘,可以提高知识获取的自动化程度,提高系统的性能,加快智能系统的开发周期。数据挖掘的目标就是在大量的数据中找出许多人们难以发现的模式,而这些隐含的模式恰恰是范例推理系统所需要的各种知识。在 CBR 系统中引入数据挖掘技术是非常必要且可行的。本书重点阐述在 CBR 中所涉及的数据挖掘技术,其研究对解决知识工程中的知识获取瓶颈问题以及对拓宽计算智能的研究及其应用都有着较高的理论价值与实际意义。

本书对基于范例推理的石油企业供应商选择的三个难点问题的研究均为五个步骤:提出问题,文献总结,提出新方法,数据实验,以及结果分析。主要的研究方法包括:使用对比分析方法发现传统方法存在的不足,提出改进方案;使用数学建模提出基于数据挖掘技术解决问题的新方法;使用实验分析方法证明提出方法的有效性;采用实

证研究方法,验证提出方法的实用性和可行性。基于以上研究方法,本书主要是研究如何提高石油企业供应商选择的准确性和效率,能够取得的创新点包括以下三点:

(1)在范例库组织和检索过程中,由于范例库中优秀供应商范例的评价属性指标方向不同,即有的是取值越高越好的正向指标,有的是取值越低越好的负向指标,因此主要难点是如何合理地组织范例库,以及如何客观确定各个特征属性的重要性。本书基于决策树和信息熵方法,组织高效的范例库结构,提取重要程度高的特征向量并客观确定权重。

(2)在范例检索过程中,由于范例的特征属性的性质多种多样,例如有些是数值型,有些是文本型,有些是布尔变量,有些是多等级变量,因此如何综合评价源范例和目标范例之间多特征属性的相似度是检索的重要难点。本书基于 k -prototypes 聚类算法,分别以距离和模作为数值和文本属性的度量标准,建立多种性质特征属性的相似度评价模型,解决属性多种语义表达的综合评价问题。

(3)在范例更新和维护的过程中,由于范例推理过程中范例库中存储范例的广泛代表性和检索效率之间是矛盾关系,因此要求采用一种具有学习能力启发式决策方法。又由于传统的供应商选择方法的结果缺乏解释性,不能从决策结论中获取具有预测性的参考信息,因此需要采用能够通过分析石油企业供应商的历史信息,挖掘历史数据中包含的模式信息的智能化方法。本书基于 BP 神经网络方法,自动地对范例库进行修正和维护,从决策结果中提取规则,最终建立基于范例推理的石油企业供应商选择系统。

通过创新点(1)属性权重的计算结果与神经网络属性权重之间的对比,得到权重的优化。通过创新点(2)的分类结果,直接推理出神经网络的属性,得到训练模型。以上两个结果可以直接用于创新点(3)——构建石油企业供应商选择系统。因此,创新点(1)、(2)是创新点(3)的先决条件。见图 1.2。

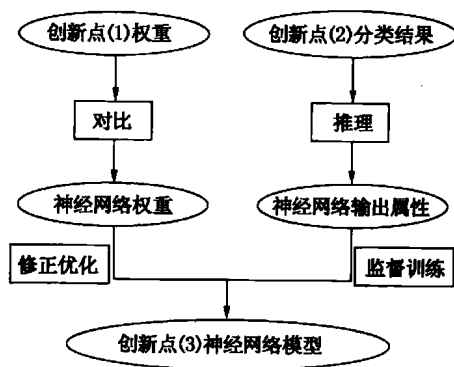


图 1.2 三个创新点之间的逻辑关系图

第二章 相关研究综述

第一节 供应商选择的主要方法

供应商选择方法的研究大致经历了三个发展阶段：定性方法、定量方法、定性与定量相结合的方法。刘晓^[8]采用分类的方法，系统地介绍和分析了自 1973 ~ 2002 年中公开发表的有关供应商选择的一些主要模型和方法，包括线性权重法、成本法、数学规划法、智能方法等。

一、线性权重法

线性权重法 (Linear Weighting Models, LWM) 是一种广泛应用于解决单资源问题的方法。它的基本原理是给每个准则分配一个权重，权重越大表明其越重要。供应商的积分为该供应商各项准则的得分与其权重乘积之和，积分最高者为最佳供应商。

层次分析法^[9] (Analytical Hierarchy Process, AHP) 是该类方法中最普遍应用的一种。它是美国运筹学家、匹兹堡大学教授 T. L. Saaty 于 1971 年提出的一种多准则决策方法，是一种定性与定量分析相结合的多准则决策分析方法，该方法简单、实用、有效，得到了广泛的运用。这种方法充分发挥人的主观能动性，在具有不确定性的环境下，依据决策者的经验、直觉和洞察力作出判断，把一些定性的因素以定量的形式表示出来。该方法可以考虑许多无法直接量化的因素，尤其是一

些对未来合作发展有长远意义的因素。因而, AHP 法被广泛应用于供应商选择^{[10][11]}和优先级评价^[12]等方面。

应用 AHP 方法选择供应商的基本步骤^[13]是: (1) 建立目标决策问题的层次结构模型; (2) 比较层次结构中供应商选择的准则, 两两比较并量化构成相应的判断矩阵, 采用特征向量法, 确定供应商选择准则的权重。由于此判断矩阵是通过人的主观判断得到的, 因此需要对判断矩阵的一致性问题进行检验、调整; (3) 针对每一个准则给出各个供应商的评价值, 计算出相应的权重; (4) 计算每一个供应商的综合权重, 并根据综合权重的大小作为最后选择供应商的依据。

AHP 法的缺点是它依靠专家的意见, 用 1~9 个标度表示, 加入个人的偏好, 主观因素较大, 易于出现判断矩阵的不一致性。另外, 在 AHP 法的应用过程中, 常常会有一些模糊和不精确因素, 对此一些学者尝试用模糊理论(Fuzzy Set Theory, FST)与 AHP 法相结合的方法来解决^{[14][15]}。另外, 文献^{[16][17]}则采用数据包络分析(Data Envelopment Analysis, DEA)与 AHP 法相结合的方法来解决供应商的选择问题。

二、成本方法

成本方法(Total Cost Approaches, TCA)是用来解决单一采购项目问题的一种常用方法, 其基本思想是对能够满足要求的供应商, 计算其采购成本, 包括销售价格、采购费用、运输费用等各项费用的总和, 通过对各个不同供应商的采购成本的比较, 选择成本较低的供应商的一种方法。

基于成本的供应商选择方法有很多种。Timmerman 提出用成本比率法^[18](The Cost Ratio, TCR)计算与成本有关的质量、运输、服务等项目的总成本来进行供应商选择。这种方法的思想是通过计算出每一项准则的成本占总成本的百分比来确定最终要选择的供应商。

Filip Roodhooft 和 Jozef Konings 于 1996 年提出了活动成本法^[19](Activity Based Costing, ABC)选择供应商, 此后 ABC 法得到一定程度