

# 基于粗集的土地定级因素权重确定方法研究

孔维华 (山东理工大学, 山东淄博 255049)

**摘要** 定级因素权重的确定是土地定级中一项重要的内容, 传统方法通常都是根据决策者的先验知识主观确定, 基于粗集理论的单个属性权重确定方法无法解决简化因素的权重问题。结合土地定级实际, 提出了基于粗集理论的综合权重确定单个因素权重的方法, 并利用主观和客观结果综合确定权重。

**关键词** 粗集; 组合权重; 定级因素

中图分类号 F301.24 文献标识码 A 文章编号 0517-6611(2009)15-07101-03

## Study on the Weight Determination Method of Land Grading Element Based on Rough Sets

KONG Wei-hua (Shandong University of Technology, Zibo, Shandong 255049)

**Abstract** The determination of grading element weight is an important content of land grading. In the traditional method, it was determined according to the experience knowledge of decision makers. The determination method of single attribute weight based on rough set theory could not resolve the weight of simplified element. Combining with the practice of land grading, a method of confirming the single element weight based on rough sets' integrated weight was put forward. And the weight was synthetically confirmed by using subjective and objective results.

**Key words** Rough set; Integrated weight; Grading element

随着市场经济和土地有偿使用制度的实施, 综合评定土地质量等级已经成为当前和未来土地管理中一项十分重要的任务, 定级的技术途径采用多因素综合评价法, 因此定级因素的权重对评定结果起着至关重要的作用。目前, 在土地分等定级中常用的确定权重方法主要有3种, 即特尔菲测定法、因素成对比较法、层次分析法<sup>[1]</sup>。这3种方法相对比较成熟但也有不足之处, 受专家和决策者主观因素的影响都较大。文献[2-3]对利用粗集理论确定单个属性权重进行了分析, 提出了主观和客观相结合的方法确定权重, 但只考虑了单个属性对决策结果的影响, 而忽视了因素之间的相互作用对决策结果的影响。文献[4]提出了组合权重的概念, 并给出了基于粗集理论的综合权重的算法模型, 该方法避免了只考虑单一属性的局限性, 但对简化单个属性的权重没有考虑。笔者结合土地定级实际, 提出了基于粗集理论的综合权重确定单个因素权重的方法, 并利用主观和客观结果来综合确定权重, 获得了比较理想的结果。

## 1 粗集理论

粗集理论是波兰华沙理工大学Z.Pawlak教授提出的用于处理不完整和不精确知识的数学工具<sup>[5]</sup>。在应用于属性权重分析时不需要先验知识, 而只利用已有信息判定属性的重要性并分别赋予不同的权重。

**1.1 知识表达系统** 一个知识表达系统可以表达为  $S = \langle U, P, V, f \rangle$ ,  $U$  是对象的集合,  $P$  是属性集合,  $V$  是属性值集合,  $f: U \times P \rightarrow V$  是一个信息函数, 指定  $U$  中每一个对象的属性值。

**1.2 知识的简化和核** 对于属性子集  $P$  属于  $R$ , 若存在  $Q$  独立,  $Q$  属于  $P$ , 且  $\text{ind}(Q) = \text{ind}(P)$ , 则称  $Q$  为  $P$  的简化, 用  $\text{red}(P)$  表示; 显然  $P$  可以有多种简化。  $P$  中所有不可省略关系的集合称为  $P$  的核, 记作  $\text{core}(P)$ 。核和简化的关系如下:

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

**1.3 属性的依赖性和权重** 对于知识库  $K = \langle U, P \rangle$ , 且

有  $P, Q$  属于  $R$ , 当  $\text{ind}(P)$  属于  $\text{ind}(Q)$ , 则称知识  $P$  依赖知识  $Q$ , 知识  $Q$  对  $P$  的依赖定义为:

$$r_P(Q) = \text{card}(\text{POS}_P(Q)) / \text{card}(U) \quad (2)$$

其中,  $\text{card}$  表示集合的基数,  $\text{POS}_P(Q)$  表示集合  $P$  在  $U/\text{ind}(Q)$  中的正区域。假设属性集  $C$  导出的分类的属性子集  $b$  属于  $B$ , 则  $b$  的重要性用二者依赖程度的差来衡量, 即:

$$\text{Imp}_b = r_B(Q) - r_{Bb}(Q) \quad (3)$$

上式表示当从集合  $B$  中去掉属性子集  $b$  后, 分类  $U/C$  的正区域受到怎样的影响。

对每个属性的重要性进行归一化处理, 就得到属性的权重。

$$W_i = \text{Imp}_{b_i} / \sum \text{Imp}_{b_i} \quad (4)$$

## 2 基于粗集的土地定级因素权重确定方法

**2.1 定级因素** 在土地定级中, 不同的定级方法对应着不同的定级因素因子, 以下以分类定级中工业用地定级为例来进行说明。根据《城镇土地分等定级规程》的要求, 通过20位专家的打分, 确定了定级因素及权重值(见表1)。

**2.2 建立决策表** 对表1进行离散化处理, 即将表1中的数据用 {0, 1, 2} 表示。对道路通达度以 0.25, 0.22 为分界点, 对外交通便利度以 0.25, 0.23 为分界点, 基础设施完善度以 0.30, 0.25 为分界点, 自然条件优劣度以 0.17, 0.15 为分界点, 产业集聚影响度以 0.15, 0.12 为分界点。离散化后对相同的行进行合并, 得到如表2所示的决策表  $T = \langle U, P \rangle$ 。其中,  $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ ,  $P = \{a, b, c, d, e\}$ 。

**2.3 因素权重的计算** 根据不可分辨关系有:

$$U/\text{ind}(p) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}\};$$

$$U/\text{ind}(p-a) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 8\}, \{6\}, \{7\}, \{9\}, \{10\}, \{11\}, \{12\}\};$$

$$U/\text{ind}(p-b) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5, 7\}, \{6\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}\};$$

$$U/\text{ind}(p-c) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\}, \{10\}, \{11\}, \{12\}\};$$

$$U/\text{ind}(p-d) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{7\}, \{8\}, \{9\},$$

基金项目 山东省自然科学基金项目(2004ZX31)。

作者简介 孔维华(1968-), 男, 河南汝南人, 副教授, 从事GIS方面的研究。

收稿日期 2009-02-27

{10} ,{11} ,{12} } ;

$U \text{ind}(p-e) = \{ \{1,6\} , \{2\} , \{3\} , \{4\} , \{5\} , \{7\} , \{8\} , \{9\} ,$

{10} ,{11} ,{12} } ;

$\text{POS}_{p-a}(P) = \{ \{1\} , \{2\} , \{3\} , \{4\} , \{6\} , \{7\} , \{9\} , \{10\} ,$

{11} ,{12} } 。

则属性 a 的重要度为:

$$\text{Imp}_a = 1 - (\text{Card}[\text{POS}_{p-a}(p)] / \text{Card}[\text{POS}_p(p)]) = 1/6$$

同理, 求出属性 b、c、d、e 的重要度为:

$$\text{Imp}_b = 1 - (\text{Card}[\text{POS}_{p-b}(p)] / \text{Card}[\text{POS}_p(p)]) = 1/3$$

$$\text{Imp}_c = 1 - (\text{Card}[\text{POS}_{p-c}(p)] / \text{Card}[\text{POS}_p(p)]) = 0$$

$$\text{Imp}_d = 1 - (\text{Card}[\text{POS}_{p-d}(p)] / \text{Card}[\text{POS}_p(p)]) = 0$$

$$\text{Imp}_e = 1 - (\text{Card}[\text{POS}_{p-e}(p)] / \text{Card}[\text{POS}_p(p)]) = 1/6$$

根据式(4) 对属性的重要度进行归一化处理, 得出相应的权重。

$$W_a = (1/6) / [(1/6) + (1/3) + (1/6)] = 1/4$$

$$W_b = (1/3) / [(1/6) + (1/3) + (1/3)] = 1/2$$

$$W_c = (1/3) / [(1/6) + (1/3) + (1/3)] = 1/4$$

$$W_d = 0$$

$$W_e = 0$$

表1 工业用地定级因素及权重

Table 1 The grading denent and weight of industrial land

定级因子 Grading	道路通达度(a) Road accessibility	对外交通便利度(b) Convenience degree of external traffic	基础设施完善度(c) Perfect degree of basic facilities	自然条件优劣度(d) Superiority of natural conditions	产业集聚影响度(e) Influence degree of industrial agglomeration
1	0.20	0.25	0.30	0.15	0.10
2	0.25	0.20	0.30	0.15	0.10
3	0.22	0.23	0.26	0.17	0.12
4	0.26	0.21	0.23	0.18	0.12
5	0.20	0.25	0.30	0.15	0.10
6	0.23	0.23	0.25	0.17	0.12
7	0.25	0.25	0.27	0.10	0.13
8	0.20	0.24	0.27	0.16	0.13
9	0.22	0.23	0.27	0.15	0.13
10	0.20	0.25	0.27	0.15	0.13
11	0.22	0.23	0.26	0.17	0.12
12	0.26	0.15	0.30	0.16	0.13
13	0.20	0.25	0.27	0.15	0.13
14	0.23	0.23	0.25	0.17	0.12
15	0.22	0.23	0.27	0.15	0.13
16	0.16	0.24	0.30	0.20	0.10
17	0.20	0.20	0.30	0.20	0.10
18	0.20	0.20	0.30	0.20	0.10
19	0.26	0.21	0.23	0.18	0.12
20	0.20	0.30	0.30	0.10	0.10
平均值	0.219	0.229	0.275	0.160 5	0.116 5

表2 定级因素决策规则

Table 2 Decision making rule of the grading factors

U	P				
	a	b	c	d	e
1	0	2	0	1	2
2	2	0	0	1	2
3	1	1	1	0	1
4	0	0	1	2	0
5	2	1	1	1	0
6	0	2	0	1	0
7	2	0	1	1	0
8	1	1	1	1	0
9	2	1	0	0	2
10	2	2	0	0	2
11	0	2	2	0	1
12	2	0	0	2	2

根据简化和核的定义可知, 属性 a、b、e 是其核, 属性 c、d 对决策结果没有影响, 应该省去。但是在土地定级中, 根据当地实际情况, 一旦选定定级因素, 它们对定级结果都是有效的, 都应该有相应的权重。通过进一步的分析发现, 当把{b,c}、{d,e} 等定级因素组合起来考虑时, 它们对决策结果的作用发生了很大的变化, 这说明因素之间存在着一定的

内在联系, 并共同影响决策结果。

$U \text{ind}[p(b,c)] = \{ \{1\} , \{2\} , \{3\} , \{4\} , \{5,7\} , \{6\} , \{8\} ,$

{9,10} ,{11} ,{12} } ;

$\text{POS}_{p(b,c)}(P) = \{ \{1\} , \{2\} , \{3\} , \{4\} , \{6\} , \{8\} , \{11\} , \{12\} \}$

则属性(b,c) 的重要度为:

$$\text{Imp}_{b,c} = 1 - (\text{Card}[\text{POS}_{p(b,c)}(p)] / \text{Card}[\text{POS}_p(p)]) = 1/3$$

$\text{POS}_{p(d,e)}(P) = \{ \{3\} , \{4\} , \{5\} , \{7\} , \{8\} , \{9\} , \{10\} , \{11\} \}$

则属性(d,e) 的重要度为:

$$\text{Imp}_{d,e} = 1 - (\text{Card}[\text{POS}_{p(d,e)}(p)] / \text{Card}[\text{POS}_p(p)]) = 1/3$$

根据式(4) 对属性的重要度进行归一化处理, 得出相应的权重。

$$W_a = (1/6) / [(1/6) + (1/3) + (1/3)] = 1/5$$

$$W_{b,c} = (1/3) / [(1/6) + (1/3) + (1/3)] = 2/5$$

$$W_{d,e} = (1/3) / [(1/6) + (1/3) + (1/3)] = 2/5$$

通过以上计算结果表明, 属性之间存在着一定的内在关系, 并通过不同的组合影响着决策结果, 只计算单个属性的权重就有片面性和局限性, 并不能真正反映属性的重要度。但是在土地定级中, 单元格总分值采用多因素分值加权求和法, 所以组合权重也必须按因素之间的重要程度分开。下面

以 b、c 为例计算单个属性的权重。此时决策表为  $T = T = \langle U, A \rangle$ 。其中  $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ ,  $A = \{b, c\}$ 。根据不可分辨关系计算得:

$$U/\text{ind}(A) = \{\{1, 6, 10\}, \{2, 12\}, \{3, 5, 8\}, \{4, 7\}, \{9\}, \{11\}\}$$

$$U/\text{ind}(Ab) = \{\{1, 2, 6, 9, 10, 12\}, \{3, 4, 5, 7, 8\}, \{11\}\}$$

$$U/\text{ind}(Ac) = \{\{1, 6, 10, 11\}, \{2, 4, 7, 12\}, \{3, 5, 8, 9\}\}$$

表3 综合权重计算结果

Table 3 The calculation results of integrated weight

因素 Element	主观权重 Subjective weight	客观权重 Objective weight	综合权重 Comprehensive weight
道路通达度 Road accessibility	0.219 0	0.200 0	0.211 4
对外交通便利度 Convenience degree of external traffic	0.229 0	0.191 2	0.213 9
基础设施完善度 Perfect degree of basic facilities	0.275 0	0.208 8	0.248 5
自然条件优劣度 Superiority of natural conditions	0.160 5	0.182 0	0.169 1
产业集聚影响度 Influence degree of industrial agglomeration	0.116 5	0.218 0	0.157 1

$$\text{POS}_{Ab}(A) = \{11\}$$

$$\text{POS}_{Ac}(A) = \emptyset$$

(上接第7084 页)

见空气间隙,让人看到更可靠放心,而插头拔没拔,较容易被忽视。建议家用照明灯的开关使用拉线开关,不要使用床头灯和床头开关,不要在床上拉电线,不要在床上使用220 V 电源的设备;当使用吊灯照明时,注意吊灯不宜太低,当人站立时,吊灯与人头之间应有至少1 m 的空气距离。床的安放不宜太靠近外墙,最好在床与外墙之间留一个小空气间隙。人在睡觉时,或在座时,不要直接靠在外墙上。打雷时,首先要做的就是关好门窗,防止雷电直击室内和球形雷飘进室内;不要爬到房顶上看天色、收衣物或检漏;不要停留在铁门、铁栅栏、金属晒衣绳旁;晾晒衣服被褥等用的铁丝不要拉到窗户、门口,以防铁丝引雷;在室内也要离开进户的金属水管、水龙头及与屋顶相连的下水管等。

**3.8 农村野外人身防雷** 雷雨天气应避开空旷的田野、大树下、电线杆、各类铁塔、铁轨、水边、高的孤立的建筑物旁和田野中的低矮窝棚内等危险区。要掌握“室内比室外安全、低处比高处安全、坐下蹲下比站立安全”这个原则,不要使自己成为比周围物体更高的“引雷针”。雷电来临时,首先考虑的是寻找一个可以避雷的地方,如有防雷设施或结构良好的

$$\text{Imp}_b = 1 - (\text{Card}[\text{POS}_{Ab}(A)] / \text{Card}[\text{POS}_A(A)]) = 11/12$$

$$\text{Imp}_c = 1 - (\text{Card}[\text{POS}_{Ac}(A)] / \text{Card}[\text{POS}_A(A)]) = 12/12$$

根据式(4) 对属性的重要度进行归一化处理,得出相应的权重。

$$W_b = (11/12) / [(11/12) + (12/12)] = 0.478$$

$$W_c = (1/3) / [(1/6) + (1/3) + (1/3)] = 0.522$$

$$\text{则: } W_b = 0.4 \times 0.478 = 0.1912, W_c = 0.4 \times 0.522 = 0.2088。$$

$$\text{同理可计算出 } W_d = 0.1820, W_e = 0.2180。$$

通过特尔菲法由专家给出的主观权重和根据大量数据的内在关系得出的客观权重各有自己的优缺点,在最后确定权重时应综合考虑二者的关系,根据经验,主观权重一般取0.6,客观权重取0.4,最后计算结果如表3 所示。

### 3 结论

应用粗集理论、组合权重和特尔菲法综合确定土地定级因素权重,能较好地解决以往权重确定当中以主观经验为主的弊端,使权重的确定更加科学合理。但在应用当中有些问题还需进一步研究,比如量化区间的选取与权重的关系,不同的组合对简化属性权重的影响大小等方面。

### 参考文献

- [1] 中华人民共和国国家质量监督检验检疫总局. 城镇土地分等定级规程[S]. 2001.
- [2] 曹秀英,梁静国. 基于粗集理论的属性权重确定方法[J]. 中国管理科学,2002,10(5):98-100.
- [3] 王洪凯,姚炳学,胡海清. 基于粗集理论的权重确定方法[J]. 计算机工程与应用,2003(36):20-21.
- [4] 卢新元,张金隆,从国栋. 基于粗集的IT 项目风险因素多元组合权重研究与聚类分析[J]. 管理学报,2005,2(5):527-531.
- [5] 曾黄麟. 粗集理论及其应用[M]. 重庆:重庆大学出版社,1996.

永久式建筑,或地形低矮的地方。如果雷电已经来到头顶而身体正处在空旷平坦的地方,就不能再站立行走和跑步,如果暂时找不到低矮的地方,就原地蹲下、坐下或躺下,双脚并拢。不能在头上、肩上顶着、举着或扛着任何长形的物件,甚至也不能打雨伞,随身所带的金属物品,应该暂时放在5 m 以外的地方,不要接听和拨打手机,不要听MP3,不宜开摩托车、骑自行车。

### 4 结语

农村地区防御雷电的技术难度远低于城市,关键是农民要有防雷和避雷的意识,注意掌握一定的防雷技术知识,并适当增加防雷设施的投入。目前,农村防雷减灾工作逐渐得到各级政府和有关部门的重视,加大了农村防雷科普知识的宣传力度,有的地方还出台了有关农村中小学校、新农村建设防雷减灾工作的规范性文件,相信有各级政府的重视,各有关部门的支持,农村的雷电灾害程度一定会大大降低。

### 参考文献

- [1] 中华人民共和国建设部.GB50057-94 建筑物防雷设计规范2000 年版[M]. 北京:中国计划出版社,2001.
- [2] 李家启,覃彬全,陈宏,等. 开县“5.23”重大雷电灾害事故分析[J]. 气象科技,2007,35(z1):55.
- [3] 梅卫群,江燕如. 建筑防雷工程与设计[M]. 北京:气象出版社,2004.