

文章编号: 1005-8451 (2010) 06-0055-04

## 基于校园一卡通平台的数据挖掘应用研究

李珊娜

(北京交通大学 信息中心, 北京 100044)

**摘 要:** 本文在研究校园一卡通平台的基础上, 以我校设计和研发的学生行为分析系统和就餐消费分析系统为例, 介绍了对校园一卡通系统的数据挖掘的初步探索, 不仅为领导决策提供了有用的参考信息, 也为将来构建完整的数字化校园决策支持系统提供了实践经验和实现方法。

**关键词:** 校园一卡通; 数据挖掘; 学生行为分析; 就餐消费分析

**中图分类号:** TP39

**文献标识码:** A

### Application study of data mining based on campus card platform

LI Shan-na

(Information Center, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** This paper gave a brief introduction of primary exploring in data mining based on the campus card platform, and the Student-behavior-analysis System and Eating-behavior-analysis System designed and implemented by our university as an example of data mining, which not only provided useful reference information for the leader making decisions, but also provided practical experiences and implementation method to construct Complete Decisions Support System of digital campus.

**Key words:** campus card; data mining; student-behavior-analysis; eating-behavior-analysis

目前校园一卡通在许多学校的应用不断增加, 随之产生的是更为丰富而庞大的一卡通数据信息, 这些数据信息包含着校园卡中许多不为人知的特点。

对校园一卡通系统中累积起来的数据进行统计、分析、挖掘, 就可以对所有一卡通用户的消费、身份认证等活动的整体状况、规律等有一个量化的概念, 对某些具体的消费行为也可以有一个较精准的判断, 这不仅可以使今后的一卡通建设有可参照的标准, 还使得校园一卡通系统可以满足从业务处理, 到中层管理的控制, 最终达到为领导决策提供支持的目的。

校园卡仅仅一年的交易数据就有近亿条记录, 文件数据量超过 10 Gbit 的大小, 如此庞大的数据对挖掘算法的复杂度、计算机性能都要求很高。本文就是在研究校园一卡通平台的基础上, 介绍目前我校设计与实现的在校园一卡通平台上进行数据挖掘的两个应用案例, 即学生行为分析系统和就餐消费分析系统。初步探索如何对一卡通海量数据进行深层次的挖掘, 为领导决策提供有价值的参考信息。

### 1 校园一卡通平台架构

下面以北京交通大学校园一卡通平台为例, 介绍校园一卡通平台架构。校园一卡通平台采用当前业界流行的三层系统架构, 即用户层、中间层和数据层, 整个系统具有良好的开放性和可扩充性, 能实现客户逻辑、业务逻辑和数据逻辑的有效分离和耦合, 一个健壮的、可扩充的、具有相当安全性的校园信息化平台架构如图 1。

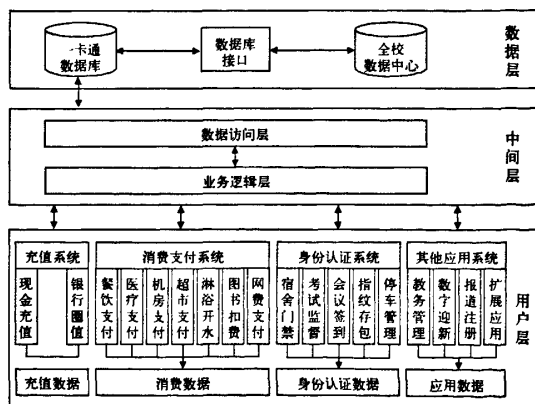


图1 校园一卡通平台架构

我校一卡通系统主要功能包括充值系统、支

收稿日期: 2009-10-12

作者简介: 李珊娜, 工程师。

付系统、身份认证系统、其他应用系统等。(1) 充值系统。目前提供现金充值和银行圈存两种方式实现对校园一卡通的充值。一卡通持卡人同时拥有校园一卡通和银行卡两张卡, 持卡人可以通过圈存系统将银行卡上的钱自助转帐到校园一卡通中。(2) 消费支付系统。一卡通具有电子钱包功能, 持卡人可以在校内任何一个与系统联网的消费网点持卡购物消费, 实现货币结算的电子化、数字化。(3) 身份认证功能。校园一卡通上不仅有本人的彩色照片, 还存有持卡人的姓名、学号、身份证号、所属院系、指纹等信息, 可以代替其它书面证件进行电子身份识别。(4) 与校内其它管理系统对接, 实现信息资源共享。教务管理系统、财务管理系统和图书管理系统等都可以利用平台预留的扩展接口与一卡通系统联网, 实现数据共享, 极大地提高管理效率。

我校根据校内各部门需求, 进一步确立一卡通数据挖掘的目标, 在现有一卡通平台的基础上设计并研发了一卡通学生行为分析系统和就餐消费分析系统, 主要是对于一卡通系统大量原始数据进行提取、加工处理、统计分析, 并得到最终的学生行为数据和就餐消费数据, 为校内各部门提供更有价值的数据依据。

## 2 学生行为分析系统

### 2.1 系统设计目标

通过对校园卡交易及身份认证数据的分析, 我们希望得到用户消费能力和身份认证的分析结果并进行预测。将用户在消费类别、消费日期、消费金额上进行聚类分析, 并利用结果结合用户资料对在校学生的相关家庭情况进行分析, 发现用户的消费习惯和潜在的消费规律, 从而可以通过这些数据为助学贷款发放的审查等工作提供数据依据。

### 2.2 系统设计

系统主要包括查询、统计、提醒3大功能。

查询功能: 可以选择对特定人群(如某一个学院、某一个年级、或者某一名学生)的充值记录、消费记录、就餐记录、上机记录、门禁记录、医疗记录和上网时长、上网流量等各种行为进行查询。

统计功能: 可以选择对一个群体(如某一个

学院、某一个年级、或者某一名学生)的充值记录、消费记录、就餐记录、上机记录、门禁记录、医疗记录、上网记录进行统计分析, 显示出相应的饼图、柱状图。图2是根据我校06级本科生的2009年10月的就餐消费情况而分析出的就餐情况分类统计图。

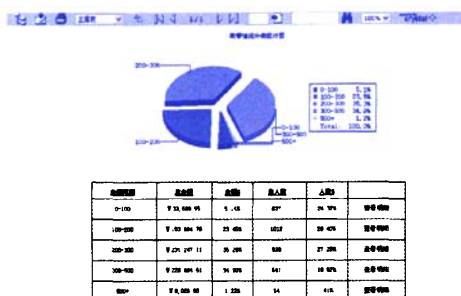


图2 就餐情况分类统计图

提醒窗口: 对于每一类行为进行分析, 将分析结果存入提醒表, 当用户进入学生行为分析系统后, 系统将自动弹出主动提醒窗口。

系统的数据处理流程是从一卡通数据库中提取原始数据到行为分析数据库, 再对行为分析数据库中的数据加工处理, 对加工过的数据按行为类别统计, 通过校园信息管理系统发布给用户。一卡通数据挖掘的流程图如图3。

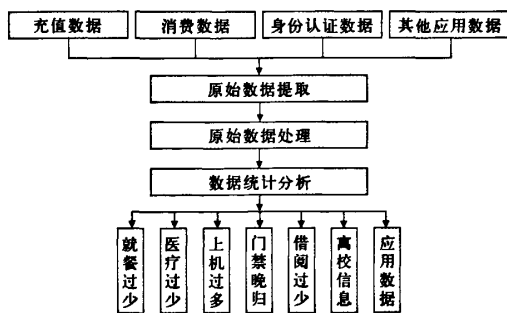


图3 一卡通数据挖掘的流程图

通过对校园卡交易数据的分析结果可以看出, 经过系统中一定条件的筛选和处理, 得到就餐过少、医疗过多、上机过多、门禁晚归的学生行为信息, 为各部门负责学生工作的教师提供了真实客观的依据, 便于做出合理正确的决策。例如: 对于就餐次数正常, 但是每餐消费金额过少的学生, 可

以为助学贷款发放的审查等工作提供辅助决策,这远比让学生写出书面申请,再去深入调查方便和客观的多。

### 2.3 几个关键问题的实现

在整个系统设计和实现过程中,对一卡通数据库中海量数据的提取以及对行为分析系统的响应性能都提出了很高的要求,我们对这些问题深入研究,并依次采取了相应的解决办法。

#### (1) 自动提取原始数据

由于校园一卡通应用子系统繁多,子系统产生的应用数据量庞大,系统主要是利用每日的热备份数据文件和归档日志,在一卡通备份库上定期执行自动恢复脚本作为学生行为分析数据库,实现一卡通数据的提取,确保一卡通生产库的性能不因数据提取受任何影响。

#### (2) 优化行为数据库

在行为分析数据库的表中建立基于学号、行为类别、操作日期的索引,大大的提高了查询速度。除此之外,还可以对原始数据进行加工处理,将刷卡记录与具体行为进行一一对应,把规定时段内多次刷卡次数统一为一次行为,确保行为分析的准确和客观。

#### (3) 提高系统响应速度

行为数据统计通常耗费大量的时间,当用户查看统计结果时再对数据进行统计会大大降低系统响应速度,有效的解决办法是使用PL/SQL编写存储过程,设置定期执行该过程,这样数据库就会利用空闲时间自动以每个持卡用户为单位,按月分类统计行为数据。当用户查询统计数据时,系统只需要查询按月分类统计结果,显著提高了系统的响应速度。

## 3 就餐消费分析系统

### 3.1 设计目标

系统将以原始就餐消费数据为基础,以各个食堂工作站为方向对其进行统计分析,计算各食堂每时每刻的就餐人数,自动产生当天的就餐曲线,分析就餐消费人数的峰值与低谷、各个食堂工作站就餐消费的对比情况等,并且把统计分析结果以报表和图表的方式直观地显示出来,为相关经营单位的生产经营提供一个比较客观的可参照

数据,有利于经营中的效率提高,做到既不浪费资源,也不会出现供不应求的情况,同时也辅助校领导对食堂工作站的有效管理做出决策。

### 3.2 系统设计与实现

系统包括4个功能:单日消费分析;双日消费对比;周消费分析;单日食堂对比分析。

系统实现架构主要分为后台数据处理模块、数据传输模块和数据整理模块3个模块。

后台数据处理模块:将一卡通生产库中的数据每天定时提取并分析后转存入行为分析数据库中的就餐数据分析表,后台数据处理模块设计为Windows系统服务形式,可以开机自动启动,后台隐藏运行,数据处理过程以日志的形式写入日志文件,不影响服务器其他应用。

数据传输模块:主要为数据整理模块提供格式化的源数据,将就餐数据分析表内的就餐数据按所需日期选择出来,以所需格式打包转发给数据整理模块,以便数据整理模块对数据进行进一步处理。从而避免数据整理模块直接和数据库交互,如果数据整理模块直接与数据库交互,则不得不为其添加庞大的Oracle客户端程序。

数据整理模块:数据整理模块根据选择的模板和日期向数据传输模块请求数据,并将接收到的数据输出到Excel表格模板,使用程序提供的各种模板来对就餐行为进行分析。系统设计框图如图4。

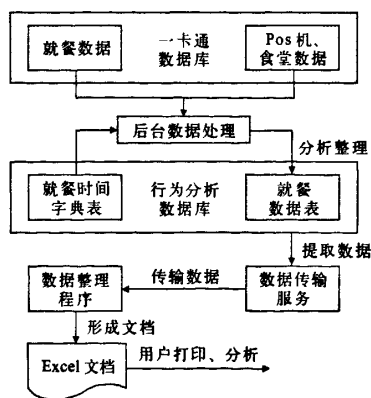


图4 系统设计框图

## 4 结束语

校园一卡通系统是数字化校园建设重要、有

文章编号: 1005-8451 (2010) 06-0058-04

## VPN 技术在站段网络建设中的应用

张 辉, 李菊萍

(西安铁路局 宝鸡工务段网络室, 宝鸡 721000)

**摘 要:** 本文首先对 VPN 的需求进行分析, 介绍了 VPN 技术和隧道技术, 提出铁路段办公网与车间班组互联方案, 采用 VPN 技术为段办公网的安全提供技术保障。

**关键词:** VPN 技术; 安全; 办公网应用

**中图分类号:** TP393

**文献标识码:** A

宝鸡工务段是西安铁路局工务系统内最大的站段之一, 管辖西宝线、宝中线、宝成线及宝天线。除机关办公使用的计算机外, 其余计算机安装在沿线各车间和班组, 分布范围较散, 且相隔距离较远。

目前段机关通过网桥接入方式接入车站机房。车间班组通过铁通的综合信息平台与段机关联通, 车间班组需要实时将信息传送到段机关及铁路局, 段机关及铁路局也需要将反馈的信息实时向车间班组下发。但段机关作为连接的中心点, 所有的数据都经过段服务器进行分发。

### 1 VPN 需求分析

随着铁路的迅速发展, 工务段管辖的车间班组相继多了起来, 信息交互也越来越频繁, 随着各种业务应用系统的实施, 重要的数据和信息在网络中传输也越来越多, 安全性要求也越来越高, 目

前仅仅依靠综合信息平台接入, 段服务器转发的组网模式已经越来越不适应工务段对信息传输平台的要求了。

从安全方面考虑, 铁通提供的综合信息平台没有经过加密处理, 重要数据和信息均是以明文在网上传输, 如果别有用心的人利用 Sniffer 等网络监听分析工具, 极易篡改、窃取甚至破坏企业数据, 给企业造成不可估量的损失; 由于传输平台没有认证功能, 企业内部员工的越权访问、误操作、有意或无意的泄密、甚至是少数员工恶意的破坏, 都会对企业的信息和数据造成很大的威胁; 由于传输平台没有访问控制和安全隔离的功能, 给外部非法人员提供了入侵的机会, 非法人员可以通过专用的黑客程序 (此类工具在 Internet 上可以任意下载), 或者盗取授权员工的访问权限, 很容易进入企业系统内部。由于工务段车间班组联网网点数目众多, 相应的受攻击的几率较大, 一旦通过计算机终端进入总部服务器, 后果将不堪设想。

从管理方面考虑, 工务段处于高速发展阶段,

收稿日期: 2009-11-13

作者简介: 张 辉, 技术员; 李菊萍, 助理工程师。

机的组成部分, 是推行数字化校园的有力载体, 其内容将随需求愈加充实, 其构成亦愈加丰富, 它可以渗透到学校的教学、科研、管理和后勤等多方面应用, 并随着科技进步和管理内容的增多而加以深化。我校在校园一卡通系统建设的成功基础上, 通过对校园一卡通系统中累积起来的各种数据进行统计分析, 得到一系列学生行为的分析结果和就餐消费分析结果, 不仅为领导决策提供了有用的参考信息, 也为将来构建完整的数字化校园决策支持系统提供了实践经验和实现方法。

### 参考文献:

- [1] 张敬涛, 李向阳, 邹秀香. 校园一卡通系统的应用研究[J]. 山东师范大学学报 (自然科学版), 2009 (23): 126-129.
- [2] 张升平. 数字化校园之校园一卡通的建设[J]. 重庆工商大学学报 (自然科学版), 2008 (25): 56-59.
- [3] 段智敏, 王如龙, 孙美青, 余 维. 基于一卡通的数字化校园资源整合研究与实现[J]. 计算机工程与科学, 2008 (30): 8-11.
- [4] 王玉芬, 张治斌, 李长江. 数据仓库在高校决策支持中的应用研究[J]. 陕西理工学院学报, 2007 (23): 17-19.