

基于校园 GIS 的空间数据挖掘系统原型

于永玲,李 向,宗思生,施进发
(郑州航空工业管理学院,河南 郑州 450015)

摘要:为了建立数字校园信息平台,必须对校园 GIS 数据进行分析 and 挖掘。文中提出将校园 GIS 与数据挖掘系统进行集成,空间数据库系统负责空间信息抽取,GIS 实现对挖掘结果进行可视化显示。搭建一个空间数据挖掘原型系统模型,利用设计的挖掘算法,挖掘出空间数据库中隐含的知识,并将空间数据挖掘结果进行可视化输出。该原型系统已成功应用于某高校校园 GIS,不仅可以挖掘显示知识(校园每个教学楼的上座率),还可以发现隐藏在校园 GIS 中的隐式知识,是对现有空间数据挖掘系统的有益改进和补充。

关键词:空间数据挖掘;系统原型;数字校园;关联规则;GIS

中图分类号:TP208

文献标识码:A

文章编号:1673-629X(2013)12-0161-03

doi:10.3969/j.issn.1673-629X.2013.12.038

Spatial Data Mining Prototype System Based on Campus GIS

YU Yong-ling, LI Xiang, ZONG Si-sheng, SHI Jin-fa

(Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China)

Abstract: In order to establish a digital campus platform, the campus GIS data must be analyzed and mined. The campus GIS and digital mining system is intergrated, the spatial database system for the extraction of spatial information, and SDM for the visual display of mining results. A spatial data mining prototype system is constructed, which can discover the implicit knowledge inside spatial database using the designed data mining algorithm. The prototype system has applied into a campus GIS, it not only discovers the explicit knowledge but also implicit knowledge hidden in the campus GIS, so that it is a beneficial improvement and complement for present spatial data mining prototype system.

Key words: spatial data mining; system prototype; digital campus; association rule; GIS

0 引言

高校是一个文化学术、教育资源比较集中的地方,随着校园办学规模的逐渐扩大,师生员工及其他设施的不断增加,使得校园信息也越来越庞大且日益复杂。而校园 GIS 就是充分利用 GIS 及计算机、网络等相关技术手段,对校园的各项资源(包括硬件和软件资源)进行全方位的管理,具有对校园地理、环境、资源、生态、校园基础设施和人员信息等复杂系统进行数字化和网络化管理、服务甚至提供辅助决策的功能。校园 GIS 每天都产生大量的数据,数据虽多,但知识贫乏、用处不大。为了从校园 GIS 数据中发现有用的知识,必须对其进行分析、挖掘^[1-5]。

目前,高校已经建成并投入使用的各项管理信息系统(MIS)已使学校的管理工作实现自动化,比如,教

务管理系统的使用使学生在任何一个地方,通过网络可以完成选课、查看课表、查看成绩等,而教师通过该系统完成教学信息的录入、查询等,所有这些任务的完成,都不受地点的限制,只需配有一台电脑,能够连接到因特网上就可以了。但是,这些管理信息系统的使用还不能使各项管理工作实现智能化。对校园 MIS 和 GIS 数据进行挖掘,可以实现对校园数据的理解,定位校园教学设施的位置,快速查找与位置相关的信息,进而发现存在于校园数据中的隐藏信息,构建校园信息平台,为用户提供数据查询和决策分析的能力,帮助管理层做出科学的决策,更好地为高校建设服务^[6-7]。

1 校园 GIS

校园地理信息系统主要面向高校师生、行政管理

收稿日期:2013-03-08

修回日期:2013-06-14

网络出版时间:2013-09-29

基金项目:航空科学基金(2012ZG55023);河南省科技攻关计划项目(112102210352,122102210200);河南省软科学研究项目(132400410621);河南省教育科学技术研究重点项目(2012B520062,2011A630043)

作者简介:于永玲(1971-),女,河南杞县人,副教授,研究方向为数据库、空间数据挖掘。

网络出版地址:<http://www.cnki.net/kcms/detail/61.1450.TP.20130929.1544.043.html>

人员,同时满足校外社会人士对本校的查询要求。主要有以下功能:

(1)利用空间和属性数据库,实现信息的可视化、管理、浏览、查询、输出。

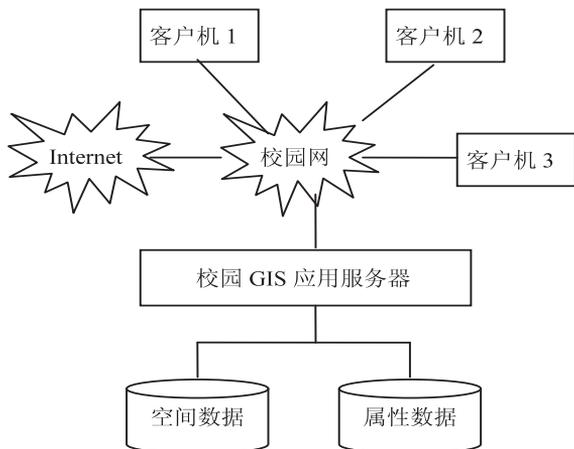
(2)提高校园的日常管理自动化水平,实现各种报表的自动生成、排序、打印等功能。

(3)提供一定的空间分析能力,为校园规划和管理决策提供依据。

(4)实现网上发布功能,世界各地各界人士都可以通过互联网了解到学校及时、真实、详尽的信息。

(5)提供方便地访问各个院系、机关、院所网页的功能。

该系统采用 C/S 体系结构,数据存储和数据统计分析处理由 GIS 服务器完成,客户端只显示处理结果,这样可以减少网络数据流量,降低对客户机的配置要求。各客户机可通过校园网实现对 GIS 服务器的访问,校外用户可通过 Internet 获得数据,从而实现信息共享。不同的用户享有不同的访问权限,只有设计人员及管理人员具有修改权限,从而实现体系的安全访问机制。校园 GIS 结构如图 1 所示。



2 基于校园 GIS 的空间数据挖掘系统原型

设计基于校园 GIS 的空间数据挖掘系统的目的就是通过对校园 GIS 数据进行统计分析,利用有效的挖掘方法挖掘出隐藏的有用知识,并将挖掘结果可视化输出。校园 GIS 可以完成空间数据的输入、存储、操作和管理,并以图形等直观形式表达输出结果,但是它的空间分析能力不够强,要借助于空间数据挖掘作为空间分析模型的一个模块,利用其强大的发现知识的功能弥补 GIS 的不足^[8-10]。

2.1 基于校园 GIS 的空间数据挖掘过程

基于校园 GIS 的空间数据挖掘就是将数据挖掘方法应用于 GIS 空间数据库,从校园 GIS 数据库中挖掘

有用的信息,挖掘过程可以概括为以下六步:

第一步,理解问题,准备数据。熟悉相关背景知识,分析用户需求。比如一个同学要去自修,他想知道距离他所在位置最近的哪个自修室有空闲地方?主管教学的领导想知道哪个专业学风较好?哪个专业学风较差?影响学生学风的因素有哪些?此阶段是数据挖掘的必经步骤,往往要花很多的时间。

第二步,选取数据。设计的数据挖掘算法不能直接应用于数据库,这就需要对数据库进行查询、统计,从相关的数据库中选取用户感兴趣的对象及其属性数据,并将它们存储为数据挖掘算法可以识别的格式。

第三步,数据预处理。此阶段主要是对数据选取阶段产生的数据进行再加工,如将定名量转换为定序量,连续数据离散化。

第四步,选择合适的挖掘算法。目前,空间数据挖掘算法有很多,考虑到空间分类数据的特点,文中采用前面介绍的空间关联规则挖掘算法。该算法利用逐步求精方法,是对经典的 Apriori 的改进。

第五步,进行空间数据挖掘。这个阶段是真正意义上的空间数据挖掘。运用前文介绍的挖掘算法,对数据源中的数据进行挖掘,提取出对用户有用的信息和知识,并将知识以可视化的方式输出。

第六步,结果解释。对数据挖掘结束后的结果进行解释,在此过程中,为了取得更为有效的知识,可能会返回前面处理步骤中。获取的有用信息和知识可以为下一步建设高校“数字校园”信息系统平台提供信息。

2.2 系统模型

文中设计的基于校园 GIS 的空间数据挖掘系统以校园 GIS 空间数据库为基础,利用 GIS 提供的数据分析和地理数据可视化的强大功能,再加上由高级语言开发的数据挖掘模块,并建立一个人机接口处理用户的指令和显示挖掘结果来共同开发数据挖掘系统。设计的数据挖掘系统模型如图 2 所示。

该模型框架大致分为三层:

第一层是数据源,要挖掘的数据既有 GIS 空间数据库,又有属性数据,还有 MIS 数据,对这些源数据按主题抽取构建数据仓库,再由空间数据仓库管理系统提供的索引、查询优化功能提取与问题相关的数据。

第二层是挖掘器,利用空间数据挖掘系统中的各种挖掘方法提取有用的数据,采用人机交互方式,允许用户根据要获取的知识类型,选用不同的挖掘方法。这里主要是应用前面介绍的空间统计分析方法和空间分类数据关联规则挖掘算法。

第三层是人机交互用户界面。利用高级语言的编程功能,编写友好的用户界面,接收用户的挖掘命令和

必要的参数,将挖掘出的有用信息和发现的知识以可视化的方式显示给用户。

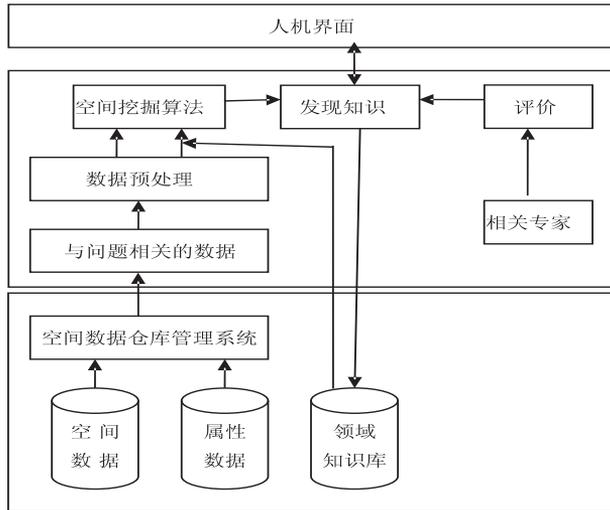


图 2 基于校园 GIS 的空间数据挖掘模型

2.3 系统实现

文中设计的空间数据挖掘原型系统是将数据挖掘系统和校园 GIS 集成,实现校园 GIS 智能化。目前,常用的数据挖掘系统和 GIS 的集成方式有三种模式,分别是松散耦合式、嵌入式和混合式。松散耦合式将 GIS 作为空间数据库,借助于高级语言对 GIS 空间数据库进行数据挖掘,该模式要求用户自主开发 GIS 空间数据的数据管理和分析功能,开发难度大;嵌入式是将空间数据挖掘器嵌入到 GIS 内部,作为 GIS 的一个模块,利用 GIS 的空间查询和分析或直接从 GIS 数据库中提取所需要的数据,利用挖掘器将有用的知识和信息挖掘出来,该模式的数据挖掘器不够灵活;混合式是将前两种方法进行结合,尽可能利用 GIS 的空间数据管理与分析功能,尽量减少程序开发的难度和工作量,又保持数据挖掘器的灵活性^[11-13]。综上几种模式,该系统采用第三种。

用户使用该系统进行数据挖掘时,首先要从数据仓库中读取数据,判断用户是否要进行数据挖掘,若是,按挖掘的主题利用前面设计的挖掘算法挖掘知识,然后,对挖掘出的知识进行评价,最后将挖掘出的知识进行可视化输出。设计的系统流程图如图 3 所示。

3 系统应用与分析

将原型系统应用于郑州一所高校校园 GIS,不仅可以挖掘显式知识——校园每个教学楼的上座率,还可以发现隐藏在校园 GIS 中的隐式知识。

3.1 按教学楼统计上座率

单击“按教学楼统计上座率”按钮,则会打开校园地图,点击你想察看的某个教学楼,就可以看到该教学楼的座位空闲情况。

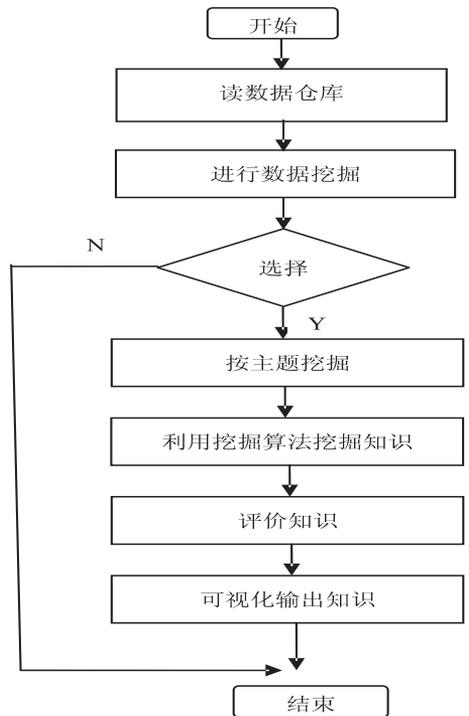


图 3 系统流程图

所有教学楼、实验楼和宿舍楼在校园的位置分布信息存储在校园 GIS 中,教学楼的使用信息来自教务管理系统,而实验楼、电子阅览室都有相应的管理系统,这些数据都是动态的,存放在相应的数据库中。通过对这些数据进行统计分析,就能得到每个教学楼、实验楼的上座率。有了这种分析结果,某个同学再想去自修,就不必逐个楼、逐个教室去查找了。借助于该系统,可以查看他想去的距离较近的教学楼上座率,然后轻松选择自修地点,特别是对那些学生人数多、校园面积大、宿舍楼距自修地点较远的校园,使用该系统的优势就更明显了。这也是现有的空间数据挖掘系统所不具备的。

3.2 空间数据关联挖掘

校园 GIS 中还隐藏有大量的信息,这些信息是隐式知识,必须利用挖掘算法才能得到。利用空间分类数据关联规则挖掘算法分别设置支持度阈值、置信度阈值、最小兴趣度、最小推荐度,就可挖掘出空间数据之间的关联规则。

从得到的强关联规则可以看出某段时期哪个年级哪个专业的学风较好,影响学生学风的因素有哪些。有了这种分析结果,负责学生的管理人员就可以针对不同的情况制定不同的对策,尽而提高学生学习的积极性,营造良好的学习风气,创造和谐校园。

4 结束语

将校园 GIS 与数据挖掘系统进行集成,设计一个

(下转第 167 页)

地位,它接收视图层用户的请求,调用模型层进行业务逻辑处理,把处理的结果返回到视图层呈现给用户。以“猪只销售记录查询”功能的控制层实现为例,控制层接收到视图层发出的需要查询指定起止日期范围内的销售记录请求,控制层利用 Entity Framework,以 LINQ 方式返回符合查询条件的 JSON 格式的数据记录,控制层把 JSON 格式的数据返回给前端视图层,视图层解析 JSON 格式的数据并以表格方式显示输出^[13]。

4 结束语

随着计算机和通信技术的飞速发展,很多行业提出了建设信息化、网络化、数字化管理平台的需求,文中以生猪养殖行业为例,采用 ASP.NET MVC 技术,研究并设计了一个适用于生猪养殖行业的数字化管理平台,该平台具有技术先进、扩展性强、性能高效和实用性好等优点。通过在陕西杨凌正泰畜牧养殖有限公司的试用,该平台显著提升了生猪养殖过程的信息化管理水平。当然,该平台目前尚有许多不完善之处,随着今后的不断改进,它将对生猪养殖过程中更多环节提供精细化的控制和管理。

参考文献:

[1] 方 遼,王 娟.关于国外生猪养殖信息化模式探究[J].

(上接第 163 页)

空间数据挖掘原型系统,不仅可以挖掘显式知识——校园每个教学楼的上座率,还可以发现隐藏在校园 GIS 中的隐式知识,并将数据挖掘结果可视化输出。目前,该原型系统已成功应用于某高校校园 GIS,是现有空间数据挖掘系统的有益改进和补充。

参考文献:

- [1] 胡圣武,李鲲鹏.空间数据挖掘的方法进展及其问题分析[J].地球科学与环境学报,2008,30(3):311-318.
- [2] 蒋良孝,蔡之华.空间数据挖掘的回顾与展望[J].计算机工程,2003,29(6):9-10.
- [3] 张 楠,曲海平,刘 念,等.空间数据挖掘的研究进展[J].微处理机,2007(2):1-4.
- [4] 蒋 旻.基于空间数据库的数据挖掘技术[J].武汉科技大学学报(自然科学版),2002,25(2):183-186.
- [5] Wang Jinlin, Chen Xi. Knowledge discovery and data mining [C]//Proc of 2009 second international workshop on WKDD. Moscow, TBD, Russia; [s. n.], 2009: 159-162.
- [6] 李金生,刘 岩,周 园,等.基于 MapInfo 的校园 GIS 查询

猪业科学,2009,26(5):18-21.

- [2] 唐 灿.下一代 Web 界面前端技术综述[J].重庆工商大学学报(自然科学版),2009,26(4):350-355.
- [3] Freeman A. Pro. ASP.NET MVC4[M]. [s. l.]: Apress, 2012.
- [4] 方 柯.基于 ASP.NET MVC 框架的 Web 应用开发研究[D].上海:华东师范大学,2011.
- [5] 林乐逸.基于 ASP.NET MVC 和实体框架的软件项目管理平台[D].上海:上海交通大学,2012.
- [6] Lerman J. Programming entity framework[M]. 2nd ed. [s. l.]: O'Reilly, 2010.
- [7] Lerman J, Miller R. Programming entity framework: DbContext[M]. [s. l.]: O'Reilly, 2012.
- [8] Lerman J, Miller R. Programming entity framework code first[M]. [s. l.]: O'Reilly, 2012.
- [9] 马鹏烜.基于 LINQ to Entity 数据访问技术的应用研究[J].现代计算机(下半月版),2011(13):41-43.
- [10] Calvert C, Kulkarni D. Essential LINQ[M]. [s. l.]: Addison Wesley, 2009.
- [11] Kuan J. Learning hightcharts[M]. [s. l.]: Packt Publishing Limited, 2013.
- [12] 何 丽,龚 旺. JQuery 在 Ajax 技术框架中页面加载的应用[J].计算机时代,2011(8):25-26.
- [13] 李志秀,张 军,陈 光,等. JQuery Ajax 异步处理 JSON 数据在项目管理系统中的应用[J].云南大学学报(自然科学版),2011,33(S2):247-250.
- 系统的设计与实现[J].测绘与空间地理信息,2008,31(6):31-33.
- [7] 鲜于丹谦.空间数据挖掘在校园 GIS 中的应用研究[J].计算机与数字工程,2008,36(1):133-136.
- [8] 王占全.基于地理信息系统空间数据挖掘若干技术的研究[D].杭州:浙江大学,2005.
- [9] 李德仁.空间数据挖掘理论与应用[M].北京:科学出版社,2006.
- [10] Wang Xianhua, Miao Zuohu. The fuzzy system technology in geospatial data mining[C]//Proc of international symposium on intelligence computation and applications. Wuhan; [s. n.], 2007.
- [11] Korting T S, Fonseca L M G, Escada M I S, et al. GeoDMA - A novel system for spatial data mining[C]//Proc of ICDMW. Pisa, TBD, Italy; [s. n.], 2008: 975-978.
- [12] 张瑞菊.空间数据挖掘方法及其与 GIS 集成模式的应用研究[D].济南:山东科技大学,2003.
- [13] 贾泽露,刘耀林,张 彤. GIS 与可视化 SDM 技术集成问题探讨[J].南京师范大学学报(工程技术版),2004,4(4):37-42.

基于校园GIS的空间数据挖掘系统原型

作者: [于永玲](#), [李向](#), [宗思生](#), [施进发](#), [YU Yong-ling](#), [LI Xiang](#), [ZONG Si-sheng](#),
[SHI Jin-fa](#)
作者单位: [郑州航空工业管理学院, 河南 郑州, 450015](#)
刊名: [计算机技术与发展](#)

ISTIC

英文刊名: [Computer Technology and Development](#)

年, 卷(期): 2013(12)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjfz201312038.aspx