

智能化交管综合业务管理系统^①

杨风召, 刘 军

(南京财经大学 电子商务系, 南京 210003)

(江苏省电子商务重点实验室, 南京 210003)

摘要: 交管信息的综合化和智能化是我国交通运输系统发展的趋势。在分析交通管理系统现状和存在问题的基础上, 提出了一个基于商务智能体系的交管综合业务管理系统总体框架, 并就其中的数据仓库设计、异构数据集成、联机分析处理和数据挖掘等关键问题给出了相应的设计方案。该设计方案已经在某省级交管系统中得到了应用。

关键词: 商务智能; 交通管理信息系统; 数据仓库; 数据挖掘

BI-Based Integrated Traffic Information System

YANG Feng-Zhao, LIU Jun

(E-Business Department, Nanjing University of Finance & Economics, Nanjing 210003, China)

(Jiangsu Key Laboratory of E-Business, Nanjing 210003, China)

Abstract: The integration and intelligence is the trend of China's Traffic Information System (TIS). Based on the analysis of the current situation of TIS and existing problems in it, the paper puts forward a framework of BI-based Integrated TIS, and design scheme of the data warehouse, data integration, on-line analytical processing and data mining. This design has been applied in a provincial TIS.

Key words: business intelligence; traffic information system; data warehouse; data mining

1 引言

随着交通需求和信息技术的迅猛发展, 交通基础设施建设、交通运营管理、运输企业管理等交通相关领域的信息化已取得重要进展, 各种交通信息系统的陆续实施已经给交通管理者和出行者带来明显的便利, 创造了一定的经济和社会价值^[1-3]。目前多数城市已经建立起各种交通信息管理系统^[4-6], 这些系统已经实现了网内的纵向数据共享。但是各个部门内部的横向数据并没有达到信息共享的要求, 大量的数据查询是集中在单独的业务系统当中。而从实际数据应用的角度来看, 对于某一事件的查询要求往往是跨业务的, 需要各方面的数据来综合分析。在现有的方式下, 很难去综合信息, 发现规律, 找出问题。

本文设计方案利用当今先进的商务智能技术将隔离在各个应用系统之中业务数据整合起来, 实现对交

通管理信息的整合、挖掘、分析、研判。深层次地提炼、挖掘交通管理信息资源, 为交通管理和决策提供科学依据, 推动交管信息化建设向更高层次发展。

2 交通管理系统存在的主要问题

以前建立的交通管理系统存在的主要问题如下:

(1) 报表功能难以满足管理决策的需要。以前建立的交通管理系统的报表反映的是各个地方局基本的业务状况, 不能按照要求进行定制。

(2) 数据关联度不高。由于各业务系统建构的时间和客观状况不一, 导致了各业务系统数据互相独立, 缺乏一种有效的关联机制将这些数据关联起来。

(3) 缺乏动态分析。目前所提供的数据应用主要是集中在报表与查询上。报表本身是一种静态的数据, 无法实现人机的交互分析。而查询虽然可以提供多种

① 基金项目:国家自然科学基金(71072172);留学人员科技活动择优资助项目(YFZ302002);江苏高校优势学科建设工程资助项目

收稿时间:2011-04-22;收到修改稿时间:2011-06-01

查询形式，但是业务人员所有的查询内容都是集中在一个层面上，无法动态实现信息的宏观汇总及细节分析。

(4) 信息分发不及时。按照现有的模式，业务部门的数据往往依赖于信息人员的提供。存在一个从业务部门提交需求→信息部门加工数据→数据交付业务部门这样一个过程。很显然，对于较为复杂的数据要求，这种模式需要花费较多时间。

(5) 业务系统各自独立运作，缺乏数据规范。现有的各业务系统在建立之初考虑的架构、功能基本上以本部门需求为主，并且各业务系统上线的时间也各不相同。因此不同的业务系统对于同样的对象定义可能存在不统一的情况。

(6) 信息分析耗费大量精力。现有的报表和查询功能总体上回答了各业务“怎么样？”的问题，如果需要继续解决“为什么？”的话，交管业务部门需要投入更多的精力在信息加工及分析上。

基于以上现状，为了打破各种业务系统之间、各地市业务数据之间的信息壁垒，为了更好地发挥这些业务数据的效能，必须在保持业务系统不受影响的前提下，打破业务系统之间的分割，从这些数据核心入手，进行后台的数据整合，建设交管综合业务数据分析系统。

3 交管综合业务分析平台体系结构设计

交管综合业务数据分析系统从原有的业务系统中将相关业务数据进行抽取、清洗、加工、整理、加载到数据仓库中，在数据仓库中形成基础的分析数据存贮。总队可以根据自身管理、业务的需要在数据仓库上建立适合自身应用的数据集市。数据仓库、数据集中蕴含的信息可以通过报表、OLAP 分析、即席查询、数据挖掘形式向分析系统使用人员展现。

交通管理综合业务数据分析系统内部结构由数据源层、数据获取层、数据分析层、业务应用层构成，主要包括两大平台，即数据集成分析支撑平台和应用服务平台（如图 1 所示）。

(1) 数据源层

数据源是指目前各类应用中已经形成的数据，如机动车/驾驶人管理信息系统数据，重点机动车/驾驶人数据、道路交通违法管理信息系统数据、道路交通事故统计分析系统数据、道路交通事故 24 小时快报系统数据、剧毒化学品、警车 GPS 数据库等，以及其他非

结构化数据和外来数据源。在必要的时候，还要考虑通过公安部数据查询接口关联公安部数据库。

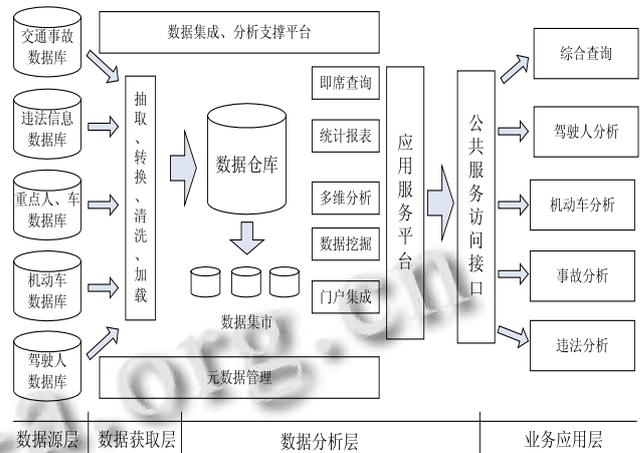


图 1 交管综合分析系统结构示意图

(2) 数据获取层

数据获取层包括三个子层次：源数据缓存库、综合数据库、专题应用数据库。通过工具以及公共应用支撑平台等方式，将综合数据库需要的数据从数据源抽取、转换到源数据缓存库；基于这些源数据，利用相关的数据管理工具进行数据的抽取转换，形成规范化和标准化的综合数据库；基于综合数据库，利用相关的数据管理工具通过映射、整合等进一步加工处理，形成面向不同应用需要的专题应用数据库。

(3) 数据分析层

根据提取的数据，结合确定的数据分析需求，在数据分析平台的支持下，对业务数据进行全方位的分析，从中挖掘有意义的业务规律知识。

(4) 业务应用层

从满足多样化的应用需求出发，以综合数据库为基础，通过应用服务平台开发的通用功能模块和专有功能模块，提供相关的应用服务，如信息查询、统计分析等多种信息服务手段。另一方面，这些功能模块不仅能提供界面访问，而且要在面向服务的架构下，开放多种服务接口，供其他应用系统调用。

(5) 数据集成分析支撑平台

数据集成分析支撑平台是一个由数据获取层和数据分析层的软硬件资源组成的服务环境。功能上包括数据的抽取、转换、清洗、加载、存储，以及数据的查询、多维分析、数据挖掘分析等。

(6) 应用服务平台

应用服务平台是在对各类软件工具进行整合集成

的基础上, 开发的各种应用服务工具组成的工具集, 全面支持数据层、应用服务提供层、应用接入层等各项功能的实现, 提供各种专题应用服务。特别是数据抽取转换装载服务, 要在省厅和省辖市统一部署, 形成统一的数据交换服务, 满足省市数据传输交换的需要。应用服务平台应具备良好的开放性、扩展性, 同时在省级交通管理综合分析系统应用过程中, 能够根据应用需求不断完善和扩充。

4 数据仓库设计

4.1 数据仓库主题的确定

数据仓库是面向在数据模型中已定义好的交管系统的主要主题领域的。在数据仓库中, 主要主题领域是以一组相关的表来具体实现的。每个表用来实现相应主题领域的一部分。一个主题的所有表通过一个公共键码联系起来; 当数据围绕这个主题组织时, 通常每个键码都有一个时间元素。

交管综合业务分析系统的数据仓库包括如下四大主题:

- 驾驶人情况
- 车辆情况
- 事故情况
- 违法情况

主题之间都存在相应的关系。同时, 每个主题也都包含若干实体, 实体间通过键码相联系, 形成星形模型或雪花模型。

4.2 异构数据的集成

按照不同的实现技术, 集成方法可以分为三类。第一类是数据转换方法, 这种方法实现的是企业数据的松散集成, 通过转换工具实现应用系统之间的数据转换和交换, 从而达到集成的目的, 是一种较低层次的集成; 第二类为数据集成方法, 在各种异构数据源的基础上, 借助于中间件系统构造一个虚拟的全局数据模式, 是一种集中式管理、分布式存储的较高层次的集成模式; 第三类是析取、转换和装载(Extract、Transform and Load, ETL)方法, 通过对异构数据源中的数据进行分析、转换和装载, 建立一个数据仓库, 是一种面向企业决策的数据集成方法。

由于交管综合业务数据分析系统的数据来源于目前交管系统中各类应用已经形成的数据, 如机动车/驾驶人管理信息系统数据, 重点机动车/驾驶人数据、道

路交通违法管理信息系统数据、道路交通事故统计分析系统数据、道路交通事故 24 小时快报系统数据、剧毒化学品、警车 GPS 数据库等, 以及其他非结构化数据和外来数据源。通过对以上三种数据集成方法的对比分析, 结合交管综合业务数据分析系统的实际, 本系统采用 ETL 方法进行数据的集成。

5 联机分析处理

数据分析系统主要是对业务的历史数据进行统计分析, 通过统计分析的结果, 得到一些规律和知识, 为决策人员提供有力的决策依据, 对外部环境进行更改或者通过人为干涉, 来减少事故或者违法事件的发生。在该系统中采用 OLAP 面向不同的应用需求提供各种数据分析服务, 包括机动车分析、驾驶人分析、重点车辆分析、重点驾驶人分析、交通事故分析、交通违法行为分析等专题应用。在统计分析过程中的提供多种多样的展示方法, 包括数据表格、平面图表、三维立体图表等。

(1) 驾驶人情况分析

驾驶人情况分析主要是从时间、空间、人口统计特征等对驾驶人的分布情况、变化情况和驾驶人增长情况进行全方位分析。

(2) 车辆情况的分析

从时间和空间分布等方面对车辆属性及动态变化情况进行分析。

(3) 事故情况分析

事故情况分析主要是对发生事故的主体及事故发生的相关属性、变化规律、特点进行时空分析, 从而找出事故发生的规律, 采取针对性的措施加以预防和管理。

(4) 违法情况分析

违法情况主要从违法主体、违法属性、时空分布等几个方面对驾驶人及车辆违法属性进行分析, 反映违法主体及时空分布特性及规律。

(5) 其他综合情况分析

对交通管理其它信息, 如危化品运输信息, 勤务信息、安全宣传信息、驾校信息、重点单位、重点车辆等信息进行全方位分析, 找出规律, 指导工作。

6 数据挖掘分析

数据挖掘是对商业数据库中的大量业务数据进行

抽取、转换、分析和其它模型化处理,从中提取辅助商业决策的关键知识的一种技术。数据挖掘的任务可分为:关联分析、时序模式分析、聚类、分类、异常检测等。在道路交通管理中可以通过对道路交通数据的分析,构建各种类型的模型,总结事故发生的规律,从而提前预警和防范。本系统拟提供以下几个方面的数据挖掘分析。

(1) 主要事故责任人的模型构建和特征分析

通过数据挖掘构建相关人员在年龄、性别、驾龄、户口性质、职业、交通方式等方面的特征与交通事故的关系模型,发现事故高发人群的身份特征和行为特征。以便在后期的交通管理中,采取培训、教育、监控等方式提高该类人群的安全意识和技术水平,减少交通事故的发生。

(2) 事故多发地段模型的构建

通过数据挖掘构建公路等级、交通信号方式、道路隔离、路表情况、路表状况、路面结构、道路线性、道路类型、照明条件等方面的特征与交通事故的关系模型,发现事故多发地段的路况特征。以便在后期的交通管理中,对该类地段采取有效的交通管理措施减少交通事故的发生。

(3) 违章违法类型与恶性交通事故的关系模型构建

一般的恶性交通事故都是由当事人违章违法行为造成的。到底哪些违法行为会较多地诱发恶性交通事故,是一个需要研究的问题。通过数据挖掘可以发现该类违法行为,以便通过修改相关法律法规,加大对此类违法行为的处罚力度,从而减少此类行为的发生,降低恶性事故的发生率。

(4) 事故多发时段模型的构建

通过数据挖掘构建不同交通时段与交通事故的关系模型,发现事故多发时段的特征。以便在后期的交通管理中,对该类时段采取有效的交通管理措施减少交通事故的发生。

(5) 诱发事故多发的天气因素的分析

通过数据挖掘构建天气类型、温度、湿度、能见度等方面的天气特征与交通事故的关系模型,发现事故多发天气特征。以便在后期的交通管理中,对该类天气时段采取有效的交通管理措施减少交通事故的发生。

(6) 事故多发车辆模型的构建

通过数据挖掘构建车辆品牌、号牌种类、车辆类型、合法状态、荷载、安全状况、行驶状态等方面的

车辆特征与交通事故的关系模型,发现事故多发车辆特征。以便在后期的交通管理中,对该类车辆采取有效的交通管理措施如强制报废、改装、维修等减少交通事故的发生。

(7) 事故综合模型的构建

通过分析构建上述因素诱发交通事故的综合模型。

(8) 其它

运用数据挖掘在交通管理中还可以根据学员的交通违法情况创建驾校评分模型,对驾校的教学情况进行考核和监督;建立交通流量模型,对交通流量进行预测;建立交通事故保险理赔模型,对交通保险相关条文的修改提出建议等。

7 结语

为了满足未来信息化社会的要求,我国提出了以智能型综合交通运输体系作为我国交通运输发展的长期战略目标,从而也指出了未来我国交通运输系统发展的趋势是智能化与集成化。交通运输系统的智能化与集成化需要实现多种交通方式之间、多个交通部门之间、多个应用系统之间的有机整合,而交通综合信息平台则是在信息化条件下实现跨部门、跨系统进行整合的主要途径。本文针对现有交通管理系统存在的问题,对交管综合业务管理系统的需求进行了分析,提出了一个基于商务智能的交管综合业务管理系统的总体框架。并在交管综合业务管理这一背景下,对数据仓库设计、异构数据集成、联机分析处理和数据挖掘等相关问题进行了研究,提出了系统实现的目标和可实现途径。本设计已经在省级交管综合业务管理系统中进行了应用,获得了很好的应用效果。

参考文献

- 1 李瑞敏,陆化普,史其信.综合交通信息平台发展状况与趋势研究.公路交通科技,2005,22(4):90-94.
- 2 史其信,郑为中.智能交通系统(TIS)共用信息平台构架及解决方案初步分析.交通运输工程与信息学报,2003,1(1):42-46.
- 3 杨晓光.面向中国城市的先进的交通控制与管理系统的研究.交通运输系统工程与信息,2004,4(4):49-53.
- 4 杨兆升,保丽霞,朱国华.深圳市综合交通信息平台系统分析与设计.公路交通科技,2005,22(2):100-103.
- 4 姬杨蓓蓓,杜豫川,孙立军.上海市城市道路交通信息平台系统分析与设计.交通与计算机,2006,24(1):54-57.
- 5 秦鸣,史磊.南昌市综合交通信息平台分析与设计.科技广场,2009,(7):123-125.