

基于案例挖掘的边坡稳定性智能评价系统研究

谢全敏, 陈立文, 夏元友

(武汉理工大学 土木工程与建筑学院, 武汉 430070)

摘 要: 虽然边坡灾害治理实践已经积累了大量的成功的或失败的边坡治理工程案例, 但这些工程案例产生的大量数据信息未被充分利用与开发, 造成了极大的资源浪费。为此, 基于边坡工程案例, 应用数据挖掘与知识发现和递归的自-组织模糊神经推理网络的方法, 初步提出了一种基于案例挖掘的边坡稳定性智能评价系统, 并通过案例挖掘的应用实例表明了该系统的有效性和可行性。

关 键 词: 边坡案例; 数据挖掘; 案例挖掘; 递归的自-组织模糊神经网络

中图分类号: TU 457; TB 115

文献标识码: A

Study on intelligent evaluation system of slope stability based on case mining

XIE Quan-min, CHEN Li-wen, XIA Yuan-you

(School of Civil Engineering and Architecture, Wuhan University of Technology, Wuhan 430070, China)

Abstract: Slope hazard is a serious natural hazard next only to earthquake and floodwater. It has brought about great losses to the economic and social development worldwide. Each year it causes a loss up to billions of dollars in the whole world. A striking amount of money has been spent on the prevention and redress of the hazard; and much more losses have been incurred indirectly in addition. Many slope treatment cases have been already accumulated for over a long time of slope treatment, but much data information of these slope cases are not completely used and developed, which results in huge resources waste. For this reason, applying the methods of the data mining, knowledge discovery in databases and recurrent self-organizing neural fuzzy inference network, an intelligent evaluation system of slope stability based on case mining is presented. Meanwhile, practical applications illustrate validity and feasibility of the system.

Key words: slope case; data mining; case mining; recurrent self-organizing neural fuzzy inference network

1 引 言

边坡灾害(滑坡、崩塌和泥石流等)对人类社会发展和经济建设的危害是世界范围普遍关注的重要问题, 每年给世界各国造成的经济损失估计达数十亿美元。在我国已受到边坡灾害威胁和可能受到边坡灾害威胁的地区约占全国陆地面积的 1/5~1/4, 每年因边坡灾害造成的损失达 60 亿元以上, 给国家和人民生命财产带来巨大损失, 其危害十分巨大。国家每年投入边坡的治理费用十分巨大, 如仅三峡库区边坡治理首期和第 2 期就分别投资了 40 亿元和 80 亿元。在国家投入巨资治理边坡的同时, 不断积累了大量的成功的或失败的边坡治理工程案例, 这些边坡治理工程案例产生的大量数据信息未

被充分利用与开发, 造成极大的资源浪费^[1-5]。其实这些数据信息中蕴涵着很多宝贵的边坡治理知识和经验, 急需去挖掘发现, 并进行开发利用, 使国家投入的巨资所产生的边坡治理案例的数据信息资源得以科技升华, 将所得到的研究成果服务于今后全国的边坡治理, 这必将产生巨大的经济效益和社会效益。为此本文根据 ANN 方法和案例推理的各自特点, 应用递归的自-组织模糊神经推理网络的方法, 初步提出一种基于案例挖掘的边坡稳定性智能评价系统^[6-9]。

2 案例推理与案例挖掘的关键技术

2.1 基于案例推理(CBR)原理及其基本步骤

基于案例推理的工作原理, 就是人们的认知心

理过程,即假定人们总是利用已有的经验和知识来解决具有相似性的新问题。实际上,真正的专家不会将问题按规则一条一条匹配,直接利用以前的经验才是最为自然的做法。只有新手才会认真回忆,仔细地按照前人或专家制定的规则、指南按部就班地做下去。当专家遇到新问题时,他们的做法往往是:识别新问题与以前遇到过的问题的相似性;询问一些有关联的贴切的问题,以确定新问题到底与哪些经验相似或应将以前的经验作些什么调整;最后得出结论。基于案例推理就是按照人们的这种自然地认知心理过程,检索出与新问题相近的案例,并补偿那些匹配不一致的地方,形成问题的辅助求解方案。基于案例推理方法应用领域广泛,表示形式和实现方法可以是多种多样的,但究其根本,基于案例推理的系统一般包括以下步骤:(1)问题的表示和存储;(2)对以前案例的检索和匹配;按照一定的算法找出当前新问题的最佳相似案例;(3)对检索出的方案,用递归的自-组织模糊神经推理网络方法进行调整得出新解;(4)评价方案的有效性;(5)调整/维护案例库。

2.2 案例仓库及其特点

案例库是由带有特定标识的一个个案例组成的,当然也可按某种原则对其进行模糊聚类。一个抽象的案例可表述为 5 个因素构成的 5 元组,即

$$Case = (P_c, R_c, A_c, T_c, E_c) \quad (1)$$

式中: P_c 为有限个因素 $(P_1, P_2, \dots, P_n)$ 构成的用来表达问题的复合因素; T_c 为期望达到的目标集; A_c 为实际达到的效果或得到的答案集; R_c 为从问题 P_c 到解答 A_c 的操作、变换,可以是抽象为矩阵或函数形式表示的关系,且有 $A_c = P_c R_c$ 成立; E_c 为对案例的评价。

案例仓库是面向主题的、集成的、稳定的、不同时间的案例的集合,用于支持经营管理中的决策制定过程。

2.3 案例仓库的结构和案例仓库管理系统

案例仓库的数据量是大约 100 倍于案例库,它的存储结构形式的设计以及数据查询和提取的效率就显得很重要。一般采用分布式结构形式,案例仓库管理系统的一项重要工作是实现对传统案例库进行提取、清理和转载案例到案例仓库中,这项工作占整个管理工作的 80%。

2.4 案例仓库中的分析工具

案例仓库虽然存有大量的案例,但提供辅助决策的信息需要利用各种分析工具,如 OLAP 工具、

DM 工具、统计分析和查询优化工具等。这些工具直接影响辅助决策的效果。

2.5 案例挖掘的方法

目前,已经形成了分类、聚类和分割等多种案例挖掘的方法。每种方法均有它的优点,也有一些不足,还需要研究更多的案例挖掘方法,以及多种方法的综合,使案例挖掘得到的知识具有更强的辅助决策效果。

3 边坡稳定性智能评价系统

决策支持技术的发展和基于模型库的 DSS 的开发困难,促使人们重新考虑 DSS 的结构体系。本文提出了案例仓库+OLAP+案例挖掘+递归的自-组织模糊神经推理网络=IDSS 的解决方案。案例仓库、OLAP 和案例挖掘是作为 4 种独立的信息处理技术出现的。案例仓库用于数据的存储和组织; OLAP 集中于数据的分析; 案例挖掘致力于决策信息的自动发现; 递归的自-组织模糊神经推理网络用于案例训练,使其更接近于人的思维模式,这 4 种技术具有内在的联系性和互补性,把它们结合起来,形成了新的决策支持系统结构,如图 1 所示。

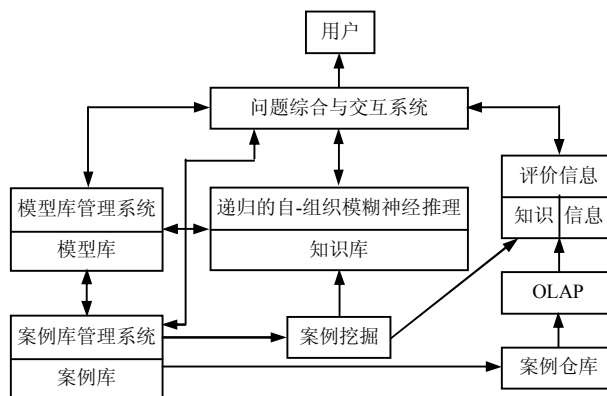


图 1 基于案例挖掘的边坡稳定性智能评价系统
Fig.1 The intelligent evaluation system of slope stability based on case mining

该结构的特点是以大量的数据为基础,系统由数据驱动。案例仓库对低层的业务数据进行集成、转换和综合,重新组织成面向主题的全局数据视图。OLAP 是在全局数据视图上进行多维数据分析。案例挖掘是从案例库或案例仓库中进行挖掘,获得辅助决策信息。由于案例仓库、OLAP、案例挖掘都是数据驱动的,故它们的共同点是寻找大量数据中的规律性,从 3 个不同角度为决策者提供辅助决策信息,这是决策支持系统新的发展方向。递归的自-组织模糊神经推理网络与专家系统是人工智能领域

中的两个热门题目,它们已成功地应用于许多领域。但它们在运行中各自遇到了许多困难,这是由它们自身机制特点所决定的,所以其应用受到了很大的限制。专家系统的致命弱点是:知识获取困难,因为它是把各种专家知识从人类专家的头脑中或其他知识源那里人工转换到知识库中,费时低效;对于动态和复杂系统,由于其推理规则是固定的,难以适应变化的情况;对于有大量数据的输入,由于数据中存在较强的干扰因素,使得专家系统的结果受到很大的影响;专家系统不能从过去处理过的事例中继续学习,这使知识获取变得更加困难。与专家系统相比,递归的自-组织模糊神经推理网络具有良好的自组织、自学习和适应能力,因而特别适用于处理复杂问题或开放系统,这正好可以弥补专家系统的不足。同样,递归的自-组织模糊神经推理网络也有其弱点:递归的自-组织模糊神经推理网络的知识是分布在整个系统内部,对用户而言是个黑箱;而且,它对于所采取的结论不能做出合理的解释,而这正是专家系统的长处。如果将递归的自-组织模糊神经推理网络与专家系统结合起来,充分发挥二者的优势,克服各自的不足,取长补短,产生巨大的合力,将会更好地满足实际需要。递归的自-组织模糊神经推理网络主要是作为专家系统的前端完成知识获取功能,将大量历史案例数据输入神经网络,经过训练学习达到稳定状态时,所获取的知识存储体现在网络神经元互连分布及其权值上,必须通过某种转换机制将其变换成专家可理解的符号规则。在这种结合模式中递归的自-组织模糊神经推理网络主要作用是知识获取,专家系统的主要作用是完成推理和解释任务。当然神经网络也可完成推理功能,那么究竟由谁来实现,可根据实际情况来确定。例如,求解问题的输入数据不全时,可采用专家系统的符号推理系统实现;若输入数据完全时,可直接利用递归的自-组织模糊神经推理网络完成推理。这种模式实现的难点在于转换方法的选择、设计和实现。智能决策的新结构体系是将决策支持系统、专家系统、案例仓库、案例挖掘结合起来,即将多种不同的辅助决策方式结合起来,起到相辅相成的作用,进一步提高辅助决策的效果。

4 实例研究

边坡稳定性往往与众多复杂因素有关。国内外许多研究者对影响边坡稳定性的因素进行了广泛的探讨,这些因素归纳起来大致为以下两个方面:

(1) 地质因素:边坡体的物理力学性质、边坡体结

构特征、结构面特征、气候条件、风化程度、地下水、地应力及地震等。(2) 工程因素:边坡开挖的几何形状(坡度、坡高)、爆破震动、机械震动、人工加载或卸载等。

为了说明了本文提出的系统的有效性和可行性,采用三峡库区研究程度较高的 48 个边坡工程案例进行分析,以其中 36 个案例作为训练样本,其余 12 个作为测试样本。36 个训练样本边坡的稳定性影响因素指标及量化结果见表 1。限于篇幅未列出 12 个测试样本边坡的稳定性影响因素指标及量化结果。

表 1 边坡案例评价指标值
Table 1 The evaluation indices values of slope cases

斜坡 编号	外形 特征	滑面 特征	滑带 强度	近期 活动	河流 作用	后缘 加载	暴雨 作用	地震 作用	人工 活动
1	1.0	0.0	0.5	0.5	0.5	0.0	1.0	0.5	0.0
2	0.5	1.0	0.0	0.5	0.5	0.0	1.0	0.5	0.0
3	0.5	1.0	0.0	0.5	0.5	0.5	1.0	0.5	0.0
4	1.0	1.0	1.0	1.0	0.5	0.5	1.0	0.5	1.0
5	1.0	0.5	0.5	0.5	0.5	0.0	0.5	0.5	0.0
6	0.5	1.0	0.5	1.0	1.0	1.0	0.5	0.0	0.5
7	0.5	1.0	0.5	0.5	1.0	1.0	0.5	0.0	0.0
8	0.5	1.0	0.5	1.0	1.0	1.0	0.5	0.0	0.0
9	0.0	1.0	1.0	0.0	1.0	1.0	0.5	0.0	0.0
10	0.5	0.5	0.5	1.0	1.0	0.0	0.5	0.0	0.0
11	1.0	0.5	1.0	1.0	0.0	0.0	0.5	0.0	0.0
12	1.0	0.5	0.5	0.0	1.0	0.0	0.5	0.0	0.0
13	1.0	1.0	1.0	1.0	1.0	0.0	0.5	0.0	0.0
14	1.0	1.0	0.5	1.0	0.0	0.5	0.5	0.0	0.0
15	1.0	1.0	1.0	1.0	1.0	0.0	0.5	0.0	0.0
16	1.0	0.5	1.0	0.5	0.5	0.5	0.5	0.0	0.0
17	1.0	1.0	1.0	1.0	0.5	0.0	0.5	0.0	0.0
18	1.0	1.0	0.5	0.5	0.5	0.0	0.5	0.0	0.0
19	1.0	0.0	0.5	0.0	1.0	0.0	0.5	0.0	0.0
20	0.5	0.5	0.0	0.0	1.0	0.0	1.0	0.0	1.0
21	0.5	0.5	0.5	0.5	1.0	0.5	1.0	0.0	0.5
22	1.0	1.0	1.0	1.0	1.0	0.5	1.0	0.0	0.5
23	0.5	1.0	0.5	1.0	1.0	0.5	1.0	0.0	1.0
24	0.5	0.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0
25	0.5	0.0	1.0	0.5	0.0	1.0	1.0	0.0	0.5
26	1.0	0.0	0.5	0.0	1.0	0.0	1.0	0.0	0.5
27	0.5	0.0	0.0	0.5	1.0	0.5	1.0	0.0	0.5
28	1.0	0.0	0.0	0.5	1.0	1.0	1.0	0.0	1.0
29	0.5	0.0	0.5	0.5	0.5	0.5	1.0	0.0	1.0
30	0.5	0.0	0.5	0.5	0.5	0.5	1.0	0.0	0.0
31	0.5	0.0	0.0	0.5	0.5	0.5	0.5	0.0	0.0
32	0.5	1.0	1.0	1.0	0.5	0.5	0.5	0.0	0.0
33	0.5	0.0	0.0	0.0	1.0	0.5	0.5	0.0	0.0
34	0.0	0.0	0.0	0.0	0.5	0.5	0.5	0.0	0.0
35	0.5	0.5	0.0	0.5	1.0	0.0	0.5	0.0	0.0
36	0.5	0.0	0.0	0.0	1.0	0.0	1.0	0.0	1.0

边坡案例按 6 级分类^[10-12],采用聚类法进行计算,36 个训练样本边坡分类结果见表 2 所示。

表 2 边坡案例聚类分级
Table 2 The classification of slope cases

样本	级别	样本	级别	样本	级别
1	IV	13	III	25	IV
2	IV	14	III	26	VI
3	IV	15	III	27	V
4	I	16	III	28	IV
5	III	17	III	29	IV
6	II	18	III	30	IV
7	II	19	V	31	IV
8	II	20	V	32	III
9	II	21	IV	33	VI
10	III	22	I	34	VI
11	III	23	I	35	V
12	V	24	V	36	V

首先,应用本系统对 36 个边坡训练样本进行训练;然后,用训练好的网络结构模型对 12 个边坡测试样本进行推理,测试结果见表 3 所示。其结果正确率为 91.6%,这说明本文提出的基于案例挖掘的边坡稳定性智能评价系统是有效的和可行的。

表 3 边坡样本测试结果
Table 3 The test results of slope samples

样本	级别	测试	样本	级别	测试	样本	级别	测试
37	V	V	41	V	V	45	II	II
38	II	II	42	IV	III	46	III	III
39	II	II	43	V	V	47	III	III
40	V	V	44	IV	IV	48	IV	IV

5 结 语

本文提出的基于案例挖掘的边坡稳定性智能评价系统主要是以知识和模型为主体,以递归的自-组织模糊神经推理网络为挖掘方法,结合边坡案例,形成定性和定量相结合的辅助决策信息。本文选取三峡库区研究程度较高的 48 个边坡工程案例进行分析,说明了递归的自-组织模糊神经推理网络作案例挖掘方法的可行性和有效性。本文的研究仅仅是个初步的工作,还有很多研究工作需要做。在基于案例挖掘的边坡稳定性智能评价系统开发的过程中存在许多困难,例如:需要大量的边坡案例,且边坡工程案例千变万化,应如何建立实用的案例仓库,以及评价过程中存在许多不确定性因素等问题还有待于进一步研究。

参 考 文 献

[1] Feng X T, Hudson J A, Li S, et al. Integrated Intelligent methodology for large-landslide prevention[J]. *International Journal of Rock Mechanics and Mining Sciences*, 2004, 41(3): 505—510.

[2] 夏元友,李梅. 基于实例类比推理的边坡稳定性评价方法[J]. 岩土力学, 2003, 24(增刊): 300—303.
XIA Yuan-you, LI Mei. A method of slope stability evaluation using case-based reasoning[J]. *Rock and Soil Mechanics*, 2003, 24(Supp.): 300—303.

[3] 谢全敏,夏元友. 危岩块体稳定性的综合评价方法分析[J]. 岩土力学, 2002, 23(6): 775—777.
XIE Quan-min, XIA Yuan-you. Comprehensive assessment method of stability of critical rock-block mass and its application[J]. *Rock and Soil Mechanics*, 2002, 23(6): 775—777.

[4] 谢全敏,边翔,夏元友. 滑坡灾害风险评价的系统分析[J]. 岩土力学, 2005, 26(1): 71—74.
XIE Quan-min, BIAN Xiang, XIA Yuan-you. Systematic analysis of risk evaluation of landslide hazard[J]. *Rock and Soil Mechanics*, 2005, 26(1): 71—74.

[5] XIE Quan-min, XIA Yuan-you. A multi-objective decision-making method for the treatment scheme of landslide hazard[J]. *Journal of University of Science and Technology Beijing*, 2004, 11(2): 101—105.

[6] 葛永利,薛华成. 基于模糊-神经网络的援例支持系[J]. 管理科学学报, 1998, 1(2): 63—64.
GE Yong-li, XUE Hua-cheng. Case based support systems using fuzzy neural technology[J]. *Journal of Management Sciences in China*, 1998, 1(2): 63—64.

[7] JUANG CHIAFENG, LIN CHINTENG. Arcurrent self-organizing neural a fuzzy inference network[J]. *IEEE Transactions on Neureal Network*, 1999, 10(4): 828—839.

[8] 马江洪. 数据挖掘与数据库知识发现[J]. 工程数学学报, 2002, 19(1): 1—13.
MA Jiang-hong. Data mining and knowledge discovery in database[J]. *Chinese Journal of Engineering Mathematics*, 2002, 19(1): 1—13.

[9] SOIBELMAN LUCIO. Generating construction knowledge with knowledge discovery in databases[J]. *Computing in Civil and Building Engineering*, 2000, 2(3): 914—921.

[10] 谢全敏,夏元友. 边坡稳定性评价的自适应模拟聚类分析方法[J]. 灾害学, 2002, 17(1): 15—19.
XIE Quan-min, XIA Yuan-you. A self-adaptive clustering method based on the simulated annealing algorithm for evaluation of slope stability[J]. *Journal of Catastrophology*, 2002, 17(1): 15—19.

[11] 谢全敏,夏元友. 基于遗传算法的边坡稳定性评价的动态聚类方法[J]. 岩土力学, 2002, 23(2): 170—172.
XIE Quan-min, XIA Yuan-you. A dynamic clustering method for evaluation of slope stability based on genetic algorithms[J]. *Rock and Soil Mechanics*, 2002, 23(2): 170—172.

[12] 夏元友,肖峰. 斜坡稳定性评价的动态聚类分析方法[C]//第 6 届全国岩石力学与工程学术会议文集. 北京: 中国科学技术出版社, 2000: 562—565.