

应用于洪涝灾害案例推理中的层次熵分析法^①

文 刚^{1,2}, 乌 云¹, 王儒敬^{1,2}

¹(中国科学院 合肥机械智能研究所, 合肥 230031)

²(中国科学技术大学 自动化系, 合肥 230039)

摘 要: 案例推理中影响匹配案例检索的关键因素之一就是特征属性权值的分配与确定。将层次熵分析法应用于洪涝灾害案例推理中, 对灾后救灾口粮的需求进行了预估。由层次分析法分析决策问题的目标及多种层次因素, 构造出合理的层次结构。结合主观判断信息得出特征属性权值的优先级。在此基础上利用熵技术对其进行修正, 确保特征属性权值具有客观性, 使得优选案例更加切合实际情况。通过实验结果证明, 该方法适用于洪涝灾害案例推理, 可以从多项历史案例中选择与目标案例匹配最佳的案例, 从而为准确预测救灾口粮需求提供了科学的参考。

关键词: 案例推理; 层次分析法; 信息熵; 特征属性权值

Application of AHP and Entropy to the Case Based Reasoning of Flood Disasters

WEN Gang^{1,2}, WU Yun¹, WANG Ru-Jing^{1,2}

¹(Department of Automation, University of Science and Technology of China, Hefei 230031, China)

²(Institute of Intelligent Machine, Chinese Academy of Sciences, Hefei 230039, China)

Abstract: The key factor that affects the case retrieval matching in Case-based reasoning is the distribution and determination of characteristic attribute weights. In this paper, AHP and Entropy was applied in the Case-based reasoning on flood disaster to estimate the need for disaster relief food rations. AHP was used to analyze and determine the variety of factors which lead to construct a reasonable hierarchy, and get the priority of attribute weights through combining the information of subjective judgments. On this basis, using the entropy technology to modify the priority, and ensure the objectivity of characteristic attribute weights to make the preferred cases fit the reality better. Experimental results show that this method can be applied to the case-based reasoning of floods disaster, and through this method, the best case for the target case can be selected from a number of historical cases. In this way, the scientific estimation of the needs for disaster relief rations can be guaranteed.

Key words: case based reasoning; AHP; entropy technology; characteristic attribute weights

近年来, 案例推理 (Case Based Reasoning, CBR) 被广泛地应用于自然灾害物资需求的预测中^[1,2]。通过类似灾害问题的历史经验和获取的救灾知识进行推理, 来解决和应对新问题。案例推理主要包括: 案例的表示与组织、特征变量抽取、案例的索引与检索、案例的调整等关键技术, 其中特征属性权值的合理分配是影响案例匹配的关键因素之一。姜枫等^[3]提出了一种基于时序的案例权重调整算法, 充分考虑了时间对

案例匹配度的影响; 王津津等^[4]在对特征项赋权中将主客观赋权法相结合, 并用遗传算法进行了优化。但上述方法都缺乏一个系统化、层次化的分析方法, 因此无法充分考虑到各属性对案例检索的影响。

目前作为系统化、层次化的分析方法层次熵分析法被广泛应用于管理评价、资源规划及经济分析等方面。国内, 徐菲菲等^[5]将层次熵分析法应用于风景名胜区规划方案的评价中, 从多项方案中选择最佳方案,

① 基金项目: 国家科技支撑计划(2008BAK49B05)

收稿时间: 2011-08-03; 收到修改稿时间: 2011-09-11

为各级、各类旅游目的地规划方案提供了参考。黄典剑等^[6]运用层次熵分析法对城市地铁突发事件应急能力进行了综合评价,并建立了城市地铁运营的风险评价模型。国外, Jalalifar 等^[7]将层次熵分析法与逼近理想解排序法相结合,建立了采矿方案的优选模型。Tavana 等^[8]则将层次熵分析法应用于载人航天模拟器的评定中,实现了基于该理论的决策支持系统。但利用层次熵分析法对自然灾害案例进行分析研究的尚未普及。由于洪涝灾害是受多种指标因素影响的复杂的自然现象,且具有流域层次特征,各指标之间互不相同,缺乏可比性等特点。基于此本文提出将层次熵分析法应用于洪涝灾害案例推理中,并对灾后救灾口粮的需求进行了预估。以多层次评价方法量化案例特征属性的隶属关系,进而计算确定出属性权重的优先级。并结合信息熵技术对其加以修正,确保特征属性权值具有客观性。通过这种主客观相结合的方法对案例的总体加以综合评价,最终检索出与目标案例匹配最佳案例,对救灾口粮的需求进行预测。

本文首先介绍层次熵分析法的基本原理。其次将层次熵分析法应用于洪涝灾害案例推理中进行了实验及分析。结果表明层次熵分析法兼顾主客观两个方面,使得优选的案例更加贴近实际要求。

1 层次熵分析法的基本原理

1.1 层次分析法确定权重

层次分析法 (Analytic Hierarchy Process, AHP) 是美国运筹学家 Saaty 提出的,解决具有复杂层次结构的多目标决策问题的一种有效方法^[9]。它的核心是将待评价复杂系统的各要素按其关联隶属度建立递阶层次模型,构造两两比较的判断矩阵,据此求解各要素重要性的排序权值,并检验判断矩阵的一致性。本文将 AHP 法应用于洪涝灾害案例推理中,借助它将决策者的经验判断定量化,确定案例特征属性权值。具体步骤如下:

Step1: 建立层次结构模型。

采用顶、中、底三层结构。最上层为目标层,记为 A,只有一个因素;中间层为准则层,记为 B;最下层为对象层,记为 C。

Step2: 构造判断矩阵。

对 step1 建立的层次结构,将隶属于同一层 n 个指标两两比较得判断矩阵 $A=(a_{ij})_{n \times n}$ 。其中, a_{ij} 表示指标 i 对

指标 j 的重要程度。本文采用 Saaty 的 1-9 比率标度法,以 {1,3,5,7,9} 分别表示 i 对 j {同样、稍微、明显、强烈、极端重要},以 {2,4,6,8} 分别表示介于上述奇数之间的重要度。

Step3: 求特征向量 (权重向量)。

向量 $W=(W_1, W_2, \dots, W_n)$ 为同一层各元素相对上一层某元素的重要性排序权值。对于矩阵 $(a_{ij})_{n \times n}$ 的特征向量,采用方根法求 n 次方根,并将其归一化得特征值 W_i (如式 (1)),进而得到相似特征向量 $W=(W_1, W_2, \dots, W_n)T$ 。

$$W_i = \frac{\bar{W}_i}{\sum_{i=1}^n \bar{W}_i} \quad (1)$$

$$\bar{W}_i = \sqrt[n]{a_{i1} \cdot a_{i2} \cdot \dots \cdot a_{in}} \quad (i=1, 2, \dots, n)$$

Step4: 计算判断矩阵的最大特征根。

在 AHP 法中计算过判断矩阵最大特征值 λ_{\max} 不需要很高的精度,一般用平均值近似计算即可,如式 (2) 所示。其中, AW 为判断矩阵 A 和特征向量 W 的乘积; $(AW)_i$ 为向量 AW 的第 i 个元素。

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i} = \frac{1}{n} \sum_{i=1}^n \frac{\sum_{j=1}^n a_{ij}W_j}{W_i} \quad (2)$$

Step5: 做一致性检验。

为保证得到的权重合理,要对每一个判断矩阵进行一致性检验,如式 (3) 所示,计算随机一致性 CR。当不满足 $CR < 0.10$ 时,需修改判断矩阵。RI 为平均随机变量指标,按表 1 所示取值。

$$CR = \frac{\lambda_{\max} - n}{RI(n-1)} \quad (3)$$

表 1 随机一致性指标 RI 值

维数 n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Step6: 计算合成权重。

以自上而下的形式,逐层地将单一准则的权重进行合成,计算出底层各元素的权重和总的一致性检验值。具体计算如式 (4) 所示。 $W^{(k)}$: 第 k 层上 nk 个元素对于总目标的合成排序权重向量; $P^{(k)}$: 第 k 层上 nk 个元素对 k-1 层上所有元素为准则的排序权重向

量; $W^{(k-1)}$: 第 $k-1$ 层上 $nk-1$ 个元素对于总目标的合成排序权重向量。

$$W^{(k)} = [W_1^{(k)}, W_2^{(k)}, \dots, W_{nk}^{(k)}]^T$$

$$= p^{(k)} W^{(k-1)} = p^{(k)} [W_1^{(k-1)}, W_2^{(k-1)}, \dots, W_{nk-1}^{(k-1)}]^T \quad (4)$$

1.2 熵技术修正权重

AHP 法将人的主观判断为主的定性分析进行量化, 用数值来表示各个候选项的差异, 以供参考。因此不能保证其计算出的权值具有客观性。为解决这一不足采用熵技术对 AHP 法得到的权重向量进行修正。具体步骤如下:

Step1: 属性归一化。

由于源案例中各个特征属性值具有取值范围及数值类型等特点, 因此熵技术修正前需进行归一化处理。将源案例中的特征属性值所组成的矩阵设为 $C=(C_{ij})_{m \times n}$, m 和 n 分别表示源案例与特征属性的个数。令 C_{jmax} 和 C_{jmin} 分别为所有案例中第 j 个属性的最大值和最小值, 则归一化后的 P_{ij} 值为式 (5) 所示。

$$P_{ij} = \frac{C_{ij} - C_{min}^j}{C_{max}^j - C_{min}^j} \quad (5)$$

Step2: 求属性的熵值 e_j 和差异系数 g_j 。

第 j 项属性的熵值 e_j 的计算公式为式 (6) 所示。其中, $k = (\ln m)-1$ 。从而得出第 j 项属性的差异系数 g_j 为 $g_j = 1 - e_j$ 。

$$e_j = -k \sum_{i=1}^m p_{ij} \ln p_{ij} \quad 0 \leq e_j \leq 1 \quad (6)$$

Step3: 求信息权重系数 μ_j , 即修正系数。

第 j 项属性的信息权重系数 μ_j 的计算公式如式(7)所示。

$$\mu_j = \frac{g_j}{\sum_{j=1}^n g_j} \quad (7)$$

Step4: 进行修正。

利用信息权重系数 μ_j 修正 AHP 法得出的属性权重系数 λ_j , 如公式(8)所示。

$$\lambda_j = \frac{\mu_j W_j}{\sum_{j=1}^n \mu_j W_j} \quad (8)$$

Step5: 计算修正后的权重向量。

通过熵技术进行修正后的属性权重向量为:

$$W = (\lambda_1, \lambda_2, \dots, \lambda_n)^T \times (W_1, W_2, \dots, W_n) \quad (9)$$

2 案例研究

2.1 洪涝灾害案例库的选取

利用民政部减灾中心提供的, 1998-2007 年间发生在安徽地区的 8 条洪涝灾害案例作为实验数据。选取受灾人口 (千人) C_1 、紧急转移安置人口 (千人) C_2 、农作物受灾面积 (千公顷) C_3 、农作物绝收面积 (千公顷) C_4 、倒塌房屋间数 (千间) C_5 、损坏房屋间数 (千间) C_6 、直接经济损失 (千万元) C_7 作为决策属性, 需救济粮数量 (吨) TC 作为目标属性。具体的特征属性值如表 2 所示:

表 2 洪涝灾害案例数据

案例编号	决策属性							目标属性
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	TC
T_1	14600	1270	1386	286	428	1878	127	889
T_2	11730	189	758	116	98	147	225	132.3
T_3	5768	106	396	50	39	24	62	74.2
T_4	11240	64	789	73	40	83	114	44.8
T_5	1849.1	7.792	105.043	22.993	9.87	17.474	27.06	5.45
T_6	9760	102	699	121	51	93	219	71.4
T_7	3680	39	3480	475.8	10.637	36.4	87.7	27.3
T_8	1336	4.032	101	7.9	2.825	4.381	26	2.82

2.2 层次结构评价指标的确定

根据民政局灾害统计类别、安徽省气象灾害年鉴等可知，洪涝灾害所造成的损失主要是从人口受灾情况，农作物受灾情况，经济损失情况等三个方面进行统计的。以此为据本文建立了如表 3 所示的洪涝灾害救灾口粮的评价决策层次结构。

表 3 洪涝灾害救灾口粮评价决策层次结构

目标层 A	准则层 B		对象层 C	
救灾口粮	人口受灾情况	B1	受灾人口	C1
			紧急转移安置人口	C2
	农作物受灾情况	B2	农作物受灾面积	C3
			农作物绝收面积	C4
	经济损失情况	B3	倒塌房屋间数	C5
			损坏房屋间数	C6
			直接经济损失	C7

2.3 参数的确定及权重的计算

1) AHP 法计算权重

各层判断矩阵的参数先由专家经验给定。其次按照 AHP 法计算步骤得出各层权重向量，结果见表 4-表 7:

表 4 准则层对目标层判断矩阵 A: B

A	B1	B2	B3	WA	$\lambda_{max} = 3.0037$ $RI = 0.58$ $CR = 0.0011 < 0.1$
B1	1	3	5	0.6483	
B2	1/3	1	2	0.2297	
B3	1/5	1/2	1	0.1220	

表 5 人口受灾情况评价判断矩阵 B1: C1-C2

B1	C1	C2	WB1	二阶矩阵不需要作一致性检验
C1	1	2	0.6667	
C2	1/2	1	0.3333	

表 6 农作物受灾情况评价判断矩阵 B2: C3-C4

B2	C3	C4	WB2	二阶矩阵不需要作一致性检验
C3	1	2	0.6667	
C4	1/2	1	0.3333	

表 7 经济损失情况评价判断矩阵 B3: C5-C7

B3	C5	C6	C7	WB3	$\lambda_{max} = 3.2174$ $RI = 0.58$ $CR = 0.0630 < 0.1$
C5	1	1/3	1/6	0.0805	
C6	3	1	1/8	0.1522	
C7	6	8	1	0.7672	

由上表可得: $W^{(0)} = [0.6483, 0.2297, 0.1220]$,

$$W^{(1)} = \begin{bmatrix} 0.6667 & 0.3333 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6667 & 0.3333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0805 & 0.1522 & 0.7672 \end{bmatrix}$$

从而可得决策属性权重向量为:

$$W = W^{(0)} W^{(1)} = [0.4322, 0.2161, 0.1531, 0.0766, 0.0098, 0.0186, 0.0936]$$

2) 熵技术修正权重

对以上 AHP 法得到的权重 W，利用熵技术进行修正。修正中的各决策属性参数值为表 8 所示。案例个数 $m = 8$ ， $k = (\ln 8) - 1 = 0.48$ 。

表 8 决策属性参数值

	C1	C2	C3	C4	C5	C6	C7
ej	0.8248	0.4837	0.7055	0.7354	0.5658	0.3383	0.5208
gj	0.1752	0.5163	0.2945	0.2646	0.4342	0.6617	0.4792
uj	0.062	0.1827	0.1042	0.0936	0.1537	0.2342	0.1696
wj	0.2412	0.3552	0.1435	0.0645	0.0136	0.0392	0.1428

根据上表可得，利用熵技术进行修正后的权重向量为:

$$W' = [0.2412, 0.3552, 0.1435, 0.0645, 0.0136, 0.0392, 0.1428]$$

3) AHP 法与熵技术修正后的权重向量对比

AHP 法与熵技术修正后的决策属性权重系数对比结果如图 1 所示。由图 1 可知，利用熵技术对权重系数进行修正后，救灾口粮的最大影响因子由受灾人口 (C1) 转变为紧急转移安置人口 (C2)。这与所需救灾口粮的主要参考依据为紧急转移安置人口的实际情况相符。

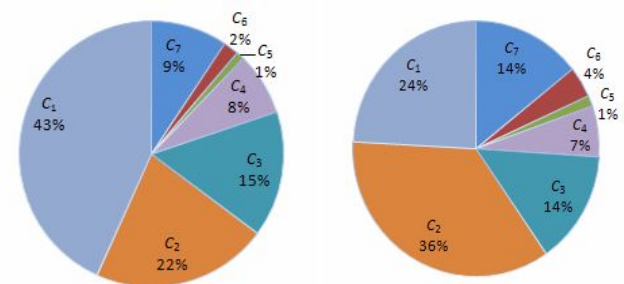


图 1 熵技术修正前后权重向量对比

2.4 案例匹配

选目标案例 $T = \{745000, 13100, 51600, 9800, 1640\}$

18000, 24000, 9.17}, 作为测试数据, 验证层次熵分析法对洪涝灾害案例推理的有效性。通过对各个源案例决策属性权重进行调整, 检索与目标案例匹配最佳的案例, 从而对救灾口粮的需求进行预测。匹配度 P 的计算公式如式 (10) 所示。 W_i 为属性权重向量中第 i 个权重因子; D_i 表示目标与源案例间第 i 个特征属性的距离; Y_i 和 y_i 分别表示源案例库和目标案例库中第 i 个特征属性的值。相对于目标案例的各个源案例的匹配度 P 的计算结果如表 2.8 所示。由表可知, AHP 法得到的最相似案例是 T8, 预测救灾口粮为 2.82 吨; 熵修正后的得到的最佳案例为 T5, 预测救灾口粮为 5.45 吨, 更接近实际的 9.17 吨。

$$P = \frac{\sum_{i=1}^n (1 - D_i) W_i^2}{\sum_{i=1}^n W_i} \quad 0 \leq W_i \leq 1 \quad (10)$$

3 结语

层次熵分析法虽然在社会、经济和管理等领域得普及。本文尝试将层次熵分析法应用于洪涝灾害案例到了广泛应用, 但在自然灾害案例推理中还没有得到推理中, 并对灾后救灾口粮的需求进行了预测。实验结果表明该方法能够从多项历史案例中检索出与目标案例最相似的案例, 科学地提供一个进行决策目标预测的参考值。本次试验选取的案例个数较少, 今后考

虑增大历史案例数据库的同时改进层次熵分析法, 并将大于匹配度临界值的案例施行多元回归模型等, 对灾后应急物资的需求进行精确的预测。

参考文献

- 傅志妍, 陈坚. 灾害应急物资需求预测模型研究. 物流科技, 2009, 32(10): 11-13.
- 王晓, 庄严明. 基于案例推理的非常规突发事件资源需求预测. 西安电子科技大学学报: 社会科学版, 2010, 20(4): 22-26.
- 姜枫, 王儒敬, 孙丙宇, 李文波, 桂圆苗. CBR 灾害救助系统一种权重调整算法. 电子技术, 2010. 9.
- 王津津, 田卫东. 案例推理中的一种组合权重方法. 微计算机信息, 2010(26): 254-256.
- 徐菲菲, 刘沛林, 白先春, 宋平. 风景名胜区规划方案的层次分析法与熵技术评价. 地理研究, 2004, 5(23): 396-400.
- 黄典剑, 李传贵. 突发事件应急能力评价-以城市地铁为对象. 北京: 冶金工业出版社, 2006.
- Jalalifar H, Behadini M, Aghajani B. The optimum rock bolt support system selection by using AHP-Entropy-TOPSIS method. Journal of Mines, Metals and Fuels, 2009, 57(9): 251-266.
- Tavana M, Hatami-Marbini A. A group AHP-TOPSIS framework for human spaceflight mission planning at NASA. Expert Systems with Applications, 2011, 38(11): 13588-13603.
- Saaty TL. Decision making the analytic hierarchy and network processes. Journal of Systems Science and Systems Engineering, 2004, 13(1): 1-35.

表 9 熵技术修正前后的匹配度对比

源案例编号	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
权重调整法								
AHP	0.0675	0.1090	0.2265	0.1571	0.7038	0.1410	0.3393	0.8123
AHP + 熵技术	0.0335	0.0994	0.2140	0.2173	0.6686	0.1502	0.3655	0.5138