

文章编号:1674-2869(2018)03-0301-05

基于案例推理的地铁施工事故案例库的建立与评价

王勇,赖芟宇*,易紫妮,李晓娟

福建农林大学交通与土木工程学院,福建福州 350108

摘要:通过网络搜集、文献查阅、单位调查等方式,对我国近15年来地铁施工事故进行统计、整理,分析事故发生的主要因素及发展趋势。将施工事故概念化和语义化,并基于各层次概念建立地铁施工事故概念树,计算其嫡系相似度、概念相似度、属性相似度以求得不同案例之间的相似度,从而实现基于案例推理的事故案例库的建立。它的建立可以为不同单位依据自身施工环境特点,检索出相似度较高的事故案例。为地铁施工安全操作的制定、事故应急措施等都具有重要作用。

关键词:案例推理;统计;地铁施工;案例库;相似度

中图分类号:TU714 文献标识码:A doi:10.3969/j.issn.1674-2869.2018.03.014

Establishment and Evaluation of Database of Subway Construction Accidents Based on Case-Based Reasoning

WANG Yong, LAI Jiyu*, YI Zini, LI Xiaojuan

School of Transportation and Civil Engineering, Fujian Agriculture and Forestry University, Fuzhou 350108, China

Abstract: The present work analyzed the subway construction accidents in China for nearly 15 years through network collection, literature review, survey and so on, finding out the main factors and the trend of these accidents. By conceptualization and semanticization of the construction accidents, the concept tree of subway construction accidents was established. Then the similarities of different accidents were calculated with their concept and property, building the accidents case database based on case-based reasoning. This database is helpful for different companies to retrieve some highly similar accident cases with their own cases based on their construction environment, which plays a significant role in the formulation of safe operation and measures towards accident emergency during subway construction.

Keywords: case-based reasoning; statistics; subway construction; case database; similarity

随着我国城市经济飞速发展,尤其是20世纪90年代以来,城市基础设施的建设与完善,城区建设规模不断扩大,吸引着来自各地的管理、技术、劳务等各种人才,使得人口总量逐年增加。截止2017年,我国城区人口超过500万的城市共11座^[1-3]。面对如此大的人口规模,传统的交通工具难以负担如此大的交通流量,而且过多的

汽车尾气排放,严重影响城市环境^[4]。因此城市轨道交通系统开始修建,并逐渐完善,以有效、便捷地缓解目前城市公共交通运输压力。

图1为我国近15年来发生的323起地铁施工事故统计^[5-7],在2003-2017年间,我国的地铁施工事故总量呈现出两段增长期。2005-2010年呈一快速增长期,2010年达到历年最高,共38起,死亡

收稿日期:2017-12-07

基金项目:国家社会科学基金(13BGL150);福建农林大学科技创新专项基金(KFA17256A)

作者简介:王勇,硕士研究生。E-mail:1030944050@qq.com

*通讯作者:赖芟宇,博士,教授。E-mail:66428267@qq.com

引文格式:王勇,赖芟宇,易紫妮,等.基于案例推理的地铁施工事故案例库的建立与评价[J].武汉工程大学学报,2018,40(3):301-305.

人数达 27; 2008 年死亡人数为历年最高, 为 36。从 2011 年至今, 又处于一段增长期, 且平均每年事故数量超过 21 人, 平均每年事故死亡人数超过 19 人; 预计未来几年内, 施工事故总量依旧会呈现增长趋势。通过数据分析其原因: 前期由于我国部分城市首次修建地铁, 施工技术与经验不足, 且相关法规与标准不完善; 后期主要是一二线城市的不断发展, 促使获批修建地铁城市大量增加, 造成事故不断增多。

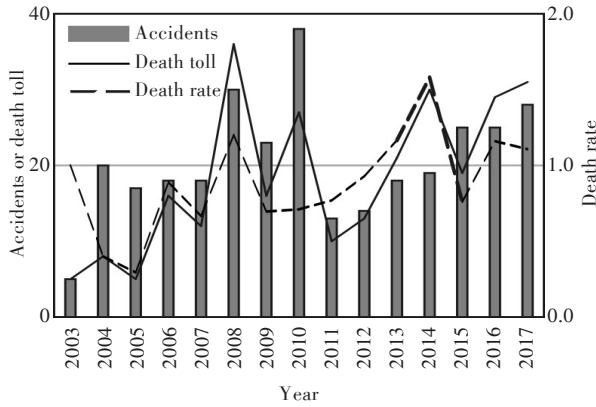


图 1 2003-2017 年我国地铁施工事故统计

Fig. 1 Statistics of subway construction accidents in China from 2003 to 2017

根据海因里希法则, 在一起死亡或重伤的事故后, 隐藏着 29 起轻伤事故以及 300 起无伤亡事故, 甚至无数的不安全的操作行为, 即三者之比为 1:29:300^[8]。在本次统计的 248 起具有较为详细记录的事故中, 三者之比为 161:11:76, 这表明仍有大量的事故数据未能统计到。基于此, 建立全面的地铁施工事故信息搜集系统, 运用基于案例推理(case-based reasoning, CBR) 机制的方法构建地铁施工事故案例库, 检索的相似案例出现时, 不仅可以为施工组织方案的制定提供参考, 增加其安全性, 还可以基于施工过程中出现的事故, 预测施工高危险系数的环节容易发生的安全事故, 以降低其发生的可能性。

1 基于 CBR 的地铁施工案例库的构建

CBR 作为三种主要的推理机制之一, 起源于 1977 年, 发展历程已近 40 年^[9]。CBR 机制机理主要是对过去事故建立案例库, 当遇到新问题时, 可以通过相似案例搜索, 力求找出与以往最相似的事故, 解析其发生的过程与解决方案, 为新问题提供有效的参考^[10], 因而是一种利用以往相似经验去推理解决新方法^[11]。若从案例库中选取

的实例相似度低, 则可以整理本次案例, 将其作为新案例输入库中, 作为以后的参考样本。这种方法广泛应用于计算机工程、分类与诊断、交通运营等各种行业^[12-13]。

自 1965 年北京地铁一号线开工以来, 我国已经开始进入现代城市交通时代。特别是在 20 世纪 90 年代, 我国地铁修建开启了第一个高潮, 但在地铁施工过程中, 由于其工程量大、施工复杂难度高, 安全管理不到位等原因^[14], 产生大量的施工事故案例。若将 CBR 方法运用到地铁施工事故案例库的建立中, 建设方从中提取有效信息选项, 案例库系统可以从中导出相对应的往年事故, 并对此相关的案例进行分析评价。而地铁施工单位作为案例库的主要参与者, 可以将施工案例搜集并整理, 导入到案例库中, 以此扩充数量, 使得以后选取与导出的案例更充分、评价信息更加完善。当施工单位在制定施工组织时, 可以先依据工程标段的各类信息, 如事故发生的时间点、事故发生段位及地理位置、施工过程中的施工方法和人员素质情况、事故结果及处理结果等, 从案例库中调取相似案例的所有信息, 既可以在施工前做好应急预案, 也可以在事故发生后采用相似且比较成功的处理措施, 其机制见图 2。若在案例库中检索不到相似度较高的案例, 则可以调整事故的信息类型, 然后将其整理好, 作为新案例重新输入到案例库系统中, 为以后的事故处置提供经验。

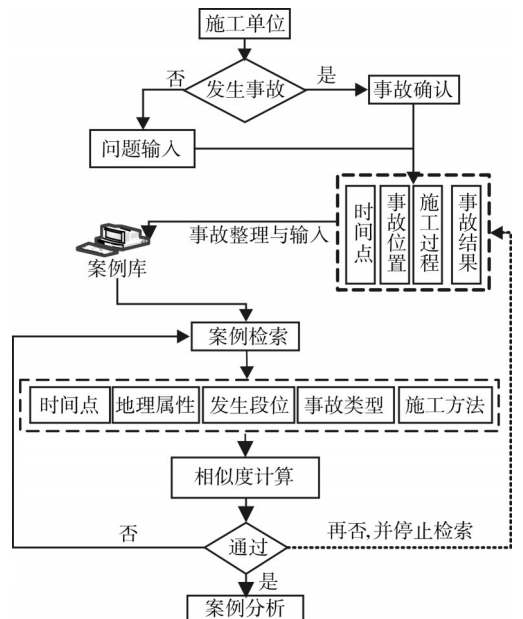


图 2 基于 CBR 的地铁施工事故案例库的运行机制

Fig. 2 Operation mechanism of CBR-based casework database for subway construction accidents

2 相似度的计算

每一地铁施工事故的发生过程可以由许多基本词语描述,因此可以将案例包含的信息概念化,而不同深度的概念以及同一深度的不同概念的相互交织,可以形成概念树^[15],并以此做为案例库建立的数据模型。在一概念树中,逻辑关系包括继承关系、组成关系、扩展关系。在一案例中,一个概念可以继承和扩展出更多的小概念。地铁施工事故扩展出不同概念,如发生地点包括特大城市、一二线城市,事故发生类型扩展出坍塌、高处坠落等。继而以坍塌为例,扩展出其他概念,形成地铁施工坍塌事故的完整体系,最后扩展出最底层的概念,如土质疏松、支护体系破坏、导洞影响等,如图3所示。在每一次的扩展中,概念树形成具有层次性的概念。

2.1 数相似度的概念及其特征

在概念树中,每个概念所处的层次称为概念层次,用 l 表示, $\max l$ 称为树的深度,记为 m 。此

$$S_t(c_1, c_2) = \begin{cases} \frac{a[l(c_1) + l(c_2)]}{[d(c_1, c_2) + a] \times 2 \times m \times \max(|l(c_1)| - |l(c_2)|, 1)} & a > 0, c_1 \neq c_2 \\ 1 & c_1 = c_2 \end{cases} \quad (1)$$

其中设 $l(c_1)$ 和 $l(c_2)$ 分别 c_1 和 c_2 的概念层次; a 为正数,表示调节参数; $S_t(c_1, c_2) = 1$ 表示概念 c_1 和 c_2 具有等价关系。从公式中可以发现,由于概念树深度与调节参数均是固定值,概念树的相似度主要受概念的概念层次和概念距离影响。若概念层次深度越大,概念距离路径越少,则两个概念之间的概念相似度则越大。

在图4中,概念树共有3层, $m=3$;发生地点和坍塌的概念距离为3,深度分别为1和2;发生地点和坍塌的概念相似度计算得:

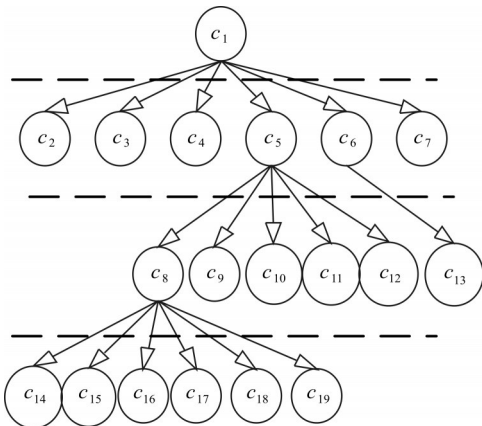


图4 概念树示意图

Fig. 4 Schematic diagram of concept tree

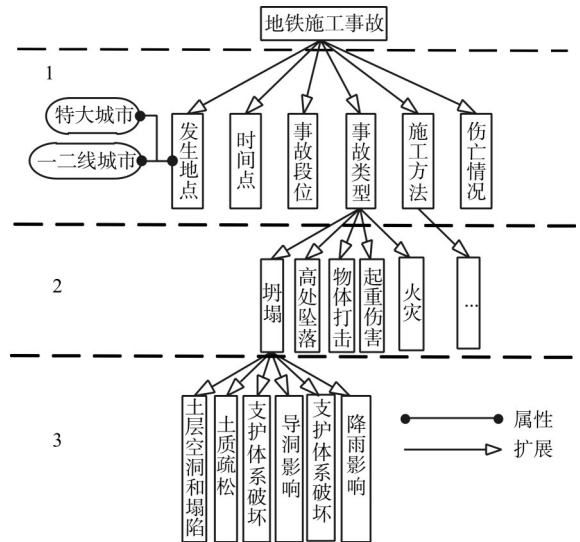


图3 地铁施工坍塌事故概念树示意图

Fig. 3 Conceptual tree diagram of subway construction collapse accident

外,概念之间的最短路径条数之和,称为概念距离,用 d 表示,则 c_1 和 c_2 的树相似度可以表示为 $S_t(c_1, c_2)$,其计算公式为:

$$S_t(c_2, c_8) = \frac{a(1+2)}{(3+a) \times 2 \times 3 \times 1} = \frac{a}{2(3+a)}$$

2.2 嫡系相似度的计算

在概念树中,前一层概念称为父概念,它表示该层引出的概念之间的相似度,称为嫡系相似度,记为 $S_s()$,其表达方式可以用公式(1)计算。

$$S_s(c_2, c_3) = 1, S_s(c_8, c_{13}) = S_s(c_5, c_6) = \frac{a \times (2+2)}{(2+a) \times 2 \times 3 \times 1} = \frac{2a}{3(2+a)}$$

2.3 概念相似度的计算

在概念树中,若 c_1, c_2 为两个不等价的概念,则其概念相似度可以表示为 $S_c()$,其计算公式可以表达为:

$$S_c(c_1, c_2) = \frac{S_t(c_1, c_2) + S_s(c_1, c_2)}{2} \quad (2)$$

2.4 属性相似度的计算

概念的属性也表达着概念的某些特征,因而概念的属性相似度也影响着案例相似度,记属性相似度为 $S_b()$,其计算方法与公式(1)类似。

3 相似度的计算过程

在评价两个案例的相似度时,不同的概念,受到多种因素的影响,因此还应根据其权重来计

算。设案例库中检索的案例 X 和输入系统的事故 Y , 案例 X 包含了 n 个概念 $\{c_{11}, c_{12}, c_{13}, \dots, c_{1n}\}$, r 个数值类型属性 $\{b_{11}, b_{12}, b_{13}, \dots, b_{1r}\}$ 和 h 个对象类型属性 $\{b_{21}, b_{22}, b_{23}, \dots, b_{2h}\}$; 案例 Y 包括 m 个概念 $\{c_{21}, c_{22}, c_{23}, \dots, c_{2m}\}$, k 个数值类型属性 $\{b_{31}, b_{32}, b_{33}, \dots, b_{3k}\}$ 和 j 个对象类型属性 $\{b_{41}, b_{42}, b_{43}, \dots, b_{4j}\}$ 。

① 求概念相似度矩阵 S :
$$S = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{m1} & \lambda_{m2} & \dots & \lambda_{mn} \end{bmatrix}$$

其中 $\lambda_{mn} = S_c(c_{1n}, c_{2m}) = \frac{S_s(c_{1n}, c_{2m}) + S_t(c_{1n}, c_{2m})}{2}$ 。

② 从矩阵 S 中选取 $\max \lambda_{mn}$, 将其所在的行与列的值, 形成新的矩阵 S' , 如此不断循环, 到该矩阵为空的时候停止。然后将每次计算得到的 $\max \lambda_{mn}$ 记为 $w = \{f_1, f_2, f_3, \dots, f_{l_0}\}$, 式中 $l_0 = \min(m, n)$ 。

③ 进而再求 $w = \{f_1, f_2, f_3, \dots, f_{l_0}\}$ 的加权平均
$$S_{cs}(X, Y) = \sum_{i=1}^{l_0} \mu_i f_i$$
, μ_i 表示概念 i 的权重值。

④ 依据数值类型和对象类型两种类别, 将概念的属性进行分类, 然后分别构造出两种类型的相似度矩阵 M 和 N :

$$M = \begin{bmatrix} \omega_{11} & \omega_{12} & \dots & \omega_{1r} \\ \omega_{21} & \omega_{22} & \dots & \omega_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{k1} & \omega_{k2} & \dots & \omega_{kr} \end{bmatrix}, \quad N = \begin{bmatrix} \rho_{11} & \rho_{12} & \dots & \rho_{1h} \\ \rho_{21} & \rho_{22} & \dots & \rho_{2h} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{j1} & \rho_{j2} & \dots & \rho_{jh} \end{bmatrix}$$

其中, 数值类型矩阵中的两属性之间的间接相似度 $w_{kr} = S_b(b_{1r}, b_{3k})$, 而对象类型之间的属性的间接相似度可以表示为 $\rho_{hj} = S_c(b_{2l}, b_{4h})$, 可由公式(2)计算得到。

⑤ 同 ① 的计算过程方法相似, 从矩阵 M 中选取 $\max \omega_{kr}$, 并清除其所在行和列中的值, 然后形成新的矩阵 M' , 不断循环计算, 直到 M' 矩阵为空的时候停止。每取得 $\max \omega_{kr}$ 时候, 记 $z = \{x_1, x_2, x_3, \dots, x_{t_1}\}$, 式中 $t_1 = \min(k, r)$ 。同理可知, 在计算矩阵 N 中, 可以得到 $v = \{y_1, y_2, y_3, \dots, y_{t_2}\}$, 同样 $t_2 = \min(h, j)$ 。

⑥ 计算 $z = \{x_1, x_2, x_3, \dots, x_{t_1}\}$ 和 $v = \{y_1, y_2, y_3, \dots, y_{t_2}\}$ 各自的加权平均值:

$$S_{ps}(X, Y) = \sum_{i=1}^{t_1} \phi_i x_i$$
, 式中 ϕ_i 是其权重值;

$$S_{po}(X, Y) = \sum_{i=1}^{t_2} \varepsilon_i y_i$$
, 式中 ε_i 为其权重值。

⑦ 将上面两种加权平均值结合起来, 可以得

到概念属性相似度, 表示为 $S_b(X, Y)$, 其公式表达为:

$$S_b(X, Y) = \delta \times S_{ps}(X, Y) + (1 - \delta) \times S_{po}(X, Y),$$

$0 < \delta < 1$ (δ 表示其相似度的权重系数)

⑧ 结合概念相似度与概念属性相似度, 按照各自权重, 计算出案例的相似度, 可以表示为 $S(X, Y)$, 即:

$$S(X, Y) = \beta \times S_{cs}(X, Y) + (1 - \beta) \times S_b(X, Y),$$

$0 < \beta < 1$ (β 表示其相似度的权重系数)

最终根据设置的相似度阈值, 判断该案例的相似度是否符合要求, 是否在系统中通过。

若 $\frac{a}{1+a} \leq S(X, Y) < \frac{2a+1}{2(1+a)}$, (其中 a 为调节参数) 则表示 X 和 Y 两个案例具有较高的相似度;

若 $\frac{2a+1}{2(1+a)} \leq S(X, Y) < 1$, 则表示 X 和 Y 两个案例具有极高的相似度;

若 $S(X, Y) = 1$, 则表示 X 和 Y 案例完全一致, 属于等价关系。

如图4中概念树中的第二层发生地点与伤亡情况的嫡系相似度 $S_c(c_2, c_7) = 1$, 表明两个概念的父概念相同。

4 案例相似度计算的应用

以2起地铁施工坍塌事故为例, 通过计算两案例之间的相似度, 确定两者之间是否有参考价值。

案例 X : 2014年12月17日17时20分, 南京地铁四号线TA06标段在安排1号竖井内钢筋绑扎时, 未按照要求, 导致钢筋笼架在施工过程中发生坍塌, 造成4死3伤的施工事故。

设计案例 Y : 2013年8月18日21点左右, 某市郊区在建的轨道交通, 某区间工地发生坍塌, 造成2死4伤的施工事故。

从这2个案例中, 在第二层中选取发生地点、发生时间、伤亡情况、事故类型4个概念, 进而扩展出郊区与车站、上下午、事故级别、坍塌等概念, 所有的概念及其属性一起形成地铁施工事故的概念树, 如图5所示。在这个案例设计中, 设取 $a = 1, \delta = 0.5, \beta = 0.5$, 案例之间相似度计算过程如下:

由 ① 可计算案例 X, Y 的概念相似度矩阵 S_1 :

$$S_1 = \begin{bmatrix} 0.56 & 0.13 & 0.13 & 0.13 & 0.13 & 0.12 & 0.12 & 0.12 & 0.12 \\ 0.19 & 0.33 & 0.33 & 0.33 & 0.33 & 0.67 & 0.21 & 0.21 & 0.21 \\ 0.19 & 0.33 & 0.33 & 0.33 & 0.33 & 0.67 & 0.21 & 0.21 & 0.21 \\ 0.19 & 0.33 & 0.33 & 0.33 & 0.33 & 0.21 & 0.67 & 0.21 & 0.21 \\ 0.19 & 0.33 & 0.33 & 0.33 & 0.33 & 0.21 & 0.21 & 0.67 & 0.21 \\ 0.19 & 0.33 & 0.33 & 0.33 & 0.33 & 0.21 & 0.21 & 0.21 & 0.67 \end{bmatrix}$$

然后取概念的权重 $\mu_i (i = 1, 2, \dots, 6) = 0.2$, 由前

面计算过程②和③,可得 $S_{cs}(X, Y) = 0.714$ 。再根据计算过程④,可分别得到案例 X 和案例 Y 之间的数值类属性的相似度矩阵 M_{11} ,和对象类属性相似度矩阵 N_{12} :

$$M_{11} = \begin{bmatrix} 0.33 & 0.20 \\ 0.20 & 0.33 \end{bmatrix}, \quad N_{12} = (0.33)$$

然后选取概念的属性权重 $\phi_1 = \phi_2 = 0.5$, $\varepsilon_1 = 1$,再按照前面的计算方式⑤至⑦,可以计算出数值型和对象型的各自加权平均:

$$S_{ps}(X, Y) = 0.33 \times 0.5 + 0.33 \times 0.5 = 0.33,$$

$$S_{po}(X, Y) = 0.33,$$

$$S_p(X, Y) = \delta \times S_{ps}(X, Y) + (1 - \delta) \times S_{po}(X, Y) = 0.353。$$

最后依据计算步骤⑧,可以计算出两个案例的相似度: $S(X, Y) = \beta \times S_{cs}(X, Y) + (1 - \beta) \times S_p(X, Y) = 0.534$

由于计算结果显示: $\frac{a}{1+a} < S(X, Y) < \frac{2a+1}{2(a+1)}$, 这表明两个案例具有较高的相似度,南京地铁施工事故案例可以为设计案例提供较高的参考价值。若在案例库检索中,与之相似度有更高值,应按照从大到小依次排列。此外,由于我国幅员辽阔,不同地区的地理环境不同,施工场地条件、人员素质、技术设备等因素也有所差异,使得计算的相似度结果会有所偏差。因此,在案例库检索时,概念树的第一层深度应建立得更加精确、详细,以提高相似度的准确性。

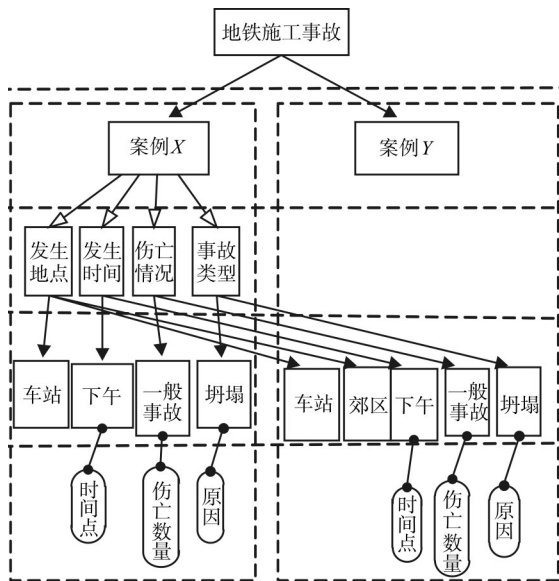


图5 地铁施工事故概念结构图

Fig. 5 Conceptual structure of subway construction accident

5 结 语

案例库的建立和完善,可以使不同单位、不同地区的地铁施工事故案例相互联系,降低传统施

工案例之间较强的独立性。若能将每个地铁施工事故案例输入案例库中,通过计算机计算其相似度,可以有效地将它们结合起来,提高案例的参考价值,相互借鉴而不是仅仅将它存在于档案库中。

通过引入概念树,特别是概念和属性同时考虑在计算中,具有一定的合理性及应用意义,可以快速判断事故之间的相似性。若相似度较高,则可以将其应用到本施工组织规划中,避免事故的发生。但是目前案例库的建立,还存在许多困难,这需要政府与相关企业进行合作,不断充实案例库的总量,发挥其借鉴与指导作用。

参考文献:

- [1] 2017中国城市人口排名名单13个城超1000万[EB/OL]. (2017-8-11)[2017-12-7]. <http://mnw.cn/news/china/1813680.html>.
- [2] 李晓娟. 地铁运营安全风险分析[J]. 工程管理学报, 2017, 31(1): 83-88.
- [3] 李志东. 地铁运营安全的风险评估与管理[J]. 价值工程, 2017(23): 54-56.
- [4] 万国胜, 梁晖, 周华杰. 地铁运行引起环境振动的评价方法[J]. 武汉工程大学学报, 2009, 31(5): 25-28.
- [5] 邓小鹏, 李启明, 周志鹏. 地铁施工安全事故规律性的统计分析[J]. 统计与决策, 2010(9): 87-89.
- [6] 李凤伟, 杜修力, 张明聚. 地铁工程建设施工事故统计分析[J]. 地下空间与工程学报, 2014(2): 474-479.
- [7] 李皓然, 李启明, 陆莹. 2002-2016年我国地铁施工安全事故规律性统计分析[J]. 都市快轨交通, 2017, 30(1): 12-19.
- [8] 谢东升, 钱七虎, 戎晓力. 地铁工程建设安全风险研究[J]. 土木工程与管理学报, 2012, 29(1): 61-67.
- [9] 李晓辉, 刘妍秀. 基于实例推理机制(CBR)综述[J]. 长春大学学报, 2006, 16(4): 68-70.
- [10] 孙锋. CBR中的相似算法[J]. 重庆工学院学报, 2008, 22(4): 67-72.
- [11] 严爱军, 钱丽敏, 王普. 案例推理属性权重的分配模型比较研究[J]. 自动化学报, 2014, 40(9): 1896-1902.
- [12] 陆莹, 李启明, 高原. 基于案例推理的地铁运营事故案例库的构建[J]. 东南大学学报, 2015, 45(5): 990-994.
- [13] 张春晓, 严爱军, 王普. 一种改进的案例推理分类方法研究[J]. 自动化学报, 2014, 40(9): 2015-2021.
- [14] 丁烈云, 吴贤国, 骆汉宾, 等. 地铁工程施工安全评价标准研究[J]. 土木工程学报, 2011, 44(11): 121-127.
- [15] 向东, 赵勇, 陈阳. 面向语义信息案例知识表达与相似度计算方法研究[J]. 计算机工程与科学, 2011, 33(12): 159-166.