

数据挖掘在《大学物理》网络教学中的应用

黄 勇, 李玉华

(华中科技大学, 湖北 武汉 430074)

摘 要:随着互联网应用的发展,学习过程被越来越多地放到网络上。但目前现有的网络课程大多模式单一,对不同的用户提供的是完全相同的学习内容,为了更好地实现因材施教,文中以《大学物理》网络教学平台的构建为例,提出使用数据挖掘技术,满足个性化学习的需求,构建个性化网络教学平台的解决方案。

关键词:大学物理;个性化学习;网络教学平台;数据挖掘

中图分类号:G434

文献标识码:A

文章编号:1673-629X(2006)11-0175-04

A Practical Study on Data Mining in Individualized Network Teaching of College Physics

HUANG Yong, LI Yu-hua

(Huazhong University of Science & Technology, Wuhan 430074, China)

Abstract: With the development of Internet, more and more learning can be done on Internet, but the modes of the available network courses are sole, providing different learners with indifferent learning materials. In order to offer individual ways to different learners, puts forward a solution of data mining to construct the network teaching of College Physics to meet the needs of individualized learning.

Key words: college physics; individualized learning; network teaching platform; data mining

0 引 言

在2000年5月,教育部启动了“新世纪网络课程建设工程”,立项320个,随着计算机辅助教学技术与计算机网络技术的飞速发展,校园网的普及为各高校开展网络教学提供了可能,而随着各高校这几年不断地扩招,在校学生与教师的人数比例越来越大,师资趋于紧张,像物理、数学等文化基础课常常采用大班教学,而教学效果与小班教学相比有一定差距,因此,开发网络教学平台不失为一个解决问题的良策。文中拟就目前《大学物理》网络教学平台的现状进行分析,并对今后的开发提出建议。

1 《大学物理》网络教学平台的现状

很多高校都设置了《大学物理》课程教学网站,现在打开 www.baidu.com 输入“《大学物理》网络课程”进行搜索,可找到相关网页有6400多个,输入“《大学物理》精品课程”进行搜索,可捞到相关网页多达10900个,其中比较有名的有:东南大学、装备指挥技术学院研制的《大学物理》网络课程(<http://www.jswl.cn/course/dxwl/>),西南交通大学《大学物理》网上学习系统(<http://bbs.xnjd.cn:82/courseware/physics/>)等。综合以上网站课程情况可以发现,普遍都设有以下栏目:教学内容(以章节为线开展教学,带有导航可在内容间选择性跳转,配有Flash动画进行形象演示并伴有声音解说);复习指南(设有基本要求、基本内容、问题讨论、典型例题、自我检测等栏目);原理应用(介绍物理知识在各领域的实际应用);物理实验(设实验原理、实验仪器、实验内容、数据处理、思考问答等栏目,使用动画生动形象地演示仪器结构、使用方法,实验过程能在网上用鼠标进行模拟操作有一定智能性,能够出错提示);物理习题(给出各章习题答案);物理进展(最新物理前沿知识、应用);在线讨论(在线答疑BBS);帮助等。有的网站还设有视频点播栏目(可点播教学与学术讲座、演示实验等录像资料),从以上分析可以看出,现有的《大学物理》网络课程充分发挥了计算机网络的优点和多媒体先进的表现手段,精心设计制作了大量图片和动画,采用音频、视频、动画、演示、仿真试验的多种教学内容呈现形式,形象生动、色彩优美,营造一种良好的学习环境和气氛,具有科学性和趣味性,设置的“在线讨论”等类似栏目能让教师与用户、用户与用户之间进行学习交流,增加了网络课程的交互性。

2 《大学物理》网络教学平台存在的问题

按照金伯格(Ginsburg, 1998)等人的分法,网络课程可以分为三代:第一代是通过网页给学习者提供教学材料

收稿日期:2006-02-08

作者简介:黄 勇(1970-),男,湖南祁阳人,硕士研究生,高级讲师,研究方向为数据挖掘与计算机辅助教学;李玉华,副教授,硕士生导师,研究方向为计算机网络及应用、人工智能。

和有关资料,以及与其他有关教育网连接;第二代是除了在网上提供学习材料外,还要求学习者通过电子邮件、电子公告栏、网上练习和测试进行异步双向交流;第三代是除了第一代、第二代外,还要求通过网上交谈室、电话会议、视频会议或 MUDs (MOOs) 系统进行同步双向交流^[1]。当前世界网络课程已正向第三代发展。国内现有的《大学物理》网络课程大部分处于第二代,有一些还处于第一代。基于网络的《大学物理》课程教学的效果也并不是很理想。主要存在以下问题:

(1) 网络教学平台智能性低、缺乏个性化设计、模式单一。

调查显示当前国内教育机构往往集中极大的精力去设计开发完善课程材料,对网络教学理论研究不够系统,相当一部分网络教学内容的讲解与呈现仍以大量的文字阅读为主,其主要教学内容组织形式可归纳为:文字+图片+表格+多媒体动画(多以 Flash 为主)+视频(流媒体、其它格式)+教师的电子讲稿(PPT)^[2],整个系统大多采取“自学课件+网上讨论答疑”的形式进行教学活动,是一种完全的“学习者驱动”的静态超文本系统,网络课件尽管也能做到图文声并茂,但缺乏互动性,学习者不能主动参与,无法实现师生之间的交流,只是将传统课堂教学简单移植到网络上。

非智能化的网络课件不能体现教与学的个性,不能根据学习者个性提供相应的指导;不同的学习者只能被动地接受完全相同的学习内容,而不能主动地根据自己的需要选择、调度、控制自己的学习过程,更不可能实现因材施教;从《大学物理》网络课程的现有成果看,还谈不上综合应用现代信息技术,及对它们进行智能化的研究、开发和实践。

(2) 缺少有效、敏感的响应与反馈、测试与评价系统。

评价与反馈是教学中不可或缺的环节。它能诊断学习者的学习成就和缺陷,并通过对成就和缺陷的评价,使学习者产生激励感、成就感和紧迫感从而形成学习动机。目前国内《大学物理》网络课程教学平台对学习者的应该达到的状态想的比较多,对学习者的现状的分析比较弱^[3]。学习者大多通过实时或者非实时的交流讨论系统和老师或者其他同学进行交流、完成教师布置的作业之后通过 E-mail 上交给教师或者上传到系统、参加基于 WEB 的在线测试练习来考察自己对所学知识的掌握情况等等。这种教学模式没有对教师提供教学效果的反馈,教师只能从学习者的提问、讨论和测试成绩,从直觉上根据经验感性地判断所有学习者的知识掌握情况和每个学习者的学习状况,并根据判断的结果制定教学策略和组织教学活动。这样,学习者的学习差距和存在的问题就不是非常清晰,也就不能很好地对症下药、解决问题。

据调查显示,在影响用户网上自主学习效果的因素中,42% 的用户认为学习监控机制是最主要的因素^[4]。目前《大学物理》网络教学的评价方式仅限于教师对学习者的

评价(作业和考试),很少考虑过程性评价。如学习者自身的评价、同伴的评价,而这两种对于学习者获得学习成就感非常有效。对于考试,网络教学平台往往是仅给出答题结果,而没有给出适当的分析、评语,没有学习诊断,没有根据学习者的问题、测试成绩给学习者指出没有掌握的知识点和学习建议并且提供给学习者针对没有掌握的知识点的补充资源,不便学习者尽快掌握。

同时,所提供的练习、作业和测试系统的题目多是事先放在 Web 页面上固定不变的,不能根据不同能力学习者的要求和学习情况自动选题组卷、变化题目,精确地对每一个学习者进行测试和评价。

(3) 导航系统不强。

目前大部分《大学物理》网上教学系统都采用超文本组织教学材料,学习者自己掌握学习的主动权,没有建立帮助学习者了解学习该课程所需要的知识水平、自己的知识层次、学习进度和学习方法的系统;不能根据学习者对知识点的理解程度及学习者已学习过的资源记录,给出学习者应学习的知识点及资源,对学习者的进一步学习起到导航的作用;学习单元之间的切换只能靠前进、后退或从头开始实现,不便于跳跃性学习;学习者只能靠记忆来确定自己的学习位置,不能记录学习者的学习路径、学习心得,学习者一旦中断学习进程,只能重新从头开始进入,容易造成学习者学习随意性大,甚至盲目地漫游,造成迷航。

3 数据挖掘技术在《大学物理》网络课程中的研究与应用

以下针对目前《大学物理》网络课程中存在的不足,说明如何把数据挖掘技术应用于《大学物理》网络课程的设计当中,将其设计成一个以学习者为中心的个性化智能教学系统。

3.1 《大学物理》网络教学平台构建模型

根据个性化教学的设计思想,《大学物理》网络教学平台模型如图 1 所示^[5]。

整个教学系统包括教学资源库和用户数据库。教学资源库具体包括:知识点库(记录着课程中的所有知识点及知识点之间的关联规则,如:并列、前提、后继等关系)、课件库、作业库、试卷库、答疑库等必需的教学资源。用户数据库则记录着用户的基本信息、学习进展、对知识的掌握(学习成绩、作业答疑)情况等学习者个性特征。

整个教学系统中最核心的部分是数据挖掘系统,由用户信息收集、个性化分析处理、智能重组 3 个模块组成。

用户信息收集(数据收集与预处理模块)模块:负责收集用户请求,跟踪每一个用户的学习过程并记录其相关信息(如学习历史、做过的练习和测试题、答疑情况、用户现有知识结构、知识点的掌握程度等)。用户的个性化特征信息是实施个性化教学策略的基础,用户数据库中存放的用户个性特征数据经过数据预处理,为个性化分析器提供实施个性化教学策略的数据来源。

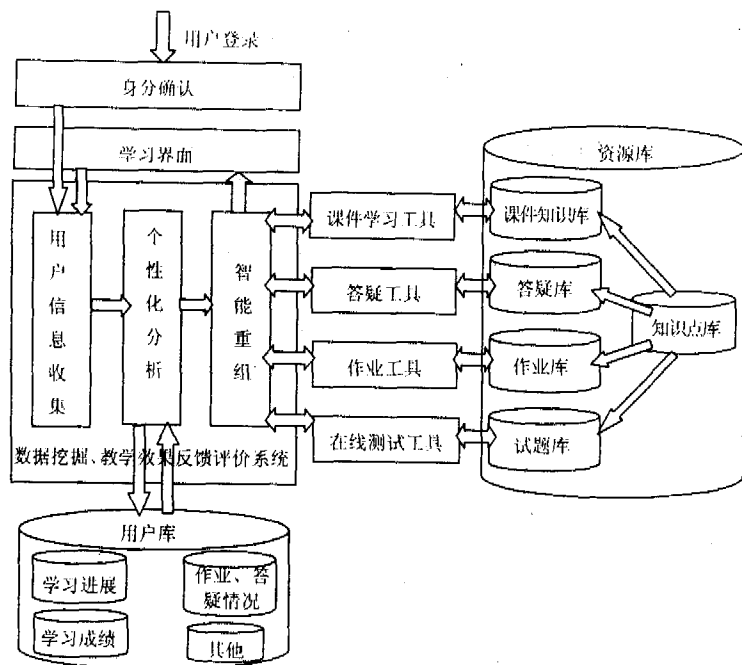


图 1 《大学物理》网上个性化智能教学系统

个性化分析模块运用数据挖掘等技术对用户数据库中记录的个性化信息进行分析,得出适合于用户当前学习的教学策略信息,提供给智能重组模块由其调用资源库中相应学习资源生成适合用户的课件、作业以及测试提供给用户,也可将对用户的学习评价及建议直接反馈给用户督促用户调节自己的学习。同时将分析所得到的个性化结果对用户数据库进行动态更新。比如说,如果一个用户在作业中在某个知识点方面出错较多,个性化分析模块就会及时发现这一现象,并记录到用户数据库中,在用户下次登录时,个性化分析模块将根据这一个性化信息通知智能重组模块对该用户就此知识点着重进行教学辅导和相关作业题的练习。

3.2 实现环境与关键技术

整个开发环境可采用 Windows 2000, SQL Server 2000, ASP.NET 进行开发设计,采用 DW + OLAP + DM → DSS 的解决方案,其中关键性的技术为数据挖掘技术。数据挖掘 (Data Mining) 就是应用一系列技术从大型数据库或数据仓库的数据 (多为不完全的、有噪声的、模糊的和随机的) 中,提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程,它又被称为知识发现 (KDD, Knowledge discovery in database)、决策支持等,所提取的知识表示为概念、规则、规律和模式等形式。

数据仓库 (DW) 用于数据的存储和组织,面向数据仓库的联机分析技术 (OLAP, On-Line Analytical Processing, 一种基于多维数据模型的数据分析和挖掘工具) 集中于数据的分析,数据挖掘则致力于知识的自动发现。由于这 3 种技术内在的联系性和互补性,它们结合起来即是一种新的 DSS (Decision Support System) 构架。

数据挖掘过程如图 2 所示:由数据预处理、数据挖掘和结果分析与运用几个主要阶段组成。

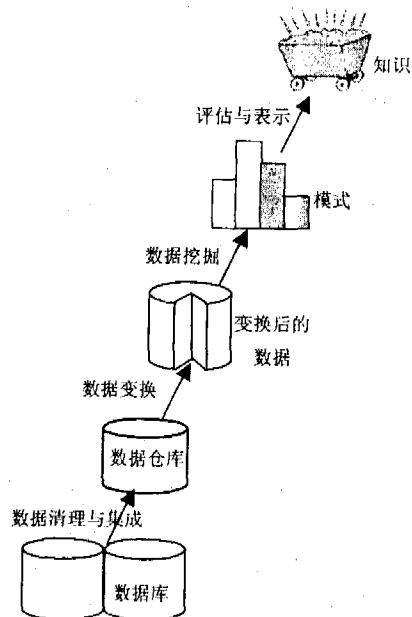


图 2 数据挖掘过程

(1) 数据预处理。

数据预处理包括数据清理、数据集成和变换、数据归约。数据清理可以去掉数据中的噪声,纠正不一致;数据集成是将多文件或多数据库运行环境中的数据进行合并处理,解决语义模糊性,处理数据中的遗漏和清洗数据等;数据归约可以通过聚集、删除冗余特性或聚类等方法来压缩数据,从而建立适应数据挖掘的目标数据集——学习者个人学习记录数据仓库,这样可以提高数据挖掘模式的质量,降低实际数据挖掘所需要的时间。

经过数据收集和预处理模块对学习者的信息进行收集,收集的基本信息包括:学习者很必要的个人信息、学习课件的信息、答疑提问的信息、作业完成情况的信息、在线测试的信息、在线学习停留时间的信息等,经预处理后提交数据挖掘阶段进行分析。

(2) 数据挖掘。

数据挖掘阶段进行实际的挖掘操作,即利用机器学习、统计分析等方法,从数据库中发现有用的模式或知识 (这里,模式是浓缩数据的信息形式,如精炼数据库、表格、产生式规则、决策树、神经网络的权值等)^[6]。

①选择数据挖掘方法。如统计分析、机器学习、模式识别方法和人工神经元方法等。

②选择数据挖掘算法。选择用来查找模式或符合数据的模型的算法,确定合适的模型和参数。数据挖掘可用到的算法有统计分析、一般分类和聚类。可以通过统计得到学习者的学习时长、学习次数、学习平均水平、提问次数及主要问题等;聚类是把一组组个体按照相似性归并成若干类别;分类将一组组个体分门别类地归入预先设定好的几个类中,分类的方法有:决策树、贝叶斯算法、K 近邻分类器等。通过分类和聚类的方法把学习者分成不同的类别,以提供相对应的个性化教学。

③数据挖掘。查找感兴趣的模式。模式一般表示为

一种特殊的形式或一套表达方式,如关联规则、分类规则或分类树、回归结构和聚类集等。

(3) 数据挖掘、结果分析表述和挖掘应用。

此阶段运用使用兴趣度度量,并与数据挖掘模块交互,以便将搜索聚焦在有趣的模式上。它可能使用兴趣度阈值过滤发现的模式。运用统计学和关联规则等方法,把挖掘分析的结果放入一个个性化数据库,当学习者下次进入系统时,系统就可根据个性化数据库提供其符合学习需求的页面。

4 结束语

网络教学平台的关键是针对用户的个性特征信息,通过系统的分析和判断,给予不同的学习环境和学习内容的呈现,通过运用数据挖掘技术可以从用户数据库及用户学习行为记录中挖掘出用户对知识点的理解程度,从而实现在学习过程中对用户的学习进行记录、指导、反馈,对用户

选择的学习策略给予支持,大幅提高《大学物理》网络教学平台的教学效果,使个性化教学真正得以实现。

参考文献:

- [1] 林君芬,余胜泉. 关于我国网络课程现状与问题的思考[J]. 教育技术通讯,2001(1):55-59.
- [2] 梁林梅,焦建利. 我国网络课程现状的调查分析与反思[J]. 开放教育研究,2002(6):13-16.
- [3] 刘莉. 远程学习者研究现状及发展趋势——远程教育专家访谈录[J]. 中国远程教育,2003(5):7-12.
- [4] 黄萍. 高校学生网络自主学习行为的调查研究[J]. 开放教育研究,2004(6):77-80.
- [5] 舒蓓,申瑞民,王加俊. 个性化的远程学习模型[J]. 计算机工程与应用,2001(9):90-92.
- [6] 康晓东. 基于数据仓库的数据挖掘技术[M]. 北京:机械工业出版社,2004.

(上接第 155 页)

法的特有功能,今后可以继续把新的数据挖掘算法引入到入侵检测中,改善检测的准确性、可靠性。

参考文献:

- [1] Lee Wenke, Stolfo S J, Mok K W. A Data Mining Framework for Building Intrusion Detection Models[C]//Proceedings of the 1999 IEEE Symposium on Security and Privacy. Los Alamitos, CA: IEEE Computer Society Press, 1999:120-132.
- [2] Wong M L, Leung K S. An Efficient Data Mining Method for Learning Bayesian Networks Using an Evolutionary Algorithm - Based Hybrid Approach[J]. IEEE Transactions on Evolu-

tionary Computation, 2004, 8(4):378-404.

- [3] 张琨,徐永红,王珩,等. 用于入侵检测的贝叶斯网络[J]. 小型微型计算机系统,2003,24(5):913-915.
- [4] Kruegel C, Mutz D, Robertson W, et al. Bayesian Event Classification for Intrusion Detection[C]//Proceedings 19th Annual Computer Security Applications Conference. Los Alamitos, CA: IEEE Computer Society Press, 2003:14-23.
- [5] 白耀辉,陈明,王举群. 利用朴素贝叶斯方法实现异常检测[J]. 计算机工程与应用,2005(34):131-132.
- [6] 牛建强,曹元大,阎惠. 基于数据挖掘的 CIDE 协同交换[J]. 计算机工程,2003,29(14):35-36.

(上接第 174 页)

代的地位与影响。信息技术极大地推动了图书馆的现代化进程,同时也带来了信息的爆炸式增长。在知识经济时代,解决好海量信息的存储、检索、开发与利用,是关系到图书馆未来的生存与发展的重大问题。因此必须制定国家数字图书馆发展战略,做好整体建设规划。同时还要制定和完善数字图书馆相关的保密、版权等方面的法律。

7 结束语

数据挖掘技术及其应用是目前国际上的一个研究热点,在数字图书馆领域,面对大量的信息,用数据挖掘技术找出数字图书馆用户感兴趣的信息加以组织利用,加强客户关系的管理,提高满意度。基于数字图书馆建设与应用数据挖掘的有利时期已经到来,优质的网络信息服务事业前景不可限量。综合应用数据挖掘技术和人工智能技术,获取用户知识、文献知识等各类知识,将是实现知识检索和知识管理发展的必经之路。

参考文献:

- [1] Poe V. Building a Data Warehouse for Decision Support[M]. [s.l.]: Prentice PTR, Prentice-Hall Inc, 1996.
- [2] 赵洗尘. 数字图书馆及其建设[J]. 现代图书情报技术, 1999(1): 28-31.
- [3] 刘霞. 关于数字图书馆建设的几个问题[J]. 图书情报知识, 1998(1): 30-32.
- [4] 王珊. 数据仓库技术和联机分析处理[M]. 北京: 科学出版社, 1998.
- [5] 刘海虹, 刘伯莹. 数据挖掘技术[J]. 丹东纺专学报, 2001(1): 15-18.
- [6] 郝先臣. 数据挖掘工具和应用中的问题[J]. 东北大学学报: 自然科学版, 2001(2): 183-187.
- [7] 卢增祥. Bookmark - 智能化网络信息服务系统[J]. 高技术通讯, 1999(6): 30-32.
- [8] 郑巧英, 杨宗英. 图书馆自动化新论——信息管理自动化[M]. 上海: 上海交通大学出版社, 1998.