

# 基于模糊粗糙集理论的建筑业综合评价

姜连馥<sup>\*</sup>1, 石永威<sup>1</sup>, 满杰<sup>2</sup>, 杨尚群<sup>3</sup>

(1. 深圳大学 土木工程学院, 广东 深圳 518060;  
2. 深圳大学 光电工程学院, 广东 深圳 518060;  
3. 对外经济贸易大学 信息学院, 北京 100029)

**摘要:** 根据模糊聚类 and 粗糙集理论的基本原理, 利用模糊聚类的相似关系与粗糙集可辨识矩阵的特性, 结合这两种方法的互补优势, 研究了将之应用于建筑业综合评价的方法和步骤. 通过模糊聚类法形成粗糙集下的决策信息系统, 采用基于可辨识矩阵的启发式约简法对属性约简, 计算出约简后各属性的权重. 利用各属性的权重分别对各省(或地区)的建筑业指标评价价值加权求和从而得出各自的建筑业综合评价价值, 对其结果的客观性和实用性所进行的分析评价证明该方法有效.

**关键词:** 模糊聚类 / 模糊相似关系; 粗糙集; 可辨识矩阵; 建筑业评价  
**中图分类号:** F407.9 **文献标识码:** A

## 0 引言

对各地区的建筑业状况进行综合实力评价, 其评价的结果对于相关产业政策的制定和战略方案的优选具有重要的理论与实践价值. 综合实力评价指标的设置反映建筑业的综合性与专业特点的基础上, 既要有前瞻性的指标, 也要有考虑与历史可比的指标; 同时, 还要考虑到指标的基础数据便于搜集, 计算简便、科学, 可操作性强. 由于评价指标体系一般具有多目标的特点, 在一定程度上存在冗余, 如何筛选出有效的评价指标, 建立合理的评价体系成为对各个地方建筑业进行综合实力评价的关键所在.

目前常用于建筑业评价的方法有专家咨询法、网络层次分析法、模糊综合评判法、主成分分析法、因子分析法、数据包络法、空间几何距离法. 这些方法的不足主要表现在: 主观性太强; 对数据的精确性要求太高且在处理过程中会损失部分既有的评价信息. 而粗糙集方法作为一种新兴的处理不确定性、不精确性的数学理论, 可以仅利用数据本身的信息, 无需任何附加信息和先验知识, 在信息不丢失以及知识库分类能力不变的前提下,

对评价指标体系进行有效约简, 提取有用的特征信息, 在处理决策信息系统的离散属性上占有绝对的优势, 可以弥补上述方法的不足. 本文依据模糊聚类和粗糙集这两者的互补性<sup>[1, 2]</sup>, 首先利用模糊聚类法的特性来实现指标的离散化, 形成粗糙集下的决策信息系统, 并利用其对各省(或地区)的建筑业综合实力进行分析评价.

## 1 基于粗糙集和模糊聚类的决策算法

### 1.1 粗糙集的基本理论

粗糙集理论是波兰数学家 Pawlak 于 1982 年提出来的, 在粗糙集理论中, 知识的表达通过信息系统(决策表)来实现. 决策表以四元组  $S = \{U, A, V, f\}$  的形式存在, 其中非空有限论域  $U = \{x_1, x_2, \dots, x_m\}$  为有限数量的对象集;  $A$  为有限数量的评价指标属性集, 由条件属性集合  $C$  和决策属性集合  $D$  组成;  $V = \bigcup_{a \in A} V_a$ ,  $V_a$  是属性  $a \in A$  的值域,  $a(x)$  是对象  $x$  在属性  $a$  上的值;  $f: U \rightarrow V_a$  为单一映射, 使论域  $U$  中任一对象的属性  $a$  有唯一的一个信息值.

收稿日期: 2006-01-19; 修回日期: 2007-07-03.

基金项目: 国家自然科学基金资助项目(70671072).

作者简介: 姜连馥<sup>\*</sup>(1955-), 男, 博士, 副教授, 1987~1993年在大连理工大学工作.

1.1.1 可辨识矩阵<sup>[3]</sup> 辨识矩阵可定义为信息系统中当决策属性不同且条件属性也不完全相同时,元素值为互不相同的属性组合;当决策属性相同时,元素值为 0,当决策属性不同,而条件属性完全相同时,元素值为 - 1,则表明系统中数据有误或提供条件属性不足.则可辨识矩阵记为

$$M = (m_{ij})_{N \times N} = \begin{cases} \{a \mid a \in f(x_i, a) \neq f(x_j, a)\}; \\ \quad \text{当 } f(x_i, D) \neq f(x_j, D), x_i, x_j \in U \\ 0; \text{ 当 } f(x_i, D) = f(x_j, D) \\ - 1; \text{ 当 } \forall a, \exists a(x_i) = a(x_j), \\ \quad f(x_i, D) \neq f(x_j, D) \end{cases} \quad (1)$$

其中  $N = |U|$ .

1.1.2 属性重要度的计算<sup>[4]</sup> 对于一个可辨识矩阵  $M = (m_{ij})_{N \times N}$ ,相应的属性  $a$  的重要度的计算公式为

$$f(a) = \sum_{i=1}^N \sum_{j=1}^N \frac{\lambda_{ij}}{|m_{ij}|} \quad (2)$$

其中  $\lambda_{ij} = \begin{cases} 0; a \in I_{m_{ij}} \\ 1; a \in m_{ij} \end{cases}$ ,  $|m_{ij}|$  表示  $m_{ij}$  包含属性的个数.

从可辨识矩阵的定义可以看出,可辨识矩阵中某项的长度越短,该项对分类所起的作用就越大,且该项出现得越频繁,该项越重要.因此,属性在可辨识矩阵中出现的次数越多,出现的项越短,则属性的重要度越大.

1.1.3 属性约简<sup>[4]</sup> 所谓属性约简就是在保持知识库分类能力不变的条件下,删除其中不相关的知识或者不重要的知识.如果约简属性  $a$  后满足式(3),则认为属性  $a$  是决策信息系统中不必要的.

$$U_1 \cup U_2 < X \quad (3)$$

其中  $U_1$  为删除属性  $a$  后引入的不相容样本数,  $U_2$  为删除属性  $a$  之前知识表中样本的数量,  $X$  为阈值,可根据实际需要确定.

1.2 模糊聚类概述

模糊聚类所处理的问题是已知论域  $U = \{x_1, x_2, \dots, x_n\}$  为待分类的对象,每个对象又由  $m$  个指标表示其性状特征,按照问题的目的要求将对象划分若干类.模糊等价关系法是其中比较直观的一种聚类方法,它主要利用了模糊理论中的相似关系.所谓相似关系是指满足自反性和对称性的二元模糊关系,采用编网法<sup>[5]</sup>可以在模糊相

似矩阵  $R$  的截矩阵  $R_\lambda$  上直接进行聚类,从而给出论域  $U$  的一个分类.其中

$$R = (r_{ij})_{n \times n} \in U^{n \times n}, r_{ij} = r_{ji}, r_{ii} = 1 (i, j = 1, \dots, n) \text{ 且 } r_{ij} \in [0, 1]$$

确定  $r_{ij}$  的值主要有以下三类方法. (1) 相似系数法: 数量积法, 夹角余弦法, 相关系数法, 指数相似系数法, 最大最小法, 算术平均最小法, 几何平均最小法; (2) 距离法: 绝对值倒数法, 绝对值指数法, 绝对值减数法; (3) 主观评分法. 本文采用绝对值减数法来确定  $r_{ij}$  的值:

$$r_{ij} = \begin{cases} 1; i = j \\ 1 - c \cdot \sum_{k=1}^m |x_{ik} - x_{jk}|; i \neq j \end{cases} \quad (4)$$

$\lambda \in [0, 1], R_\lambda = \{A_i \mid A_i = \{x_j \mid r_{ij} \geq \lambda, j = 1, \dots, n\}, i = 1, \dots, n\}$ . 显然,  $\lambda$  的取值直接影响到聚类的类别,其值越大,分类越细,反之亦然.

2 基于模糊粗糙算法的建筑业综合评价指标和评价步骤

2.1 建筑业综合实力评价指标的建立

影响地区建筑业综合实力的因素多且复杂,遵循指标选取的专业性、可操作性、综合性和完备性原则,本文在大量相关文献<sup>[6-8]</sup>的基础上选取其中 12 个指标作为反映各省市建筑业综合实力的指标体系.具体所取指标及其说明如表 1 所示.

2.2 基于模糊粗糙集的建筑业综合评价步骤

选取建筑业综合评价指标之后,结合上述粗糙集和模糊聚类的基本理论,本文归纳出基于模糊粗糙集的建筑业综合实力评价步骤如下:

- (1) 根据建立的评价指标,对样本(对象)的指标赋值,并将样本及样本的指标值(属性值)组成一个信息矩阵;
- (2) 对指标值(属性值)做量纲一化及同向化的规格化处理;
- (3) 对规格化后的信息矩阵建立模糊相似关系  $R$ ,用编网法在  $R$  的  $\lambda$  截矩阵  $R_\lambda$  上直接进行聚类;

(4) 通过模糊等价聚类的编网法对原信息系统进行离散化处理: 信息表中条件属性值离散化是依属性逐个处理的,进行样本聚类的样本都是一维变量,然后选取适合聚类目标的分类作为决策属性值与离散化后的信息矩阵组合成决策信息系统;

表 1 建筑业综合实力评价指标

Tab. 1 Indicators of comprehensive evaluation of construction industry

指标名称	编号	指标说明	单位
固定资产投资	$a_1$	由于固定资产投资的增长和建筑业总产值的增长总体上是正相关的,这一指标可反映建筑业总产值的增长情况	亿元
建筑业从业人员	$a_2$	建筑业从业人员数量在一定程度上反映了建筑业产业规模	万人
亿元产值能耗	$a_3$	反映建筑产品在形成过程中对能源消耗和对生态环境影响的情况,是反映建筑业可持续发展能力的一个重要指标	万吨标准煤
建筑垃圾量	$a_{11}$		万吨
人均利润	$a_4$		亿元/万人
人均产值	$a_5$	虽然总量指标(如固定资产投资)可较好地反映建筑业产业规模,但不能全面反映建筑业的真实状况.有些地区建筑业规模虽较大,但人均值却偏小,说明这些地区建筑业仍然停留在以粗放型增长方式为主的发展阶段.故本文中引入这些人均指标反映建筑业产业水平	亿元/万人
人均房屋竣工面积	$a_6$		$m^2$
从业人员人均报酬	$a_{12}$		元
动力装备率	$a_7$	对发达国家的建筑业而言,促进其增长的各要素中,技术进步的贡献居首,而技术装备水平则是技术进步的一个重要标志,主要体现为技术装备率和动力装备率两个方面	kW/人
技术装备率	$a_8$		元/人
产值利润率	$a_9$	体现建筑业产值的盈利水平,是反映建筑业发展效率的重要指标	%
全员劳动生产率	$a_{10}$	劳动生产率是反映一个产业生命力及现代化程度最重要的综合指标,该指标的变化可有效反映这些企业的技术进步情况	元/人

(5) 计算决策信息系统的可辨识矩阵,然后得出各属性的重要度并排序;

(6) 删除重要性最小的属性  $a$  并判断约简操作是否满足式 (3),若满足则继续选择约简后知识表中重要性最小的属性进行约简,否则停止操作;

(7) 将约简后属性的重要性进行归一化处理,得到各属性的权重,然后采用约简后的各属性值的规格化数据作为指标评价矩阵;

(8) 依次对各省市计算指标评价价值与权重的加权,确定各评价对象的综合评价价值.

### 3 应用实例——我国八省市建筑业综合实力评价

本文选取建筑业总产值之和约占全国 57%

的八省市:北京( $x_1$ )、上海( $x_2$ )、天津( $x_3$ )、重庆( $x_4$ )、江苏( $x_5$ )、浙江( $x_6$ )、广东( $x_7$ )、山东( $x_8$ )为评价对象,评价对象的指标属性值数据来源于 2006 年各省市的统计年鉴(限于篇幅,原始数据表没有列出).

#### 3.1 评价指标值的规格化

极差规格化公式:  $u'_{ij} = \frac{u_{ij} - m_j}{M_j - m_j}$ ,  $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, m$ . 其中  $m_j, M_j$  分别为对信息矩阵的第  $j$  列计算  $m_j = \min(u_{1j}, u_{2j}, \dots, u_{nj})$ ,  $M_j = \max(u_{1j}, u_{2j}, \dots, u_{nj})$  消除了原指标数据量纲的影响,注意到指标  $a_3$  和  $a_{11}$  均为逆向指标,需要采用公式  $u''_{ij} = 1 - u'_{ij}$  将之与其他指标同向化,所得规格化后的信息矩阵如表 2 所示.

表 2 规格化后的数据

Tab. 2 Normalized data

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$
$x_1$	0.20	0.06	0.83	0.30	1	0.59	0.93	0.33	0.28	0.79	0.91	0.62
$x_2$	0.29	0.13	0	1	0.77	0.47	0.31	0.24	1	1	0.81	1
$x_3$	0	0	1	0.18	0.47	0.13	1	1	0.06	0.91	1	0.07
$x_4$	0.06	0.18	0.82	0	0.03	0.41	0	0.05	0.15	0.16	0.86	0
$x_5$	0.87	1	0.70	0.09	0.23	0.56	0.15	0.06	0	0.49	0.41	0.25
$x_6$	0.85	0.89	0.18	0.22	0.33	1	0.02	0	0.41	0.75	0.74	0.36
$x_7$	0.75	0.43	0.13	0.25	0.20	0.25	0.43	0.24	0.71	0.46	0.70	0.29
$x_8$	1	0.84	0.86	0.15	0	0	0.09	0.02	0.49	0	0	0.06

### 3.2 模糊聚类

取论域  $U = \{x_1, x_2, \dots, x_8\}$ , 按照式 (4) 求相似系数 (取  $c = 0.1$ ), 得模糊相似矩阵

$$R = (r_{ij})_{8 \times 8} = \begin{pmatrix} 1 & 0.58 & 0.67 & 0.56 & 0.47 & 0.49 & 0.53 & 0.30 \\ 0.58 & 1 & 0.35 & 0.39 & 0.33 & 0.47 & 0.58 & 0.19 \\ 0.67 & 0.35 & 1 & 0.57 & 0.40 & 0.34 & 0.45 & 0.31 \\ 0.56 & 0.39 & 0.57 & 1 & 0.65 & 0.54 & 0.59 & 0.62 \\ 0.47 & 0.33 & 0.40 & 0.65 & 1 & 0.74 & 0.67 & 0.70 \\ 0.49 & 0.47 & 0.34 & 0.54 & 0.74 & 1 & 0.71 & 0.57 \\ 0.53 & 0.58 & 0.45 & 0.59 & 0.67 & 0.71 & 1 & 0.59 \\ 0.30 & 0.19 & 0.31 & 0.62 & 0.70 & 0.57 & 0.59 & 1 \end{pmatrix}$$

将  $R$  中的元素进行排序为

$1 > 0.74 > 0.71 > 0.70 > 0.67 > 0.65 > 0.62 > 0.59 > 0.58 > 0.57 > 0.56 > 0.54 > 0.53 > 0.49 > 0.47 > 0.45 > 0.40 > 0.39 > 0.35 > 0.34 > 0.33 > 0.31 > 0.30 > 0.19$

根据这 8 个省市建筑业综合实力聚类分析的结果及总分情况, 并做出适当的调整, 发现当  $\lambda = 0.65$  时, 将  $U$  分为三类较为合理, 此时  $x_1, x_3$  相连接,  $x_4, x_5, x_6, x_7, x_8$  相连接, 因此  $U$  被分为  $\{x_2\}, \{x_1, x_3\}, \{x_4, x_5, x_6, x_7, x_8\}$ . 同样采用模糊等价聚类的编网法对原信息系统中的条件属性值依属性逐个进行样本聚类得到离散化的信息矩阵. 最后得到模糊离散化的决策信息系统如表 3.

表 3 模糊离散化后的决策信息系统

Tab. 3 The decision-making information system by the result from fuzzy aggregation analysis

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$	$a_{11}$	$a_{12}$	$d$
$x_1$	1	1	2	2	3	2	3	2	1	3	3	2	2
$x_2$	2	1	1	3	2	2	2	2	3	3	3	3	3
$x_3$	1	1	3	2	1	1	3	3	1	3	3	1	2
$x_4$	1	1	2	1	1	2	1	1	1	1	3	1	1
$x_5$	3	1	2	2	1	2	1	1	1	2	2	1	1
$x_6$	3	3	1	2	1	3	1	1	1	3	3	1	1
$x_7$	3	2	1	2	1	1	2	2	2	2	3	1	1
$x_8$	3	3	2	2	1	1	1	1	1	1	1	1	1

### 3.3 决策信息系统的属性重要度计算及约简

由式 (1) 可以计算决策信息系统表 3 的可辨

识矩阵 (见表 4).

表 4 可辨识矩阵

Tab. 4 Discernibility matrix

$U$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$
$x_1$	0							
$x_2$	$a_1 a_3 a_4 a_5$	0						
$x_3$	0	$a_1 a_3 a_4 a_6$ $a_7 a_8 a_9 a_{12}$	0					
$x_4$	$a_4 a_5 a_7 a_8$ $a_{10} a_{12}$	$a_1 a_3 a_4 a_5 a_7$ $a_8 a_9 a_{10} a_{12}$	$a_3 a_4 a_6$	0				
$x_5$	$a_1 a_5 a_7 a_8$ $a_{10} a_{11} a_{12}$	$a_1 a_3 a_4 a_5 a_7 a_8$	$a_1 a_3 a_6 a_7$	0	0			
$x_6$	$a_1 a_2 a_3 a_5$ $a_6 a_7 a_8 a_{12}$	$a_1 a_3 a_4 a_5 a_6$ $a_7 a_8 a_9 a_{12}$	$a_1 a_2 a_3$	0	0	0		
$x_7$	$a_1 a_2 a_3 a_5 a_6$ $a_7 a_9 a_{10} a_{12}$	$a_1 a_3 a_4 a_5 a_6$ $a_9 a_{10} a_{12}$	$a_1 a_2 a_3 a_7$ $a_8 a_9 a_{10}$	0	0	0	0	
$x_8$	$a_1 a_2 a_3 a_6 a_7$ $a_{10} a_{11} a_{12}$	$a_1 a_2 a_3 a_4 a_5 a_6 a_7$	$a_1 a_2 a_3 a_7$	0	0	0	0	0

根据基于可辨识矩阵的启发式算法,按照式

(2) 计算 12 个属性的重要性:

$$f(a_1) = \sum_{i=1}^8 \sum_{j=1}^8 \frac{1}{m_{ij}} = \frac{1}{7} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{8} + \frac{1}{9} + \frac{1}{9} + \frac{1}{10} + \frac{1}{9} + \frac{1}{8} + \frac{1}{12} + \frac{1}{7} + \frac{1}{6} + \frac{1}{7} + \frac{1}{7} = 1.8837$$

$$f(a_2) = 1.1329, f(a_3) = 1.5464, f(a_4) = 0.9750, f(a_5) = 1.4552, f(a_6) = 1.2679, f(a_7) = 2.0921, f(a_8) = 1.7131, f(a_9) = 1.0385, f(a_{10}) = 1.5603, f(a_{11}) = 0.7369, f(a_{12}) = 1.4552$$

由此可得出属性重要性的降序排列为  $a_7-a_1-a_8-a_{10}-a_3-a_{12}-a_5-a_6-a_2-a_9-a_4-a_{11}$ . 然后按照 2.2 节中第 (6) 步对指标进行约简,根据本文对建筑业综合实力评价的精度要求,取  $X=5%$ . 先后删去  $a_{11}$  及  $a_4$ ,当删除  $a_9$  时由于  $U_i \cap U_j = 5.78% > 5%$ ,因此结束约简,最后可约简  $a_{11}$  和  $a_4$ .

### 3.4 综合评价

采用公式  $w_{a_i} = \frac{f(a_i)}{\sum f(a_i)}$  先对约简后的指标

的重要性进行归一化,得到各个指标权重:

$$w_{a_1} = \frac{1.8837}{15.1452} = 0.1244, w_{a_2} = 0.0748, w_{a_3} = 0.1021, w_{a_5} = 0.0961, w_{a_6} = 0.0837, w_{a_7} = 0.1381, w_{a_8} = 0.1131, w_{a_9} = 0.0686, w_{a_{10}} = 0.1030, w_{a_{12}} = 0.0961.$$

然后采用约简后各属性值的规格化数据作为指标评价矩阵如  $X$  所示.

$$X = \begin{pmatrix} 0.20 & 0.06 & 0.83 & 1 & 0.59 & 0.93 & 0.33 & 0.28 & 0.79 & 0.62 \\ 0.29 & 0.13 & 0 & 0.77 & 0.47 & 0.31 & 0.24 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0.47 & 0.13 & 1 & 1 & 0.06 & 0.91 & 0.07 \\ 0.06 & 0.18 & 0.82 & 0.03 & 0.41 & 0 & 0.05 & 0.15 & 0.16 & 0 \\ 0.87 & 1 & 0.70 & 0.23 & 0.56 & 0.15 & 0.06 & 0 & 0.49 & 0.25 \\ 0.85 & 0.89 & 0.18 & 0.33 & 1 & 0.02 & 0 & 0.41 & 0.75 & 0.36 \\ 0.75 & 0.43 & 0.13 & 0.20 & 0.25 & 0.43 & 0.24 & 0.71 & 0.46 & 0.29 \\ 1 & 0.84 & 0.86 & 0 & 0 & 0.09 & 0.02 & 0.49 & 0 & 0.09 \end{pmatrix}$$

然后对指标评价矩阵  $X$  逐行与指标权重  $w_{a_i}$  加权求和,得到我国八省市的建筑业综合评价值分别为  $V_{x_1} = 0.5855, V_{x_2} = 0.4968, V_{x_3} = 0.5139, V_{x_4} = 0.1743, V_{x_5} = 0.3172, V_{x_6} = 0.4488, V_{x_7} = 0.3894, V_{x_8} = 0.3291$ .

### 3.5 结果分析

从最后的综合评价结果可以看出,北京、天津

和上海这三市的建筑业处于领先水平,紧随其后的浙江及广东两省在建筑业也显示出了很强的实力,山东和江苏二省的建筑业水平处于再次的位置,而重庆市的建筑业水平则处于落后的地位. 然后通过各评价指标的权重以及各省市的建筑业综合评价值再结合各省市的建筑业评价指标值可以得出以下结论:

(1) 衡量建筑业技术化水平的指标动力装备率和技术装备率对各省市建筑业的综合实力得分影响最大,北京和天津这两市能在建筑业综合实力排名位居前列是与它们的高技术化水平分不开的.

(2) 固定资产投资对建筑业的影响十分突出,这说明建筑业的发展和整体经济状况密切相关.

(3) 全员劳动生产率是衡量建筑业的生命力及现代化程度最重要、最综合的指标,其中上海、天津的劳动生产率最高,重庆和山东最低,因此提高劳动生产率,是山东和重庆提高建筑业综合实力的关键.

(4) 从亿元产值能耗来看,上海的亿元产值能耗最高,这对于上海市在建筑业综合实力排名上落后于北京和天津有着至关重要的影响,同时对于上海市实现其建筑业的可持续发展以及促进上海市建筑业由传统的“粗放型”向“集约型”转化都是极为不利的. 另外广东和浙江也需要对此引起特别重视,以提升自身的建筑业综合实力.

(5) 地处西部的重庆市无论是在技术水平还是在全员劳动生产率上都明显落后于其他省市,技术落后不仅会导致劳动生产率的降低,而且随着我国的建筑工程向大型化、复杂化、高技术化方向发展,重庆市很可能会因为技术竞争力不足失去许多机会,最后必然会削弱其整体实力.

## 4 结 语

建筑业综合实力值受多种不确定因素的影响,合理确定各因素的权重是科学评价的关键所在. 本文充分利用模糊聚类与粗糙集理论在处理不确定和不精确问题方面的互补性,根据建筑业综合实力评价指标的统计数据所提供的信息确定出相应的权重,实例表明该评价模型具有较好的客观性和实用性,从而为建筑业综合评价提供了一种新的方法. 今后如能进一步与其他理论如灰色系统理论相结合<sup>[9]</sup>,则能使该评价模型涵盖面

更宽,应用范围更广.

## 参考文献:

- [1] JENSEN R, SHEN Qiang. Fuzzy-rough attribute reduction with application to web categorization [J]. *Fuzzy Sets and Syst*, 2004, **141**: 469-485
- [2] SWINIARSKI R W, SKOWRON A. Rough set methods in feature selection and recognition [J]. *Pattern Recog Lett*, 2003, **24**: 833-849
- [3] LASHIN E F, MEDHAT T. Topological reduction of information systems [J]. *Chaos, Solitons and Fractals*, 2005, **25**: 277-286.
- [4] 芦晓红,陈世权,吴今培. 基于可辨识矩阵的启发式属性约简方法及其应用 [J]. *计算机工程*, 2003, **29**(1): 56-59
- [5] 陈水利,李敬功,王向公. 模糊集理论及其应用 [M]. 北京: 科学出版社, 2005
- [6] 管昌生,王蕾蕾. 主成分分析法在建筑企业评价中的应用 [J]. *计算机与数字工程*, 2006, **34**(1): 61-64
- [7] 张长根,李涛. 因子分析法对建筑业综合实力的评价 [J]. *建筑经济*, 2003, **12**: 47-49
- [8] 陆宁,蔡爱云,黄永安,等. 我国建筑业可持续发展综合评价 [J]. *重庆建筑大学学报*, 2006, **28**(4): 94-97
- [9] WANG Xiao-sheng. Some relationships of rough sets and grey systems [J]. *Fuzzy Syst and Math*, 2006, **20**(6): 129-135

## Comprehensive evaluation of building industry based on fuzzy aggregation and rough set theory

JIANG Lian fu<sup>\* 1</sup>, SHI Yong wei<sup>1</sup>, MAN Jie<sup>2</sup>, YANG Shang-qun<sup>3</sup>

( 1.College of Civil Eng., Shenzhen Univ., Shenzhen 518060, China;

2.College of Optoelectr. Eng., Shenzhen Univ., Shenzhen 518060, China;

3.School of Inf. Technol. & Manage. Eng., Univ. of Int. Bus. and Econ., Beijing 100029, China )

**Abstract** According to the basic principles of fuzzy aggregation and rough set theory, the similar relationship of fuzzy aggregation was combined with the characteristics of discernibility matrix of rough set and the application of fuzzy aggregation and rough set theory was studied in comprehensive evaluation of building industry based on their complementary advantages. The decision information system was formed through fuzzy aggregation, then the heuristic algorithm of attribute reduction based on discernibility matrix was used to decide the sequence of attribute reduction and ascertain attribute weight. According to the weight of each attribute, the attribute values of building industry in each province or region were summed to get an integrated value. Through analyzing and evaluating the objectivity and practicality of the results, the validation of the evaluating method used is demonstrated.

**Key words** fuzzy aggregation/ fuzzy similar relationship; rough set; discernibility matrix; evaluation of building industry