

基于匹配属性相似度的应急决策方案推荐方法

董庆兴¹, 李 赛¹, 张大斌², 李延晖¹

(1. 华中师范大学 信息管理学院, 武汉 430079; 2. 华南农业大学 数学与信息学院, 广州 510642)

摘要: 针对目标案例与历史案例的非一致属性集合问题, 基于匹配度生成子案例库, 使用基于核密度的群体序关系分析法对目标案例各属性进行赋权. 利用符号型属性和数值型属性的特征, 通过不同相似度计算方法分别计算各属性下的相似度, 集结各属性相似度信息和权重信息得到非一致属性集合的历史案例与目标案例的相似度, 最终得到应急方案的推荐结果. 最后通过算例表明了该方法的有效性和可行性.

关键词: 应急决策; 案例检索; 序关系分析法; 相似度计算

中图分类号: C934

文献标志码: A

Emergency decision response plan recommendation method based on similarity of matched attributes

DONG Qing-xing¹, LI Sai¹, ZHANG Da-bin², LI Yan-hui¹

(1. School of Information Management, Central China Normal University, Wuhan 430079, China; 2. School of Mathematics and Information, South China Agricultural University, Guangzhou 510642, China. Correspondent: DONG Qing-xing, E-mail: qxdong@mail.ccnu.edu.cn)

Abstract: A method based on the attribute set of target case is proposed to retrieve historical cases from the case warehouse, so as to get the qualified historical case to form a sub-case warehouse. Then, a G1 and Kernel Density based group decision method is used to allocate the weights of the target case's attributes. Considering the features of numerical attributes and symbolic attributes, the corresponding similarity functions are used to calculate the similarities of cases under different attributes. Thus the similarities between historical cases and target case can be obtained. Finally, the historical response action to the emergency decision target case can be recommended according to the rank of their similarities. A numerical example is given to illustrate the effectiveness and the feasibility of the proposed method.

Keywords: emergency decision; case retrieval; G1 method; similarity calculation

0 引言

突发事件, 是指突然发生, 造成或者可能造成严重社会危害, 需要采取应急处置措施予以应对的自然灾害、事故灾难、公共卫生事件和社会安全事件^[1]. 由于突发事件具有事发突然、不确定性高、发展演变情境复杂等特点, 当突发事件发生时, 应急决策主体必须依据事件不断发生、发展的态势, 针对具体的情景, 快速响应并作出有效的应对决策, 最大限度地减少突发事件所造成的损失^[2-4]. 因此, 依据当前突发事件发生时的一些特征和信息, 如何借鉴历史案例应急响应的相关信息, 生成当前突发事件的可行的应急方案, 可以避免重复决策耗费的时间和成本^[5-6]. 案例

推理(CBR)是一种通过访问知识库中历史案例(过去发生的类似突发事件)的求解过程和结果, 为目标案例(当前突发事件)的解决提供经验和知识的方法, 已经被广泛应用到了应急决策中^[7]. 姜艳萍等^[8]针对森林火灾突发后快速应急响应决策问题, 给出了一种相似度计算的应急响应方案选择方法; 李永海等^[9]提出了一种基于案例决策理论的解决广义不确定型决策问题的案例决策方法; Zhang等^[10]给出一个集成了CBR推理方法的灾难救助辅助决策支持的理论知识管理框架; Fan等^[11]针对风险项目管理提出了一种基于CBR原理的风险应对策略框架. 但是已有的基于CBR的应急响应决策方法在计算案例之间的相似度

收稿日期: 2015-06-07; 修回日期: 2015-09-13.

基金项目: 国家自然科学基金项目(71501080, 71501054); 中央高校基本科研业务费专项基金项目(CCNU15A05050, CCNU14Z02016).

作者简介: 董庆兴(1988—), 男, 讲师, 从事群体评价、群体决策等研究; 李赛(1989—), 男, 硕士生, 从事机器学习、数据挖掘的研究.

时,都是假设案例具有一致的属性(指标)集合,进而得到案例间的相似度.然而,在现实突发事件应急管理实践中,当前面临的突发事件(目标案例)的属性集合与历史案例的属性集合未必完全一致,因此不能直接通过一致属性集合模式下的传统相似度函数计算案例间的相似度.其次,基于CBR的技术,要求案例库中存储大量的案例来满足用户对知识的需求,然而随着案例库中案例数量的增长,案例相似度计算量会大幅增加,从而不满足应急决策的时间要求.此外,由于突发事件的各属性重要性往往不同,如何合理地为目标案例的属性进行赋权,直接影响到应急方案的推荐效果.

本文提出一种基于目标案例属性的历史案例检索方法,将满足条件的历史案例组合成新的子案例库;然后使用基于核密度的群体序关系分析法对目标案例各属性进行赋权,利用符号型属性和数值型属性的特征,通过不同相似度计算方法分别计算各属性下的相似度,集结各属性相似度信息和权重信息得到非一致属性集合的历史案例与目标案例的相似度,最终得到应急方案的推荐结果.

1 问题描述

CBR的推理过程是利用基于记忆的历史案例来求解新问题的类比推理过程,对用户输入的目标案例与系统中的案例库中的历史案例集合进行相似度计算,将与目标案例最相似的案例返回给用户,并对历史案例的解决方案进行重用.本文为叙述方便,用 P^C 表示当前需要解决的应急决策问题,即目标案例. $Z = \{(P_1, A_1), (P_2, A_2), \dots, (P_n, A_n)\}$ 表示由 n 个历史案例构成的集合,其中 (P_j, A_j) 表示第 j 个案例, P_j, A_j 分别表示 (P_j, A_j) 的问题和解决方案.

一个案例通常是由多个属性来描述且每个案例的属性集合可能不完全一样,例如,在高速公路车祸应急救援历史案例中,该案例所涉及的问题可由公路类型、天气类型、事故车辆数量、有无危险品等属性来描述.设 S^C 为目标案例所对应的属性集合,即 $S^C = \{s_1^C, s_2^C, \dots, s_{m_c}^C\}$,其中 m_c 为目标案例问题属性的个数, s_i^C 为目标案例 P^C 的第 i 个属性, $i \in [1, m_c]$, $V(s_i^C)$ 为目标案例 P^C 在属性 s_i^C 上的取值.历史案例 P^j 对应的属性集合为 $S_j = \{s_1^j, s_2^j, \dots, s_{m_j}^j\}$,其中 m_j 为历史案例 P^j 所具有的属性个数; s_i^j 为历史案例 P^j 的第 i 个属性, $i \in [1, m_j]$; $V(s_i^j)$ 为历史案例 P^j 在属性 s_i^j 上的取值.依据现实情况,案例的属性值可分为数值型和符号型两种形式.前例中属性“事故车辆数量”的属性值一般是数值型的,属性“天气类型”的属性值一般是符号型的,如采用“晴天”、“雨雪”或“大雾”等符号来表述.

2 应急决策方案推荐方法

传统的基于案例推理的决策方法是在目标案例与历史案例具有一致的属性集合的基础上计算案例间的相似度,最后通过提取相似度最大的历史案例的实施方案作为决策参考方案来完成决策.但是,在实际中,突发事件与历史案例由于事件自身的特征,都有自己独特的属性集合.属性集合的不完全相同,导致其属性值向量未必等长,或者不同的属性值向量在同样索引下的值代表不同的属性.假设目标案例 P^C 具有 a, b, c 三个属性,对应的属性值向量为 $n^C = (3, 4, 1)$;某个历史案例 P^1 具有 a, b, d 三个属性,对应的属性值向量为 $n^1 = (3, 4, 1)$;历史案例 P^2 的属性为 (a, b) ,对应的属性值向量为 $n^2 = (3, 4)$.但是,由于 P^C, P^1, P^2 具有不一致的属性集合, P^C 与 P^1 和 P^2 的相似度不能直接通过一致属性集合模式下的传统相似度函数计算案例间的相似度.因此,需要通过属性检索的方式,对历史案例进行筛选组成新案例库,再利用匹配属性计算方案间的相似度.

2.1 历史案例的选取

在应急决策过程中,若要计算目标案例与每一个历史案例的相似度,在案例库中的历史案例较多的情况下,计算的时间成本较大,很难满足应急决策的时限性要求.为此,需要对历史案例数据库进行筛选,在筛选过的子案例库中进行相似度计算.

定义 1 若属性 s_i 同时存在于目标案例 P^C 的属性集合 S^C 和历史案例 P^j 的属性集合 S_j 中,则称 s_i 为 P^C 和 P^j 的匹配属性.由 P^C 和 P^j 的所有匹配属性构成的属性集合称为 P^C 和 P^j 的匹配属性集合,记为 $S_{(c,j)}$. P^C 和 P^j 的匹配属性集合 $I_{(c,j)}$ 中的匹配属性个数定义为 P^C 和 P^j 的匹配度.

依据匹配属性和匹配属性集合的定义,有如下筛选规则:一般情况下,目标案例与各个历史案例的匹配度是不一致的,若某个历史案例 P^j 与目标案例 P^C 没有共同属性,则该历史案例与目标案例的匹配度为0;若某个历史案例与目标案例的匹配度比较大(即匹配属性集合较大),则该历史案例与目标案例的相似度倾向于高,反之倾向于低.因此,与目标案例匹配度高的历史案例应该被优先选出构建子案例库 Z_{sub} .其具体算法流程如下.

Require: 案例库中的所有历史案例;

Ensure: 筛选后的案例组成的子案例库 Z_{sub} .

1) 给定目标案例 P^C ,对应的属性集合为 $S^C = s_1^C, s_2^C, \dots, s_{(m_c)}^C$;

2) 设案例库中历史案例的个数为 len ,实例化一个元素全部为0、长度等于 len 的数组 Index ,实例化

新的案例库 $Z_{\text{sub}}, Z_{\text{sub}} = \emptyset$;

3) for each $j \in 1 : \text{len do}$;

4) 设 $\text{count} = 0, m_j$ 为历史案例 P_j 中的属性个数;

5) If $s_i^j \in s^C$, then 当前历史案例的属性同时存在于目标案例中;

6) $\text{count} = \text{count} + 1$;

7) end if;

8) end for;

9) $\text{Index}[j] = \text{count}$, $\text{Index}[j]$ 代表第 j 个历史案例与目标案例的匹配度;

10) 若 Index 中的元素全部为0, 则将目标案例直接提交给专家进行处理, 否则对 Index 从大到小排序, 选出 $\text{Index}[j]$ 较大的值对应的历史案例 P_j 添加到 Z_{sub} 中构成新的子案例库.

由上述算法即可构造出筛选过的子案例库, 接下来则可以计算该子案例库中案例与目标案例间的相似度.

2.2 目标案例的属性赋权

在案例相似度计算过程中, 属性的权重用来反映属性的相对重要性, 需要专家集体决策, 因此如何有效地集结专家判断是一个重要问题. 本文首先通过序关系分析法得到各专家对目标案例的属性权重判断, 由于目标案例的属性所取的权重值服从未知的概率分布, 通过一种非参数的核密度估计方法将专家个体给出的判断本身视为该未知分布下的一个样本, 并拟合出相应的概率密度分布; 然后判断专家的准确性并取概率最大的值作为属性权重. 该方法可以有效避免个别误差较大的权重对最后集结的权重的影响^[12].

2.2.1 目标案例的属性赋权

序关系分析法在属性赋权中应用广泛, 其优势在于无需进行一致性检验, 计算过程比较简单. 其核心思想是: 首先对被评价对象由大到小排序; 然后比较相邻两个被评价对象的相对重要程度, 得到被评价对象的权重, 基本流程如下^[15].

1) 不失一般性, 设 $x_1, x_2, \dots, x_m (m \geq 2)$ 是 m 个指标, 若指标 x_i 相对某评价准则 (或目标) 的重要程度不劣于 x_j 时, 则记为 $x_i \succeq x_j$ (符号 \succeq 表示不劣于关系). 将指标 $x_1, x_2, \dots, x_m (m \geq 2)$ 相对某评价准则建立关系式, 有

$$x_1^* \succeq x_2^* \succeq \dots \succeq x_m^*, \quad (1)$$

则称指标 $x_1, x_2, \dots, x_m (m \geq 2)$ 之间按 \succeq 建立了序关系. 这里 x_i^* 表示 $x_1, x_2, \dots, x_m (m \geq 2)$ 按序关系 \succeq 排定顺序后的第 $i (i = 1, 2, \dots, m)$ 个评价指标. 为书写方便, 以下仍记 x_i^* 为 x_i .

2) 给出 x_{k-1} 与 x_k 间的相对重要程度的比值判断. 设专家关于评价指标 x_{k-1} 与 x_k 的重要程度之比的理性判断为

$$r_k = \frac{w_{k-1}}{w_k}, \quad (2)$$

其中 $k = m, m-1, \dots, 3, 2$. r_k 的取值如表 1 所示.

表 1 r_k 赋值参考表^[13]

r_k 取值	意义
1.0	指标 x_{k-1} 与 x_k 具有相同重要性
1.2	指标 x_{k-1} 比 x_k 稍微重要
1.4	指标 x_{k-1} 比 x_k 明显重要
1.6	指标 x_{k-1} 比 x_k 强烈重要
1.8	指标 x_{k-1} 比 x_k 极端重要

3) 权重系数 w_k 的计算. 若专家给出 r_k 的理性赋值, 则 w_m 为

$$w_m = \left(1 + \sum_{k=2}^m \prod_{i=k}^m r_i \right)^{-1}. \quad (3)$$

而由式 (2) 可知

$$w_{k-1} = r_k w_k, \quad (4)$$

从而可得各指标的权重.

2.2.2 基于核密度的专家判断集结

核密度集结方法利用专家对目标案例的各个属性所可能取的权值判断得到相应的权值集合, 通过对属性的权值集合运用平滑的高斯核密度函数进行拟合, 可以得到每一个属性权值的概率密度曲线 $f(x)$, 其中 x 为该属性所可能取的权值, $f(x)$ 为该属性取的权值 x 所对应的概率. 不失一般性, 令使 $f(x)$ 取最大时的 x 即为该属性所取的权值.

定义 2 设 x_1, x_2, \dots, x_n 为取值于 \mathbf{R} 的独立同分布随机变量, 其所服从的分布密度函数为 $f(x), x \in \mathbf{R}$. 定义函数

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \quad (5)$$

为密度函数 $f(x)$ 的核密度估计, 其中 $K(\cdot)$ 称为核函数, h 为预先给定的正数, 通常称为窗宽或光滑参数. 不失一般性, 本文采用的核函数为高斯函数

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \quad (6)$$

设 $E = \{e_1, e_2, \dots, e_q\}$ 为专家集合, 其中 e_i 为第 i 个专家, q 为专家个数. 当专家 e_i 在对目标案例 P^C 的属性集合 (即 $S^C = \{s_1^C, s_2^C, \dots, s_{m_c}^C\}$, m_c 为目标案例的属性个数) 的属性进行赋权时, 首先利用序关系分析法对 P^C 的属性进行有序排列, 通过式 (2) 和 (3) 得到表示专家 e_i 对属性 s_j^C 的赋权 $w_i^j, j = 1, 2, \dots, m_c$. 由此可以得到属性 s_j^C 的权值向量为 $w_j = (w_1^j, w_2^j, \dots, w_q^j)$. 不失一般性, 假设每位专家都是独立给出判断, 因此属性 s_j^C 的权值向量的每一个元素也互相独立, 运用高斯核密度函数可以得到属性

s_j^C 的权值向量 $\mathbf{w}_j = (w_1^j, w_2^j, \dots, w_q^j)$ 的概率密度曲线 $f(x_j^C)$, 选择使得该密度函数取极大值时对应的 x_j^C 作为属性 s_j^C 的权值, 即

$$v_j^C = \arg \max f(x_j^C). \quad (7)$$

重复以上步骤, 有目标案例的权值向量 $\mathbf{v}^C = (v_1^C, v_2^C, \dots, v_{m_c}^C)$. 最后进行归一化, 有

$$w_j^C = \frac{v_j^C}{\sum_{j=1}^{m_c} v_j^C}, \quad (8)$$

其中 $j = 1, 2, \dots, m_c$. 从而可以得到目标案例 P^C 的属性的权重向量 $\mathbf{w}^C = (w_1^C, w_2^C, \dots, w_{m_c}^C)$.

2.3 案例相似度计算

在计算目标案例 P^C 与历史案例 Z_{sub} 中每一个历史案例 P_j 的相似度之前, 首先得到 P^C 与 P_j 的匹配属性集合 $S_{(C,j)} = \{s_1, s_2, \dots, s_h\}$, h 为 $S_{(C,j)}$ 里的匹配度. 属性有符号型和数值型两种, 因此对 $S_{(C,j)}$ 中的每一个属性 s_i , P^C 和 P_j 在该属性下的相似度 $\text{Sim}_i(P^C, P_j)$ 的计算有以下两种方法:

1) 数值型属性相似度计算. 当属性 s_i 的属性值为数值型时, 采用下式计算:

$$\text{Sim}_i(P^C, P_j) = 1 - \frac{|V(s_i^C) - V(s_i^j)|}{\alpha_i^{\max}}. \quad (9)$$

其中: $V(s_i^C), V(s_i^j)$ 分别为 P^C 和 P_j 在属性 s_i 上的取值, $\alpha_i^{\max} = \max\{V(s_i^C), \max\{V(s_i^j) | s_i \in S_{(C,j)}\}\}$.

2) 符号型属性相似度计算. 当属性 s_i 的属性值为符号型时: 假设 s_i 为“天气情况”属性, 若 P^C 与 P_j 的天气情况不一样, 则 P^C 和 P_j 在该属性下的相似度 $\text{Sim}_i(P^C, P_j) = 0$; 若天气情况一样, 则直接令 $\text{Sim}_i(P^C, P_j) = 1$ 是常见的处理方法^[13]. 但是, 这种方法首先会造成案例间相似度的极化现象, 相似度为 0, 或者为 1. 例如, 案例库中共有 10 个案例, “天气情况”属性 s_i 共有 3 种情况, 分别为: “晴天”、“雨雪”、“大雾”. 记 $n(V(s_i) = \text{“晴天”}) = 4$, $n(V(s_i) = \text{“雨雪”}) = 2$, $n(V(s_i) = \text{“大雾”}) = 4$ 分别为这 3 种天气类型在案例库中出现的次数. 若案例库中案例 P_a 和 P_b 的天气属性都为“雨雪”, 案例 P_m 和 P_n 的天气属性都为“晴天”, 则按照传统的计算方法, $\text{Sim}_i(P_a, P_b) = \text{Sim}_i(P_m, P_n) = 1$. 但是, 由于“晴天”比“雨雪”在案例库中出现的频繁, P_a 和 P_b 的相似度应该大于 P_m 和 P_n 的相似度. 因为如果两种案例恰好出现了某种非常罕见的天气, 则其相似度应该更高. 因此, 本文在计算符号型属性之间的相似度时, 需要考虑符号型属性特征值的出现频率. 梁昌勇等^[16]将遗传算法和条件概率整合到案例推理中, 提出了一种适应离散变量诊断决策案例的案例推理方法. 但是, 鉴于突发事件决策对时间紧迫性的要求, 本文的符号属

性相似度计算方法不仅满足了时间紧迫性, 而且避免了符号型属性的极化现象. 设符号型属性 s_i 共有 r 个水平, 记为 $D_i = \{d_1, d_2, \dots, d_r\}$, 则每一个水平出现的概率为

$$f(V(s_i) = d_k) = \frac{n(V(s_i) = d_k)}{l_{\text{sub}}}. \quad (10)$$

其中: $k = 1, 2, \dots, r$, l_{sub} 为案例库中历史案例的个数, $n(V(s_i) = d_k)$ 为 s_i 属性下水平 d_k 出现的次数. 因此, 按照如下两种情况即可计算得到目标案例 P^C 与历史案例 P_j 在符号型属性 s_i 下的相似度 $\text{Sim}_i(P^C, P_j)$:

① 当 $V(s_i^C) \neq V(s_i^j)$ 时, 有

$$\text{Sim}_i(P^C, P_j) = 0; \quad (11)$$

② 当 $V(s_i^C) = V(s_i^j) = d_k$ 时, 有

$$\text{Sim}_i(P^C, P_j) = \frac{\min\{f(V(s_i) = d_v) | d_v \in \{D_i\}\}}{f(V(s_i) = d_k)}. \quad (12)$$

由上述相似度计算方法可得, 历史案例 P_j 与目标案例 P^C 的相似度为

$$\text{Sim}(P^C, P_j) = \sum_{k=1}^h w_k \text{Sim}_k(P^C, P_j). \quad (13)$$

其中: h 为 P^C 和 P_j 的公共属性集合 $S_{(C,j)} = \{s_1, \dots, s_h\}$ 的属性个数, $\text{Sim}_k(P^C, P_j)$ 为 P^C 和 P_j 在公共属性集合 $S_{(C,j)}$ 的属性 s_k 的相似度, $k = 1, 2, \dots, h$, w_k 为属性 s_k 对应的权值.

综上所述, 基于匹配属性相似度的应急决策方案的生成步骤归纳如下.

Step 1: 首先基于属性匹配, 从历史案例库中筛选出与目标案例具有匹配属性的案例作为新的子案例库.

Step 2: 根据式 (3)~(6), 通过专家判断与密度集结为目标案例的属性进行赋权, 最后由式 (8) 进行归一化, 得到目标案例属性的权值向量.

Step 3: 根据式 (9)、(11) 和 (12) 分别计算目标案例与历史案例的数值型和符号型属性的相似度, 然后由式 (13) 得到目标案例与历史案例的相似度.

Step 4: 对目标案例与历史案例的相似度进行排序, 从而得到应急决策方案的推荐结果.

3 算例分析

随着我国经济社会的发展, 各省域、市域路网规模逐渐扩大, 密度不断提高. 与此同时, 道路交通突发事件的发生也随之增多. 道路交通中的突发事件因发生的时间、空间上的不确定性, 一旦发生极易引起交通拥挤、堵塞, 甚至导致交通事故, 不仅降低了本道路运输效率, 还会造成严重的次生交通堵塞问题, 因此路网突发事件对应急决策的时间要求很高. 某省

交通运输联网管理中心接到了路网突发事件的报警电话, 通过先遣人员的初步情报收集, 发现这是一起包含“雨雪”天气、公路类型为“高速”、事故车辆数量在“15”左右、死亡人数估计在“10”左右、无“堵车”的伴随“有危险品”的车祸突发事件. 其中“事故车辆数量”和“死亡人数”为连续型属性, “天气”、“公路类型”、“有无堵车”、“危险品”为符号型属性. 由此可得目标案例 P^C 的属性集合为: $S^C = \{s_1^C = \text{公路类型}, s_2^C = \text{天气类型}, s_3^C = \text{事故车辆数}, s_4^C = \text{死亡人数}, s_5^C = \text{有无堵车}, s_6^C = \text{危险品}\}$.

Step 1: 历史案例的筛选. 通过属性匹配, 从案例库中筛选出与目标案例比较相似的案例作为新的子案例库, 如表 2 所示.

Step 2: 通过专家判断与密度集结为目标案例的属性进行赋权. 在确定了目标案例包含的属性之后,

指挥中心邀 5 名专家 $E = \{e_1, e_2, e_3, e_4, e_5\}$ 对 S^C 中的属性进行主观判断, 如表 3 所示. 然后由式 (3)~(6) 计算可得各专家对目标案例的属性赋权, 如表 4 所示.

表 2 通过属性匹配形成的子案例库

案例	案例属性							
	公路类型	天气类型	事故车辆数量	死亡人数	堵车	危险品	受伤人数	火灾
P^C	高速	雨雪	15	9	无	有	—	—
P_1	普通	晴天	4	9	无	有	3	无
P_2	高速	雨雪	12	2	有	有	—	—
P_3	普通	大雾	4	3	有	有	5	无
P_4	高速	大雾	9	13	无	有	—	—
P_5	普通	雨雪	1	17	无	无	—	—
P_6	高速	大雾	2	0	有	—	—	—
P_7	普通	晴天	3	5	无	—	2	—
P_8	高速	晴天	1	—	—	有	1	—
P_9	高速	雨雪	7	—	—	—	—	无
P_{10}	普通	—	2	1	—	—	3	有

表 3 专家对目标案例属性重要性的排序及判断

专家	序关系	重要性程度
e_1	$s_3^C \succ s_4^C \succ s_6^C \succ s_1^C \succ s_5^C \succ s_2^C$	$s_3^C/s_4^C = 1, s_4^C/s_6^C = 1, s_6^C/s_1^C = 1.2, s_1^C/s_5^C = 1.2, s_5^C/s_2^C = 1.4$
e_2	$s_6^C \succ s_4^C \succ s_3^C \succ s_1^C \succ s_5^C \succ s_2^C$	$s_6^C/s_4^C = 1.6, s_4^C/s_3^C = 1, s_3^C/s_1^C = 1.2, s_1^C/s_5^C = 1, s_5^C/s_2^C = 1$
e_3	$s_4^C \succ s_1^C \succ s_6^C \succ s_2^C \succ s_3^C \succ s_5^C$	$s_4^C/s_1^C = 1.4, s_1^C/s_6^C = 1, s_6^C/s_2^C = 1.4, s_2^C/s_3^C = 1.2, s_3^C/s_5^C = 1$
e_4	$s_6^C \succ s_4^C \succ s_3^C \succ s_1^C \succ s_2^C \succ s_5^C$	$s_6^C/s_4^C = 1, s_4^C/s_3^C = 1.6, s_3^C/s_1^C = 1.2, s_1^C/s_2^C = 1, s_5^C/s_2^C = 1$
e_5	$s_3^C \succ s_4^C \succ s_6^C \succ s_1^C \succ s_5^C \succ s_2^C$	$s_3^C/s_4^C = 1, s_4^C/s_6^C = 1.2, s_6^C/s_1^C = 1.2, s_1^C/s_5^C = 1.2, s_5^C/s_2^C = 1$

表 4 各个专家对目标案例属性的赋权

专家	公路类型	天气类型	事故车辆数量	死亡人数	堵车	危险品
e_1	0.167	0.098	0.199	0.199	0.138	0.199
e_2	0.137	0.137	0.164	0.164	0.137	0.262
e_3	0.189	0.135	0.112	0.264	0.112	0.189
e_4	0.124	0.124	0.149	0.239	0.124	0.239
e_5	0.148	0.124	0.213	0.213	0.124	0.178

通过高斯密度函数对每一个属性可能取的值进行密度集结后, 得到每一个属性的概率密度曲线, 如图 1 所示. 取每一个属性的权值在该属性的概率密度取最大时所对应的值, 由式 (8) 归一化后得到目标案例属性的权值向量为 $w^C = (0.1457, 0.1296, 0.1711, 0.2182, 0.1277, 0.1913)$.

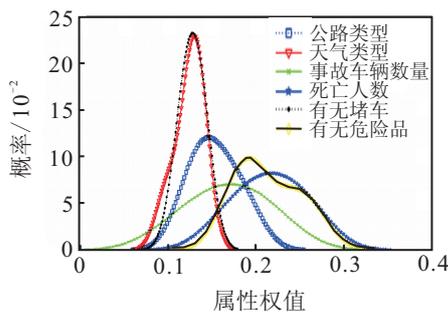


图 1 目标案例属性权值的概率密度曲线

Step 3: 相似度计算. 由式 (9)~(13) 求得目标案例与各历史案例的相似度值为

$$\text{Sim}(P^C, P_1) = 0.519, \text{Sim}(P^C, P_2) = 0.541,$$

$$\text{Sim}(P^C, P_3) = 0.315, \text{Sim}(P^C, P_4) = 0.511,$$

$$\text{Sim}(P^C, P_5) = 0.385, \text{Sim}(P^C, P_6) = 0.272,$$

$$\text{Sim}(P^C, P_7) = 0.297, \text{Sim}(P^C, P_8) = 0.285,$$

$$\text{Sim}(P^C, P_9) = 0.355, \text{Sim}(P^C, P_{10}) = 0.139.$$

从而根据相似度可以对历史案例进行排序, 因此历史案例决策方案的推荐结果为 $A_2 > A_1 > A_4 > A_5 > A_9 > A_3 > A_7 > A_8 > A_6 > A_{10}$. 可将上述历史案例对应的决策方案作为生成当前应急决策方案的参考集, 提交给相应的决策者以快速生成当前决策方案.

4 结 论

鉴于突发事件的独特性质, 本文提出了目标案例与历史案例的非一致属性集合问题, 属性的非完全匹配不能直接采用传统的相似度算法计算案例间的相似度. 此外, 为了提高应急决策方案生成的速度, 本文提出了匹配度这一概念, 基于匹配度提取比较相似的历史案例组成子案例库; 然后通过密度函数方法集结专家组的权重判断, 并给出最终的目标案例的属性权重; 最后设计目标案例与历史案例的匹配属性集合的数值型属性与符号型属性的相似度计算方法, 从而得到应急决策方案的推荐结果. 本文在应急决策方案的生成流程中做了以下工作:

1) 针对所提出的目标案例与历史案例的非一致属性集合问题, 基于匹配度提取匹配度较大的历史案

例组成新的子案例库,减少了案例相似度计算量,提高了应急决策方案的生成速度。

2) 通过 G1 法得到专家组对目标案例属性的权重判断后,采用密度函数得到目标案例各属性的权重概率分布曲线,由分布曲线得到最终的属性权重分配,不仅体现了应急决策管理人员对属性的主观判断,而且可以有效避免个别误差较大的权重对最后集结权重的影响。

3) 在计算案例间相似度时,本文针对数值型属性和符号型属性的自有特征,分别提出不同的相似度计算方法,尤其在符号型属性相似度计算方法中,考虑了符号型属性在历史案例库中的出现频率,不是直接赋值为 0 或 1,避免了符号型属性相似度的极化现象。

今后的研究工作将针对历史事件的方案实施效果,通过专家群决策来进行应急方案生成的研究。

参考文献(References)

- [1] 中华人民共和国突发事件应对法[EB/OL]. [2015-02-21]. <http://www.gov.cn/flfg/2007-08/30/content732593.htm>.
(The emergency response law of the People's Republic of China[EB/OL]. [2015-02-21]. <http://www.gov.cn/flfg/2007-08/30/content732593.htm>.)
- [2] 鄂文帅, 寇纲, 彭怡, 等. 面向突发事件的模糊多目标应急决策方法[J]. 系统工程理论与实践, 2012, 36(6): 1298-1304.
(Wu W S, Kou G, Peng Y, et al. A fuzzy multi-criteria emergency decision-making method[J]. Systems Engineering — Theory & Practice, 2012, 36(6): 1298-1304.)
- [3] Kou G, Ergu D, Shi Y. An integrated expert system for fast disaster assessment[J]. Computers & Operations Research, 2014, 42: 95-107.
- [4] Liu Y, Fan Z, Yuan Y, et al. A FTA-based method for risk decision-making in emergency response[J]. Computers & Operations Research, 2014, 42: 49-57.
- [5] 李永海, 樊治平, 袁媛. 考虑应急方案实施效果的突发事件应急方案生成方法[J]. 控制与决策, 2014, 29(2): 275-280.
(Li Y H, Fan Z P, Yuan Y. Method for generating emergency alternative with considering implementation effects of emergency alternatives[J]. Control and Decision, 2014, 29(2): 275-280.)
- [6] 王清, 赵勇, 饶从军. 基于阈值的案例决策方法及其在创新设计中的应用[J]. 控制与决策, 2010, 25(10): 1562-1566.
(Wang Q, Zhao Y, Rao C J. Case-based decision method based on threshold and its application in innovative design[J]. Control and Decision, 2010, 25(10): 1562-1566.)
- [7] Aamodt A, Plaza E. Case-based reasoning: Foundational issues, methodological variations, and system approaches[J]. AI Communications, 1994, 7(1): 39-59.
- [8] 姜艳萍, 樊治平, 袁娜. 基于相似度计算的森林火灾应急响应方案选择方法[J]. 系统工程, 2010, 28(11): 104-109.
(Jiang Y P, Fan Z P, Yuan N. A method for selecting forest fire emergency response alternatives[J]. Systems Engineering, 2010, 28(11): 104-109.)
- [9] 李永海, 樊治平, 李铭洋. 解决广义不确定型决策问题的案例决策方法[J]. 系统工程学报, 2014, 29(1): 21-29.
(Li Y H, Fan Z P, Li M Y. Case-based decision analysis method for general uncertain decision making problem[J]. J of Systems Engineering, 2014, 29(1): 21-29.)
- [10] Zhang Dong-song, Zhou Li-na, Jr Nunamaker J F. A knowledge management framework for the support of decision making in humanitarian assistance/disaster relief[J]. Knowledge and Information Systems, 2002, 4(3): 370-385.
- [11] Fan Z, Li Y, Zhang Y. Generating project risk response strategies based on CBR: A case study[J]. Expert Systems with Applications, 2015, 42(6): 2870-2883.
- [12] Hall P, Wand M. On the accuracy of binned kernel density estimators[J]. J of Multivariate Analysis, 1994, 56(2): 165-184.
- [13] 王学军, 郭亚军. 基于 G1 法的判断矩阵的一致性分析[J]. 中国管理科学, 2006, 14(3): 65-70.
(Wang X J, Guo Y J. Analyzing the consistency of comparison matrix based on G1 method[J]. Chinese J of Management Science, 2006, 14(3): 65-70.)
- [14] 李存华, 孙志挥, 陈耿, 等. 核密度估计及其在聚类算法构造中的应用[J]. 计算机研究与发展, 2004, 41(10): 1712-1718.
(Li C H, Sun Z H, Chen G, et al. Kernel density estimation and its application to clustering algorithm construction[J]. J of Computer Research and Development, 2004, 41(10): 1712-1718.)
- [15] Fan Z, Li Y, Wang X, et al. Hybrid similarity measure for case retrieval in CBR and its application to emergency response towards gas explosion[J]. Expert Systems with Applications, 2014, 41(5): 2528-2529.
- [16] 梁昌勇, 顾东晓, 程文娟, 等. 含非连续性信息多属性案例中的决策知识发现方法[J]. 中国管理科学, 2014, 22(4): 83-84.
(Liang C Y, Gu D X, Cheng W J, et al. A decision knowledge discovery method for multi-attribute cases with non-continuous features[J]. Chinese J of Management Science, 2014, 22(4): 83-84.)

(责任编辑: 孙艺红)