

# 基于粗糙集的隧道施工风险案例推理研究

胡 珉, 张链翔

(上海大学悉尼工商学院, 上海 200072)

**摘要:** 隧道工程具有复杂性和不确定性, 其施工过程存在很大风险。该文调研隧道工程历史案例, 从问题、情景描述、求解方法和结果 4 个方面进行案例库的设计和归纳。根据粗糙集理论, 利用差别矩阵实现对案例决策表特征值属性约简和属性权重的学习。与传统方法相比, 该方法有利于减少主观影响, 提高风险评估准确率和效率, 保证了隧道施工的安全性。

**关键词:** 粗糙集; 隧道施工; 风险; 案例推理

## Research of Case-based Reasoning for Tunnelling Risk Using Rough Set

HU Min, ZHANG Lian-xiang

(Sydney Institute of Language and Commerce, Shanghai University, Shanghai 200072)

**【Abstract】** As the complexity and uncertainty of the tunnel engineering, there is a considerable risk in the entire construction process. After investigating many history cases, this paper sums up these cases from four aspects, including question, scene, solution and result. To improve the efficiency and accuracy of case retrieving in case-based reasoning, the rough-set theory is introduced to reduce case attributes and get attribute weights. Compared with traditional technique, this technique reduces the subjective effects and improves risk assessment of the accuracy and efficiency, and guarantees the safety of the tunnel construction.

**【Key words】** rough set; tunnelling; risk; case-based reasoning

### 1 概述

隧道施工期间存在诸多不确定因素, 如何有效减少由此带来的工作量成为隧道工程急需解决的问题。文献[1]于2000年提出“隧道工程的建设过程就是全面的风险管理和风险分担的过程”, 如何对隧道工程进行风险分析被越来越多地关注。而隧道工程的复杂性导致风险分析时存在知识结构化程度不高的问题, 使规则推理或决策树推理等方法不能完全满足实际工程需求。案例推理的方法模仿人类解决问题和不断学习的方式, 充分利用过去的经验处理将要出现的新问题, 具有知识获取容易、符合专家思维过程、具有持续不断的学习能力等优点。由于隧道工程具有经验积淀较丰富的特点, 使得利用案例推理进行风险分析成为可能。

案例推理工作可以概括为如下 5 个步骤<sup>[2]</sup>:

(1) 案例描述, 对新问题以案例的形式进行描述;

(2) 案例检索, 计算案例库中案例与新问题的匹配度, 获得新问题的相似案例;

(3) 案例重用, 将相似案例的解决方案重用到新问题, 根据领域知识和目标约束条件, 修改相似案例的解决方案, 使之适用于新问题, 作为新问题的建议解输出;

(4) 校验修改, 对新问题的解进行检验或验证, 将验证后的解作为最终解输出;

(5) 案例学习, 保留获得最终解的新问题, 将其存入案例库中, 实现案例知识的学习。

案例推理结构如图 1 所示, 其中 3 个关键点如下:

(1) 案例库。是 CBR 系统中的主要知识库, 案例库的设计是整个案例推理的基础。

(2) 案例检索环节。是案例推理的关键步骤, 最常用的案例检索策略是最近相邻策略或  $k$  近邻算法。如何选择有效的特征属性, 并为每个特征属性赋予一个合理权重, 客观地反映案例和待解问题的相似度, 以提高检索的准确性, 是问题的关键。

(3) 案例推理中的学习功能。即不断往案例库中增加新的案例。以案例推理为基础的系统需要保持并管理一组数量较大的案例, 为了限制案例库的无限膨胀, 需要合理的办法来删除冗余案例, 限制案例库的大小。

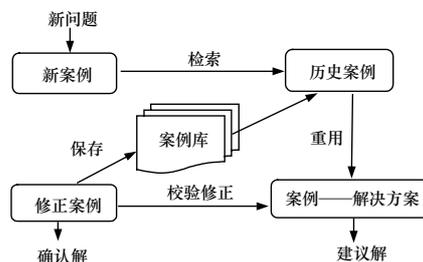


图 1 案例推理结构

### 2 案例描述

案例是一段带有上下文信息知识, 该知识表达了推理机在达到其目标的过程中能起关键作用的经验。案例一般由 3 个主要部分组成: 问题或情景描述, 解决方案和结果。因

**作者简介:** 胡珉(1970—), 女, 副教授、博士, 主研方向: 人工智能, 地下工程信息处理; 张链翔, 学士

**收稿日期:** 2008-03-08 **E-mail:** minahu@163.com

此, 对一个隧道施工风险案例描述如下:

(1)案例类别。为了有效组织案例库中的案例, 本文按隧道掘进设备、隧道的功用、施工阶段等进行类别划分, 可以方便索引和检索。

(2)案例属性。案例属性分为条件属性和辅助属性, 其中, 条件属性是在推理过程中起重要作用、与决策结果有因果关系的属性; 辅助属性与决策结果无明显因果关系, 仅用来更全面地对案例进行描述。由于隧道工程涉及面广、风险相关属性众多, 因此本文利用粗糙集方法进行分类, 确定哪些是主要属性、哪些是辅助属性。

(3)解决方案。指风险的分类结果和案例所对应的方法及措施。

(4)结论。该案例采用上述决策结果后, 产生的结果情形的描述。

### 3 属性选择

#### 3.1 属性初始化

粗糙集理论处理具有不完整、不一致及不确定性特征的信息时有特别的优势。基于粗糙集理论的属性约简是一种高效数据挖掘算法, 该算法利用布尔代数式的化简结果, 得到所需约简和规则, 约简后的属性集应与原属性集具有同样的分辨能力<sup>[3]</sup>。

根据粗糙集理论, 如果一个信息系统可以表示为

$$S: (U, A, V, f)$$

其中,  $U$  是一个有限对象集;  $A$  是有限属性集, 分为条件属性集  $C$  和决策属性集  $D$ , 即  $A=C \cup D$ ,  $C \cap D = \emptyset$ ;  $V$  是属性值组成的集合;  $f$  是一个  $U \times A \rightarrow f$  信息函数, 它指定  $U$  中每个对象的属性值。

则假设属性子集  $R \subseteq C$ , 若

(1)对于  $D$  有  $POS_R(D) = POS_C(D)$ , 其中,  $POS_C(D)$  表示  $C$  正区域<sup>[4]</sup>, 即对象集  $U$  的所有条件类。

(2)不存在任何子集  $R_1 \subseteq R$ , 使得  $POS_{R_1}(D) = POS_C(D)$ 。

则称  $R$  为  $C$  的一个约简。

为了方便属性约简, 对案例中相关数据进行离散化处理, 如表 1 所示。

表 1 数据离散化处理结果示例

参数	盾构 类型(T)	隧道 功用(U)	地质 状况(S)	风险所处工序(P)				风险 特征(C)
				出洞	推进	进洞	旁通道	
1	单圆土 压平衡	轨道 交通	粘土	旋喷法	穿地面构筑物	旋喷法	旋喷法	周边环境 不良影响
2	双圆土 压平衡	公路 隧道	粉土	搅拌法	穿地下构筑物	搅拌法	搅拌法	隧道质量 问题
3	泥水 平衡	能源 工程	砂土	SMW 法	穿地下桩基	SMW 法	降水法	施工安全 问题
4	网格 挤压	水环境 工程	岩土	降水法	穿越江河	降水法	冰冻法	盾构机械 设备故障

在属性量化的基础上, 原有案例能被重新表示为如表 2 所示的形式。

表 2 案例量化表示

案例	盾构 类型(T)	隧道 功用(U)	地质 状况(S)	风险所处工序(P)				风险 特征(C)	原因 分析(I)
				出洞	推进	进洞	旁通道		
1	4	3	1	3	-	-	-	1&2&3&4	管片碎裂
2	1	1	1	1&2	-	-	-	1&2	轴线偏移
3	1	2	1	4	-	-	-	2	轴线偏移
4	1	1	2	-	-	1	-	1&2	管片碎裂
5	1	1	1	-	-	1&2	-	1	轴线偏移
6	2	1	2	-	2	-	-	1	轴线偏移
7	2	1	1	-	-	2	-	1&2&4	轴线偏移
8	1	1	1	-	2	-	-	2	管片碎裂
9	1	1	1	-	-	1&2	-	1&2	管片碎裂
10	1	1	1&2	1&2	-	-	-	2	轴线偏移

#### 3.2 属性约简

在属性量化的基础上, 可以在隧道风险案例推理的众多属性中找出关键属性, 即属性约简过程。属性约简可以通过构造差别矩阵, 由差别矩阵导出差别函数并对其化简得到。差别矩阵的单个元素算法如下:

$$m_{ij} = \begin{cases} \text{属性} & (A_i \neq A_j) \cup (f_i \neq f_j) = \text{true} \\ \emptyset & (A_i \neq A_j) \cup (f_i \neq f_j) = \text{false} \end{cases} \quad (1)$$

根据式(1)可以求出核属性, 得到差别矩阵<sup>[5]</sup>如图 2 所示。

	案例 2	案例 3	案例 5	案例 6	案例 7	案例 10
案例 1	$T \cup U \setminus P \setminus C$	$T \setminus U \setminus S \setminus P \setminus C$				
案例 4	$P \setminus C$	$U \setminus S \setminus P$	$S \setminus P \setminus C$	$T \setminus P \setminus C$	$T \setminus P \setminus C$	$S \setminus P \setminus C$
案例 8	$P \setminus C$	$U \setminus C$	$P \setminus C$	$T \setminus C$	$T \setminus P \setminus C$	$P$
案例 9	$P$	$U \setminus P \setminus C$	$C$	$T \setminus S \setminus P \setminus C$	$T \setminus P \setminus C$	$S \setminus P \setminus C$

图 2 差别矩阵

根据差别矩阵, 找出核心属性, 简化冗余数据, 得到约简后的数据表。

在案例数据表中属性较多, 可能出现一些冗余属性, 例如上述矩阵中,  $S$  和  $U$  可以省略, 不会影响数据表的知识结构。因此, 采用以下步骤进行属性约简<sup>[6]</sup>:

- (1)在差别矩阵中删除包含核属性的项;
- (2)统计频率最大的剩余属性;
- (3)核属性+最大剩余属性的集=相对简化集;
- (4)删除该属性后, 转第(3)步继续, 从而获得多个相对简化集。

每个简化集作为案例的一个索引, 有多个简化集, 也就是有多个案例索引, 案例索引的建立是案例检索的基础。例如, 图 2 的简化结果为  $(P, C)$ 。

通过粗糙集技术可以有效利用差别矩阵进行属性简约。可以利用属性相似度阈值处理案例库的冗余, 有选择地删除多余案例, 以兼顾系统的能力与效率。

### 4 案例检索

当一个基于隧道工程案例的系统遇到新问题时, 系统会在案例库中寻找相似的案例或案例集, 此过程成为检索。案例检索策略的选择和检索过程直接决定着整个系统的质量和效率。

最近相邻法是目前应用最广泛的检索策略, 将案例视为

散布在以各属性为坐标轴的  $n$  维空间中, 而案例库是已知点的集合。即将隧道工程中可能出现的风险因素一一列出来并设计相关推理机制。当一个新问题被提出时, 先在案例空间中确定其坐标, 然后找到距离最近的“案例点”作为参考。计算过程如下:

假设有一个隧道新范例(查询范例), 其中包含一组属性{隧道类型  $a_1$ , 隧道功用  $a_2$ , ..., 风险工序  $a_n$ }, 用式(2)计算该案例和案例库中案例的匹配度。

$$\text{案例匹配度} = \frac{\sum(\text{相似度}[i] \times \text{权重}[i])}{\sum \text{权重}[i]} \quad 0 \leq \text{权重}[i] \leq 1 \quad (2)$$

式(2)中的相似度<sup>[7]</sup>用式(3)取得。

$$\text{相似度}[i] = (1 - \text{距离}[i]) \times \text{权重}[i] \quad (3)$$

式(3)中距离的计算方法如下: 如果为属性为连续值则按式(3)计算; 如果为离散属性, 例如沉降程度等类型、隧道功用、土质、工序、工法等属性时, 则完全匹配距离为 0, 否则距离为 1。

$$\text{距离}[i] = \frac{|y_i - Y_i|}{Y_i} \quad (4)$$

其中,  $i$  表示第  $i$  个匹配的关键特征 ( $1 \leq i \leq n$ ); 相似度 $[i]$  表示案例间第  $i$  个关键特征的相似度; 权重 $[i]$  表示第  $i$  个关键特征的权重因子;  $Y_i$  表示案例库中案例的第  $i$  个特征属性的值;  $y_i$  表示当前发生风险的案例的第  $i$  个特征属性的值。

一般情况下, 权值由专家设定, 必然受其主观意识的影响。本文借助粗糙集理论从现实数据中客观地得出特征属性的重要程度, 并将其作为属性权重值<sup>[8]</sup>。

采用上述粗糙集理论中信息表达系统的定义, 属性的重要性可用式(5)进行判断, 即

$$\beta_a = \text{card}(\text{pos}_s(D)) / \text{card}(U) - \text{card}(\text{pos}_{c-a}(D)) / \text{card}(U) \quad (5)$$

其中,  $\text{card}(\cdot)$  表示集合的基数。

属性  $a$  的重要度通过属性集中有无  $a$  属性的导致分类  $U/D$  的正域大小的差异来表示, 在此基础上, 将所有属性的重要程度进行归一化处理, 即

(上接第 238 页)

由图 2 可以看到, SymbolDraw 是暴露给客户使用的边界类, 用以实现 GIS 客户端铺设符号的要求。客户的使用接口是 symdraw(本文省略了客户使用的参数, 仅表示结构), SymbolDraw 组合了符号系统的符号绘制接口 ISymDrawFeature, 符号绘制接口 ISymDrawFeature 封装了绘制基本要素点、线、面上铺符号的方法接口, 而将具体实现都放在相应子类中。

在当前版本中定义了 2 种实现子类, 分别对应矢量浏览使用的 SymDrawFeature 和制图精确绘制使用的 SymDrawFeaturePrecise。矢量浏览对应数字设备的显示, 要求速度必须很快, 能满足用户拖动、放大、缩小的需要, 尽量减少延迟的时间。制图对应的是打印输出设备, 要求必须精确、美观, 每个符号都要贴合地理要素。可见, 上述 2 种实现算法是不一样的, 因此, 它们实现的接口函数不同。但对于边界类客户端来说, 只有一种接口, 系统只要根据客户状态在 Runtime 阶段装入适当的实现算法即可。当系统需要实现别的需求算法时, 只要实现新算法绘制子类, 用户接口和边界类不需要做任何改动就能动态装入新算法。

$$W_a = \frac{\beta_a}{\sum_{i \in C} \beta_i} \quad (6)$$

利用属性重要度进行相似度量, 度量过程中利用权值进行修正, 针对表 2, 根据式(6)和式(5), 可以计算出核心属性  $P, C$  的权重分别为 5/9 和 4/9。

## 5 结束语

本文从案例推理的角度, 探索盾构隧道工程风险评价方法, 将粗糙集理论引入盾构施工领域, 对隧道工程案例属性进行有效约简并确定了各属性的权重, 使其成为对隧道工程进行有效风险管理的新方法。该方法与规则推理相结合, 建立一个以知识管理为基础的施工风险管理系统——盾构施工风险控制知识管理系统, 辅助施工人员进行风险管理, 有助于盾构施工安全高效地进展。

## 参考文献

- [1] 黄宏伟. 隧道及地下工程建设中的风险管理研究进展[J]. 地下空间与工程学报, 2006, 2(1): 13-20.
- [2] Golding A R. Review of Case-based Reasoning[J]. AI Magazine, 1995, 16(2): 85-96.
- [3] 张冬玲. 基于粗糙集理论的属性约简算法的实现[J]. 计算机应用, 2006, 26(6): 78-79, 82.
- [4] 孙翎, 张金隆, 迟嘉昱. 基于粗糙集的 CBR 系统案例特征项权值确定[J]. 计算机工程与应用, 2003, 39(30): 44-46.
- [5] 杨明, 杨萍. 差别矩阵浓缩及其属性约简求解方法[J]. 计算机科学, 2006, 33(9): 181-183, 269.
- [6] 陈堂敏. 面向用户的知识量最佳属性约简算法在数控机床故障预测中的应用[J]. 机械科学与技术, 2006, 25(2): 163-167.
- [7] 骆敏舟, 周美立. 实例推理检索中相似度量方法的研究[J]. 合肥工业大学学报: 自然科学版, 2001, 24(6): 1091-1094.
- [8] 李茹, 任海涛, 刘开瑛. 基于案例推理中特征权重自动学习方法研究[J]. 山西大学学报: 自然科学版, 2004, 27(3): 245-248.

## 3 结束语

本文设计并实现大型 GIS 平台 VegaGIS 可视化系统架构。该架构已在多个领域(包括金水工程、城市应急、城市管理和公安系统)的大型工程中得到应用, 并在 2004 年和 2005 年获得了科技部 GIS 专业软件测评的表彰和推荐。

## 参考文献

- [1] 鄂伦, 张晶, 赵伟. 地理信息系统[M]. 北京: 电子工业出版社, 2002.
- [2] Vlissides J M, Linton M A. Unidraw: A Framework for Building Domain-specific Graphical Editors[J]. ACM Transactions on Information System, 1990, 8(3): 237-268.
- [3] Johnson R E. Documenting Frameworks Using Patterns[C]// Proceedings of Conference on Object-oriented Programming Systems, Languages and Applications. Vancouver, British Columbia, Canada: [s. n.], 1992.
- [4] Ormsby T, Napoleon E, Burke R, et al. Getting to Know ArcGIS Desktop[M]. [S. l.]: ESRI Press, 2004.