

文章编号: 1001-0920(2005)09-1047-05

虚拟企业伙伴选择的粗糙集方法

周庆敏^a, 殷晨波^b

(南京工业大学 a 信息科学与工程学院; b 机械与动力工程学院, 南京 210009)

摘要: 将粗糙集理论应用于虚拟企业合作伙伴选择中, 提出了基于粗糙集理论的虚拟企业伙伴选择的模型和方法。该方法根据各潜在伙伴企业的样本数据集建立决策系统, 以伙伴选择的评价指标作为属性, 从中挖掘出反映评价指标本质关系的重要属性以及伙伴选择知识规则。这些规则很好地描述了有限样本中所反映出的属性之间的本质特征, 运用这些规则可对伙伴选择数据库中的其他样本有效地进行伙伴选择。应用实例表明, 该方法是正确有效的。

关键词: 粗糙集理论; 决策系统; 虚拟企业; 伙伴选择

中图分类号: TP18 **文献标识码:** A

Rough Set Approach to Partnership Selection in Formation of Virtual Enterprises

ZHOU Qingmin^a, YIN Chenbo^b

(a School of Information Science and Engineering; b School of Mechanical and Power Engineering, Nanjing University of Technology, Nanjing 210009, China. Correspondent: ZHOU Qingmin, Email: mse@njut.edu.cn)

Abstract: Rough set theory is applied to partnership selection in formation of virtual enterprises. A rough set approach for partnership selection is presented. The decision system of partnership selection is established by using the sample sets of the potential partnership enterprises. The evaluation indexes are expressed as attributes of the decision system, from which the important attributes are extracted and the knowledge rules of partnership selection are induced. Those rules describe the essential characteristics between the attributes of the decision system. The partners are effectively selected from the other samples in partnership selection database by using those rules. An example of enterprises partnership selection is given to show the reasonableness and effectiveness of the proposed method.

Key words: Rough sets; Decision system; Virtual enterprises; Partnership selection

1 引言

虚拟企业是由多个分散异构企业组成的动态联盟, 其目的是为了实现在企业间技术和资源的共享、快速反应市场和增强竞争能力。在动态联盟中拥有主要核心资源并最先发现市场机遇的企业为盟主, 其他参与经营的企业为合作伙伴, 因为伙伴选择决定虚拟企业的总体竞争能力及其是否能成功运作的问题, 所以国内外将伙伴选择问题作为虚拟企业研究的一个重要内容^[1-4]。但目前虚拟企业合作伙伴的选择问题仍没有系统科学的方法。盟主面对一个市场机遇和众多可选择的潜在伙伴, 常常感到无所适

从。有时可供选择的伙伴及评价指标可能非常多, 为了分清主次, 通常设定相应的优先级, 甚至人为加权重, 往往造成选择中主观成分太多, 而影响伙伴选择的优化结果。

虚拟企业合作伙伴选择的过程实际上是一个具有不确定性的混合多属性决策问题。粗糙集理论作为研究不确定知识表达、学习、归纳的新型数学工具^[5], 其重要特点是不需要预先给定某些特征或属性的数量描述, 而直接从给定问题的描述集合出发, 在信息不确定情况下, 仍可挖掘出大量对决策有帮助的知识信息, 粗糙集理论提供了一整套比较成熟

收稿日期: 2004-10-18; 修回日期: 2005-03-11

作者简介: 周庆敏(1963—), 女, 河北唐山人, 教授, 从事计算智能、计算机集成制造等方面的研究; 殷晨波(1963—), 男, 江苏无锡人, 教授, 博士, 从事先进制造技术、车辆工程等方面的研究。

的样本学习式决策方法

本文提出了虚拟企业伙伴选择的粗糙集模型和方法,该方法通过各潜在伙伴企业的样本数据集建立决策系统,以伙伴选择的评价指标作为属性,从中筛选出能反映评价指标本质关系的重要属性,并从中挖掘出数据间的关系而形成伙伴选择知识规则。实践表明,这些规则很好地描述了有限样本中所反映出的属性间关系的本质特征。利用这些规则可对伙伴选择数据库中其他序列样本有效地进行伙伴选择。该方法中不需要任何先验知识和加权,得到的决策结果客观、合理。

2 伙伴选择问题的粗糙集模型

2.1 伙伴选择的决策系统表示方法

假定对于盟主某个已分解的任务,有 n 个愿意合作的企业构成合作伙伴候选集 $U = \{x_1, x_2, \dots, x_n\}$; 在伙伴选择过程中首先要确定影响伙伴选择的因素,即评价指标。通常,盟主选择各方面与自己互补的企业为伙伴,所以评价指标没有统一的标准。一般来说,评价指标主要包括各候选企业的生产能力、质量管理能力、制造成本、财务状况、信誉、地域等定量或定性的影响因素,这些影响因素作为条件属性集 $C = \{a_1, a_2, \dots, a_m\}$; $D = \{d\}$ 为伙伴选择决策属性集。伙伴选择的决策系统可表示为 $S = (U, C, D)$, 且 $C \cap D = \emptyset$ 。

2.2 伙伴选择决策系统的核与约简

决策系统的核与约简是粗糙集理论中两个重要概念。粗糙集理论认为知识是基于对对象分类的能力,分类的过程是将相差不大的对象分为一类,它们的关系是不可分辨关系,也称等价关系。约简是能保证决策分类质量的最小属性集合,一族等价关系可能有多个约简;核是不能去除的最重要的属性集合,是构成任何一个约简的必须成分,全部约简的交集定义为核,即

$$\text{core}(P) = \bigcap \text{red}(P). \quad (1)$$

其中: $\text{core}(P)$ 为伙伴选择决策系统的一族等价关系 P 的核, $\text{red}(P)$ 表示 P 的所有约简。对于一个伙伴选择决策系统可能存在多个相对约简,人们总是关注具有最少属性的约简,即最小约简,因为可通过最小约简得到决策系统的最小约简的规则集。但找出一个决策系统的最小约简已被证明是 NP-hard 问题^[6]。而决策表的属性核包含在所有约简中,所以找出属性核是约简计算的基础。本文通过构造决策系统的可辨识矩阵和分辨函数,得到伙伴选择决策系统的核和约简。

2.3 可辨识矩阵和分辨函数

数学家 Skowron 提出的可辨识矩阵实际上是

一种信息表示技术,对于伙伴选择决策系统 $S = (U, C \cup D)$, $C = \{a_j | j = 1, 2, \dots, m\}$, $D = \{d\}$, 若用 $M_D(i, j)$ 表示可辨识矩阵中第 i 行 j 列的元素,则可辨识矩阵定义为

$$M_D(i, j) = \begin{cases} \{a_k | a_k(x_i) \neq a_k(x_j)\}, & d(x_i) \neq d(x_j); \\ 0, & d(x_i) = d(x_j). \end{cases} \quad (2)$$

从定义可以看出,可辨识矩阵是一个依主对角线对称的矩阵,且对角线上的元素为零。核是可辨识矩阵中所有单个元素组成的集合^[7]。由可辨识矩阵唯一确定了一个分辨函数 F , F 实际上是一个布尔函数,对于每个属性 $a \in C$, 若 $a(x, y) = \{a_1, a_2, \dots, a_m\} \cap \emptyset$, 指定布尔函数 a_1, a_2, \dots, a_m , 采用 $a(x, y)$ 表示;分辨函数定义为

$$F = \bigwedge_{(x, y) \in U \times U} a(x, y). \quad (3)$$

研究发现,分辨函数的极小析取范式中的所有合取式是条件属性 C 的所有决策属性 D 简化集^[7],也就是说,每个合取式对应一个约简后伙伴选择的决策规则。

2.4 规则的置信度

对于伙伴选择的决策规则,利用粗糙隶属函数的值来表示每条决策规则的置信度^[8]。设决策规则 $r_{ij}: \text{des}(X_i) \rightarrow \text{des}(Y_j)$, 其中: X_i 和 Y_j 分别代表 U/C 和 U/D 中的各个等价类, $Y_j \cap X_i = \emptyset$, $\text{des}(X_i)$ 和 $\text{des}(Y_j)$ 分别表示 X_i 和 Y_j 对于各条件属性值和各决策属性值的特定取值;规则 r_{ij} 的置信度为

$$\mu(X_i, Y_j) = \frac{\text{card}(Y_j \cap X_i)}{\text{card}(X_i)}. \quad (4)$$

式中 $\text{card}(\bullet)$ 表示取集合中元素的个数,可以看出, $0 < \mu(X_i, Y_j) \leq 1$ 。当 $\mu(X_i, Y_j) = 1$ 时, r_{ij} 是确定的;否则 r_{ij} 是不确定的; $\mu(X_i, Y_j)$ 越大,表示规则 r_{ij} 的置信度越大。值得一提的是,粗糙集的规则置信度是客观计算的,因此避免了人为给定等主观因素的影响。

2.5 伙伴选择粗糙集模型

基于粗糙集的伙伴选择模型如图 1 所示。盟主企业在分析自身的核心资源后,根据联盟战略选择相应的合作伙伴评价指标。在确立了评价指标后,以这些评价指标作为伙伴选择决策系统的条件属性,每个潜在伙伴成员对应决策系统中的一条记录。每一列表示属性和属性值。因为粗糙集理论只能处理离散数据,所以伙伴选择决策系统中的连续属性值需要进行离散化处理。离散化处理后,决策系统进行合并整理,使其不含相同属性及属性值的重复成员。

通过构造可辨识矩阵和分辨函数计算核和所有的约简集, 一个伙伴选择决策系统可能同时存在几个约简集。不同的约简集可产生不同的决策规则, 在实际应用中, 很难判定哪个约简集导出的决策规则更有用, 因此以所有约简集为初始节点, 建立层状节点网络模型, 对于各节点, 提取大于置信度阈值的伙伴选择规则。置信度将规则分成确定性和不确定性决策规则, 经规则合成后写入规则集中。对规则集中的规则进行冗余性检验和多义性检查, 避免规则的重复和不一致性。利用粗糙集方法挖掘的规则集很好地反映出评价指标属性间的本质特征, 可运用它对伙伴选择数据库中其他序列样本有效地进行伙伴选择。

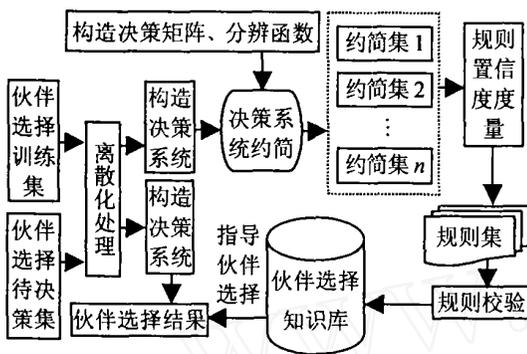


图 1 伙伴选择粗糙集模型

3 基于粗糙集模型的伙伴优化选择方法

3.1 参数的选择

当盟主识别出市场机遇后, 通过分析该产品的基本构成, 将任务分解成几个模块, 根据自身的制造能力、设计能力和经济效益等, 确定哪些模块应该由自己完成, 哪些需要寻找合作伙伴, 在众多愿意进行合作的伙伴中挑选出可供评价和选择的参数, 构造评价指标体系, 然后进行评价指标多属性决策, 优选合作伙伴。为了使评价结果更加具有代表性和客观性, 必须尽可能地建立完整的评价指标体系。伙伴选择的评价指标很多, 如: 响应时间、生产质量、产品报价、技术服务等, 它们是国际市场竞争成功的关键因素^[2,3]。同时, 财务信誉、信息化能力、企业的管理水平与文化、环境地理因素、机遇实现能力及后勤保障因素等也是影响虚拟企业联盟成败的关键。在实际操作中, 可根据联盟目标和盟主企业特点确立评价指标, 以这些评价指标作为伙伴选择决策系统的属性集, 在属性集中, 有些是定量的, 而有些是定性的。粗糙集理论处理伙伴选择决策系统时, 如果定量评价指标属性值为连续的, 则必须经过离散化, 寻找最优的离散化方法问题目前仍处于研究中。通常采取的离散化方法有等距离划分法和等频率划分法, 这

两种方法相对简单。对于定性评价指标属性, 则采用若干评价等级来描述。

3.2 算法设计

输入: 伙伴选择决策系统

$$S = (U, C, D), \quad (5)$$

其中: $U = \{x_1, x_2, \dots, x_n\}$ 为合作伙伴候选集, C 为评价指标构成的条件属性集, $D = \{d\}$ 为伙伴选择决策属性集; 规则置信度阈值 μ_0 ;

输出: 伙伴选择决策结果;

Step 1: 检索决策系统中每个样本实例, 将相同的实例合并, 各实例样本个数存入数组 K 中;

Step 2: 计算决策系统可辨识矩阵 M_D ; 由 M_D 求出分辨函数 F 和 C 相对于 $\{d\}$ 的核 C_0 ;

Step 3: 计算 F 的极小析取范式中的所有合取式, 得到条件属性 C 的所有决策属性 D 简化集 $\text{red}(S)$;

Step 4: 以 $\text{red}(S)$ 为初始节点, 建立层状递减节点模型。将具有相同条件属性数 i 的节点放在同一层, 层与层之间条件属性数递减 $i = i - 1$, 直到 $i = 0$ 终止;

Step 5: 对于层状模型的各节点, 分别计算节点各规则的置信度 $\mu(X_i, Y_j)$, 提取 $\mu(X_i, Y_j)$ 大于置信度阈值 μ_0 的规则, 将其写入规则集 DB 中。

Step 6: 在 DB 中校验规则冗余性和多义性。IF 规则的结论相同, 而条件属性多少不同, THEN 选择规则置信度高的规则; IF 规则置信度相同, THEN 优先选用上层次的节点规则; IF 规则考核指标属性相同而结论不同, THEN 在用户窗口显示该条规则, 以便进一步分析评判;

Step 7: 校验过的规则作为伙伴选择知识库按规则置信度高低分层次存储;

Step 8: 对于待决策的伙伴选择数据集中各评价指标相关数据和信息进行离散化处理, 用知识库中的规则对数据集中记录逐条进行评价和伙伴决策选择。

3.3 应用算例

本文以江苏某工程机械制造厂为应用背景, 将粗糙集方法运用于该企业伙伴的选择。根据该企业的产品类型特点及动态联盟的目标, 建立一个用于该企业伙伴选择的评价指标体系, 即生产质量 Q 、产品报价 C 、技术服务 S 、交货准时性 T 、生产管理 M 、信息化能力 I 、财务信誉 B 、地理环境 E 八个伙伴选择评价指标。

现取 60 组通过认定的数据作为学习样本 $U = \{x_1, x_2, \dots, x_{60}\}$, 条件属性 $C = \{Q, C, S, T, M, I, B, E\}$ 。首先对连续属性值进行离散化处理, 本文采

用等频率区间法, 制定划分区间个数为 5, 用对象总数除以区间个数得到每个区间的样本数, 根据每个区间的样本数确定区间边界. 如候选伙伴的产品报价 C 属性直接关系到盟主企业的赢利, 其值的离散结果为 $C = \{0, 1, 2, 3, 4\}$, $C = 0$ 表示候选企业的报价比基本价位低 15% 以上, $C = 1$ 表示报价比基本价位低 5% ~ 15%, $C = 2$ 表示报价在基本价位 $\pm 5\%$ 区间, $C = 3, 4$ 分别表示报价比基本价位高 5% ~ 15% 和 15% 以上. 对于定性属性本文采用 5 个评价等级来描述, 如信息化能力 $I = \{0, 1, 2, 3, 4\} = \{\text{很高, 高, 一般, 低, 很低}\}$. 将合作伙伴决策选择分为重要合作伙伴 0、一般合作伙伴 1 及不作为合作伙伴 2, 所以决策属性 $\{d\} = \{0, 1, 2\}$, 如表 1 所示.

表 1 用决策系统表示的合作伙伴选择事例集

U	Q	C	S	T	M	I	B	E	d
x_1	1	0	1	1	0	2	0	1	0
x_2	2	1	2	2	1	2	1	2	1
x_3	2	2	1	2	0	2	2	1	1
x_4	2	4	4	3	2	3	2	3	2
\vdots									
x_{57}	1	1	2	1	0	1	1	2	0
x_{58}	3	1	4	2	3	4	3	2	2
x_{59}	2	2	2	0	1	3	1	3	1
x_{60}	0	2	2	0	1	3	1	3	0

计算其可辨识矩阵

$$M_D = \begin{bmatrix} 0 & QCSTMBE & QCTB & QCSTMIBE & \dots & 0 & QCSTMIBE & QCSTMIBE & 0 \\ & 0 & 0 & QSTMIBE & \dots & QTM I & QSM IB & 0 & QCTIE \\ & & 0 & CSTMIE & \dots & QCSTIBE & QCSMIBE & 0 & QSTMIBE \\ & & & 0 & \dots & QCSTMIBE & 0 & CSTMBE & QCSTMIBE \\ & & & & \dots & \vdots & \vdots & \vdots & \vdots \\ & & & & & 0 & QSTMIB & QCTMIE & 0 \\ & & & & & & 0 & QCSTMIBE & QCSTMIBE \\ & & & & & & & 0 & Q \\ & & & & & & & & 0 \end{bmatrix}$$

可见, M_D 是一个对角线上元素为零的对称矩阵. 由 M_D 可得属性核 C_0 为 Q, C, T 及其分辨函数

$F =$

$$\begin{matrix} (Q & C & S & T & M & B & E) \\ (Q & C & T & B) & (Q & C & S & T \\ M & I & B & E) & (Q & C & S & T \\ M & I & B & E) & (Q & C & S & T \\ M & I & B & E) & (Q & S & T & M \\ I & B & E) & (Q & T & M & I) \\ (Q & S & M & I & B) & (Q & C & T \\ I & E) & (C & S & T & M & I & E) \\ (Q & C & S & T & I & B & E) \\ (Q & C & S & M & I & B & E) & \dots \\ (Q & S & T & M & I & B & E) \\ (Q & C & S & T & M & I & B & E) \\ (C & S & T & M & B & E) & (Q & C \\ S & T & M & B & E) & (Q & S & T & M \\ I & B) & (Q & C & T & M & I & E) \\ (Q & C & S & T & M & I & B & E) \\ (Q & C & S & T & M & I & B & E) \end{matrix}$$

$$Q = (Q \ C \ T \ S) \ (Q \ C \ T \ M) \ (Q \ C \ T \ I) \ (Q \ C \ T \ B).$$

4 个简式 $\{QCTS\}$, $\{QCTM\}$, $\{QCTI\}$ 和 $\{QCTB\}$ 为条件属性 C 对决策属性 $\{d\}$ 的约简集. 约简集中没有属性 E , 由此可见, 在这里属性 E (企业地理环境) 是决策的冗余属性, 去掉属性 E 后并未改变决策系统的分类能力. 属性的约简集有 4 个, 以 4 个约简集为初始节点, 建立层状节点模型. 模型共分 5 层, 第 1 层节点为 4 个约简集 $\{\{QCTS\}, \{QCTM\}, \{QCTI\}, \{QCTB\}\}$; 第 2 层为 $\{\{QCT\}, \{QCS\}, \{CTS\}, \{QTS\}, \{QCM\}, \{CTM\}, \{QTM\}, \{QCI\}, \{CTI\}, \{QTI\}, \{QCB\}, \{CTB\}, \{QTB\}\}$ 共 13 个节点; 第 3 层为 $\{\{QC\}, \{CT\}, \{QT\}, \{QS\}, \{CS\}, \{TS\}, \{CM\}, \{QM\}, \{TM\}, \{QI\}, \{CI\}, \{TI\}, \{CB\}, \{QB\}, \{TB\}\}$ 共 15 个节点; 第 4 层为 $\{\{Q\}, \{C\}, \{T\}, \{S\}, \{M\}, \{I\}, \{B\}\}$ 共 7 个节点; 第 5 层为 $\{\emptyset\}$ 作为算法循环结束的标志. 对于模型各节点, 提取规则置信度大于置信度阈值 μ_0 的规则, 设 $\mu_0 = 0.85$, 限于篇幅, 本文只给出节点 QTM 上规则置信度大于 0.85 的规则集合, 如表 2 所示.

将由各节点得到的大于置信度阈值 0.85 的决策规则逐个添加到规则集 DB 中, 经检查校验规则

表2 节点 QTM 的规则集合

	规 则				置信度/%	样本数
1	$Q = 1$	$T = 1$	$M = 0$	$d = 0$	98	12
2	$Q = 0$	$T = 0$	$M = 1$	$d = 0$	100	6
3	$Q = 0$	$T = 1$	$M = 2$	$d = 0$	89	8
4	$Q = 2$	$T = 0$	$M = 1$	$d = 1$	87	5
5	$Q = 1$	$T = 2$	$M = 2$	$d = 1$	93	4
6	$Q = 1$	$T = 1$	$M = 1$	$d = 1$	100	10
7	$Q = 3$	$T = 2$	$M = 2$	$d = 2$	100	7
8	$Q = 3$	$T = 3$	$d = 2$		100	8

冗余性和多义性,最后形成该企业合作伙伴选择决策的知识规则库,共有68条规则按规则置信度高低分层次存储在数据库中。运用这些规则对伙伴选择数据库中其他序列样本进行伙伴选择决策,如收集到候选伙伴的一组评价指标数据,采用等频率区间方法得到离散的属性值为: {0, 1, 1, 0, 1, 2, 0, 3} 其属性 $Q = 0, T = 0, M = 1$,与第2条规则匹配,得出 $d = 0$,由此做出可作为重要合作伙伴关系的判断,这与实际应用结果相一致。本文对50组样本进行选择决策测试,其正确率为93.2%。

4 结 论

虚拟企业的伙伴选择本身是个相对复杂的决策过程,由于其影响因素的多种多样,并且在这些因素间还存在着彼此制约或相互依赖的关系,需要采用科学的决策方法,做出最优的决策选择。本文尝试性地将粗糙集理论运用到虚拟企业伙伴选择问题中,提出了伙伴选择决策模型和方法,方法中不需要任何先验知识和加权,实际应用研究表明,基于粗糙集理论的伙伴选择方法更加客观、合理。需要注意的是,由于评价指标中存在许多难以量化的指标,给伙伴选择最优化决策带来困难。评价指标的离散化问题还需要进一步的研究。

参考文献(References)

- [1] Olivier P, Claude G. A Model to Support Collaborative Work in Virtual Enterprises[J]. *Data and Knowledge Engineering*, 2004, 50(1): 63-86
- [2] Martinez M T. Virtual Enterprise — Organisation, Evolution and Control [J]. *Int J of Production Economics*, 2001, 74(1/3): 225-238
- [3] 廖貅武,唐焕文. 动态联盟中伙伴选择的证据推理方法[J]. *计算机集成制造系统-CMS*, 2003, 9(1): 57-62 (Liao X W, Tang H W. An Evidential Reasoning Approach for Partner Selection in Dynamic Alliance[J]. *Computer Integrated Manufacturing Systems*, 2003, 9(1): 57-62)
- [4] 曹洪医,汪定伟. 用GA求解动态联盟中伙伴选择的多目标优化模型[J]. *控制与决策*, 2002, 17(3): 274-277. (Cao H Y, Wang D W. Genetic Algorithm for a Multiobjective Optimization Model of Partner Selection in Virtual Enterprise [J]. *Control and Decision*, 2002, 17(3): 274-277.)
- [5] Pawlak Z. Rough Sets Theory and Its Applications to Data Analysis[J]. *Cybernetics and Systems*, 1998, 29: 661-668
- [6] Ziarko W. The Discovery, Analysis and Representation of Data Dependencies in Databases [A]. *Knowledge Discovery in Databases [C]*. Cambridge: AAAI/MIT Press, 1990: 213-228
- [7] 张文修,吴志伟,梁吉业,等. *粗糙集理论与方法*[M]. 北京: 科学出版社, 2001. 75-130 (Zhang W X, Wu Z W, Liang J Y, et al. *Rough Set Theory and Its Method* [M]. Beijing: Science Press, 2001: 75-130)
- [8] 李永敏,朱善君,陈湘晖,等. 基于粗糙集理论的数据挖掘模型[J]. *清华大学学报*, 1999, 39(1): 110-113 (Li Y M, Zhu S J, Chen X H, et al. Data Mining Model Based on Rough Set Theory [J]. *J of Tsinghua University*, 1999, 39(1): 110-113)