

分布式处理技术下公交信息感知挖掘系统设计

黄正锋¹, 任刚², 华璟怡³, 陆丽丽²

(1.宁波大学海运学院, 浙江宁波 315211; 2.东南大学交通学院, 江苏南京 210096; 3.公安部交通管理科学研究所, 江苏无锡 214151)

摘要: 采取集中式处理技术对大规模公交信息进行处理时, 传输与运算耗时长, 不利于实现公交信息实时发布的需求。相比之下, 公交车载智能终端可以针对单车信息进行独立感知与挖掘。由此构建基于分布式处理技术的公交信息感知与挖掘系统框架, 包括公交车载智能终端、公交信息服务平台、公交调度中心三大模块。进而明确系统实施流程, 并探讨公交信息挖掘软件的信息映射模型。由于分布式处理技术下的公交车载智能终端与调度中心可联网共享及交互信息, 因此该系统除了实现公众信息实时发布之外, 还能为交通部门的公交优先管制技术提供基础数据支撑。

关键词: 智能公交系统; 分布式处理技术; 公交信息感知; 公交信息挖掘

Design of a Bus Information Perception and Mining System Based on Distributed Processing Technology

Huang Zhengfeng¹, Ren Gang², Hua Jingyi³, Lu Lili²

(1.Faculty of Maritime and Transportation, Ningbo University, Ningbo Zhejiang 315211, China; 2.School of Transportation, Southeast University, Nanjing Jiangsu 210096, China; 3. Traffic Management Research Institute of the Ministry of Public Security, Wuxi Jiangsu 214151, China)

Abstract: When processing massive bus information with centralized processing technology, the transmission and calculation is time-consuming and undermines the broadcasting of bus information in real time. In contrast, the onboard intelligent terminal can perform independent perception and mining of single-bus information. This paper develops a framework of bus information perception and mining system based on distributed processing technology, which consists of three key modules: bus onboard intelligent terminal, bus information service platform, and bus dispatching center. Meanwhile, the implementation process of the system and information mapping model for the bus information mining software are discussed. Since the onboard intelligent terminal is connected with the dispatching center to share and exchange information through the distributed processing technology, the system can not only realize real-time bus information broadcasting but also provide base data support for bus priority control in other transportation sectors.

Keywords: intelligent public transportation system; distributed processing technology; bus information perception; bus information mining

收稿日期: 2014-04-28

基金项目: 国家高技术研究发展计划(863计划)“公交车辆行程时间预测技术”(2012AA112304)

作者简介: 黄正锋(1986—), 男, 浙江金华人, 博士, 讲师, 主要研究方向: 交通规划与控制。

E-mail: huang321000@gmail.com

城市交通拥堵问题伴随着城镇化进程产生, 是影响居民出行的一种城市病。间断性供需失衡在城市交通网络中一直存在, 对交通拥堵的完全规避难以实现, 因此缓解拥堵措施应力求维护大多数出行者的利益。智能公交系统的建设就是该原则的体现, 以先进理论为指导, 通过技术改造提升公交乘客时间价值, 吸引私人小汽车出行者选择公共交

通出行, 改善道路拥堵局面。目前, 中国智能公交系统建设仍停留在初级规划与试实施阶段, 还难以成为国家公共交通优先发展战略的有力保障。公交信息服务是提升公交服务质量的重要手段, 当前智能公交系统的发展应当以解决公交信息感知与挖掘技术为首要任务, 以此为基础拓展到公交信息服务层面的应用。

1 公交信息感知与挖掘技术发展现状

发达国家公交信息感知与挖掘技术的研发较早,已嵌入到一些成熟的公交运营管理系统中。在美国,这套技术被应用于城市公共交通管理局(Urban Mass Transportation Administration, UMTA)推出的先进的公共交通系统(Advanced Public Transport Systems, APTS),包括多个子系统:通信系统、地理信息系统、车辆自动定位(Automatic Vehicle Location, AVL)系统、乘客自动计数(Automatic Passenger Counting, APC)方法、公交运营软件、交通信号优先控制方法和乘客信息服务系统^[1]。另外,其他发达国家在公共交通建设中也应用了这套技术,例如,日本东京都交通局开发的都市公共交通综合运输控制系统(Centralized Transit Control System, CTCS),以及欧洲多数城市构建的公交专用车道系统。

在中国,上海、北京等大城市通过安装电子站牌、车载定位系统(Global Position System, GPS)等举措逐步迈入公交智能化时代。但是与国外先进的公共交通系统相比,中国的智能化水平仍然很低,很少利用动态交通信息和车辆定位信息对系统进行深层次的开发。例如,相关的公交智能系统多数单纯使用距离与车速的比值求得车站间运行时间致使准确率降低,因此所提供的信息服务并不可靠。换言之,不成熟的公交信息感知与挖掘技术导致公交公司几乎没有采用公交数据采集、处理、交互与发布的成套技术。虽然有些传统方法采集了公交IC卡数据,并集中发送至调度中心进行统一处理,但多数属于事后的公交数据分析技术。之所以难以做到实时信息分析处理与发布,是由于公交IC卡数据量过于庞大,很多大中城市持卡乘客比例达到60%以上(南京市已接近80%),公交IC卡信息一日的数据量可达数百万条,如果集中进行处理将超出传统数据分析技术的实时性要求^[2]。此外,如果结合GPS设备采集得到的自动定位和速度检测信息进行共同传输和分析,则会使实时处理的难度加大。

中国的智能公交系统所涉及的数据感知与挖掘技术多数属于集中式处理方法,该方法可应对公交车辆少、乘客有限的情况,但难以适应中国公交车辆多、容量大的环境。据统计,北京、上海在“十一五”末年分别拥有公交营运车辆达2.5万辆和1.8万辆^[3-4],高峰时段几乎满载运行。因此,中国有必要开发一套能实现实时信息处理、发布

功能的公交信息感知与挖掘技术。

2 公交信息感知与挖掘系统

2.1 分布式处理技术

基于公交车载智能终端配置,将收集到的信息由短距离传输技术输入到智能终端进行信息挖掘。终端处理器内置公交信息挖掘软件,使得每辆公共汽车都有自主数据实时处理能力,可减少公交调度中心服务器的数据处理任务,从而降低运算时耗。此技术克服了集中式处理面对大数据所遇到的高时耗缺陷,达到节省时间成本的效果,以真正实现信息发布的实时化。由于分布式处理技术下的公交车载智能终端与调度中心可联网共享及交互信息,因此该系统除了实现实时的公众信息服务功能之外,还能为交通部门的公交优先管制技术提供基础数据支撑,可适应中国智能公交系统深层次的开发需求。

2.2 系统框架

公交信息感知与挖掘系统总体目标是:以公共汽车内检测装置所感知的动态数据为基础,以公交车载智能终端信息挖掘软件开发为核心,服务于公交信息发布及公交优先管制。

公交信息感知与挖掘系统由三大模块组成(见图1):

1) 公交车载智能终端。

作为公交信息独立处理器,是分布式公交信息感知与挖掘系统的体现,规避了因集中式处理而产生的信息滞后现象,提高了信息的实时处理与发布能力。公交信息感知装置和公交信息挖掘软件是其中两块重要组件。

公交信息感知装置包括GPS定位测速装置、IC卡信息读取装置、手机信息读取装置、视频检测监控装置等。信息感知装置能够检测并采集多种实时公交运营数据,包括公交载客量、到达时间、行驶速度等,用于公交信息融合及预测。

公交信息挖掘软件是公交车载智能终端的核心软件,用于处理采集得到的原始公交数据。信息挖掘软件常由数据库和数据平台组成^[5]。数据库是按照数据结构来组织、存储和管理数据的仓库,在本文中分为公交地理信息系统(Geographic Information System-Public Transit, GIS-PT)地图数据库(由联网生成)、实时采集数据库、历史数据库和反馈数据库,从而实现数据采集、挖掘、管理、应用的分层次数据结构。数据平台提供分

析、挖掘、共享交互和接口标准规范等各种数据服务。其中：数据接口服务模块对不同采集设备的数据以及与外部系统共享的数据进行标准化处理，解决汇集数据与实际情况偏差的问题；数据融合处理模块对多源同类型数据进行融合分析，为下一步数据应用提供标准支撑；公交状态估计预测模块在对历史数据进行挖掘和实时采集数据进行分析的基础上，不断完善信息提取模型，提高交通模型预测的准确性；数据归档服务模块用于从数据库中提取相应的数据进行处理，将处理后的数据发送至数据库进行存储。

2) 公交信息服务平台。

可将公交车载智能终端挖掘得到的信息直接联网传输至此进行发布，为各类出行者提供公交服务信息。公交信息发布途径大致包括：手机应用软件供出行车联网查询公交出行信息，帮助出行者进行出行方式及路径选择；交通枢纽触摸屏提供公交路径等多种信息查询，便于乘客选择公交线路；电子站牌显示公交车辆预计到达时间等提醒信息，为乘客选择公交车次服务；电话语音记录公交车辆到达各个编号车站的预计时间等信息，供乘客手机拨号查询；广播电台实时发布突发事件造成的公交延误等信息，有益于特殊事件下的决策调整。

3) 公交调度中心。

将整个城市的GIS-PT地图数据库及实

时获得的公交信息作为输入参数，既服务于公交调度，还有助于公交优先管制。公交优先管制技术是在共享公交信息的基础上提出的公交服务水平提升技术。例如，联网对信号控制机进行配时优化、主动控制交叉路口的公交延误时间等。

2.3 系统实施流程

公交信息感知与挖掘系统实施流程如图2所示。公交车载智能终端作为信息处理发布中主要调用的处理器，成为该系统的底层处理器，而调度中心只是在公交系统外部交通管制过程中才触发运行的上层处理器，由此形成分布式系统。公交车载装置收集得到GPS定位测速数据、视频监控数据、IC卡和手机信息等识别数据，通过近距离传输技术输入到公交车载智能终端，统一在实时联网更新的GIS-PT地图数据库平台上进行数据操作，由信息挖掘软件进行地图匹配、样本过滤、参数校正等数据处理分析，最终获得可发布的信息。一方面通过GSM(Global System for Mobile Communications)/GPRS(General Packet Radio Service)/3G(Third Generation)无线通信和IP网络将数据传输到交通枢纽触摸屏、电子站牌、到站查询电话端、广播电台站、手机移动终端等。另一方面将信息传输至公交调度中心，进行其他交通管制优化。例如，通过信号机终端调整获

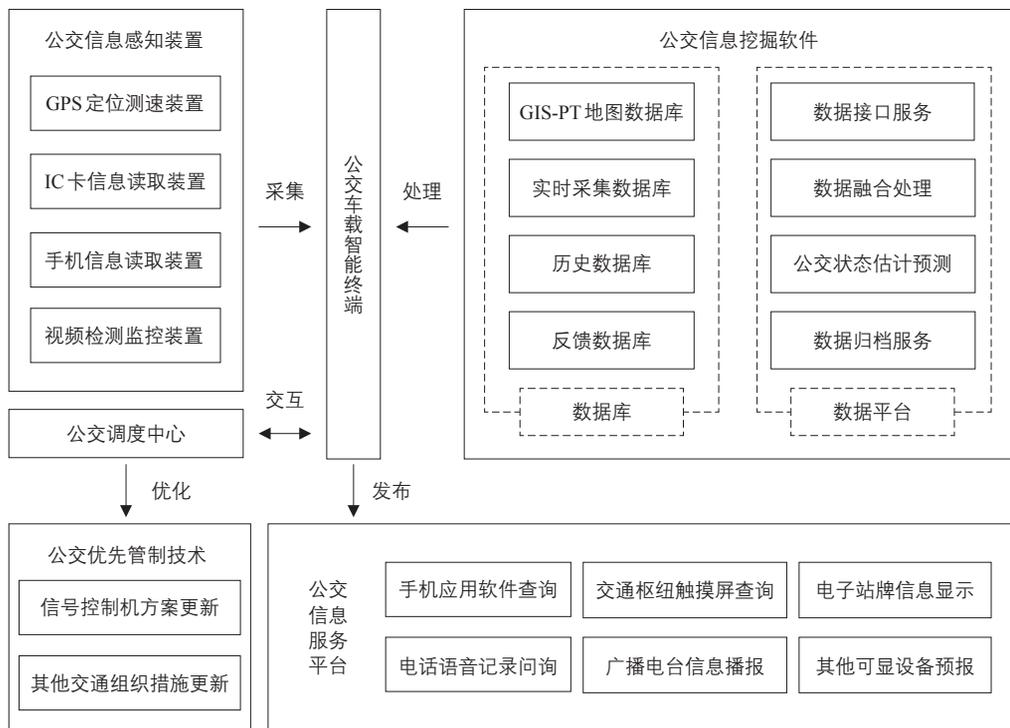


图1 公交信息感知与挖掘系统框架

Fig.1 Framework of the bus information perception and mining system

得公交优先信号，或者进行突发事件下的公交调度和交通组织优化。此类事件包括大型活动交通组织，提供的信息有助于交通组织者进行地面公交车辆调度及道路交通疏导，对潜在拥堵交通流进行主动控制。

2.4 核心软件的信息映射模型

公交信息挖掘软件根据获取的感知信息，经由信息挖掘过程处理，生成公交发布或共享信息。其中，输入的感知数据与输出的发布数据之间存在映射关系。建立合理的信息映射模型是实施挖掘过程的理论指导，是实现公交信息挖掘软件功能的有力保障，是体现系统可靠性的核心环节。

公交信息挖掘软件的输入数据类型包括视频监控图像识别数据、GIS-PT 地图数据、手机射频消费数据、乘客 IC 卡消费数据，以及 GPS 定位测速数据；输出数据类型包括手机应用软件数据、电子站牌显示数据、交通枢纽触摸屏数据、电话语音记录数据以及广播电台播报数据。图 3 给出了软件输入、输出数据变量映射，在此对重要的输出数据及其所利用的输入数据进行说明，揭示公交信息挖掘过程中的信息映射规律。

1) 公交路径运营时间(TRT)。

通过 GIS-PT 地图界面为乘客提供由等车、换乘、公交运行等时间组成的全程出行时耗信息，能够在手机应用软件和交通枢纽触摸屏上显示。这项数据的获取不仅需要采集对象车辆的所有输入数据，而且需要采集其他相互影响的公交车辆联网共享数据。

2) 预计到达车站时间(AST)。

特定线路运营车辆的到站预测时间可通过手机应用软件、电子站牌、交通枢纽触摸屏和电话语音方式进行发布。服务于车站等车乘客，方便乘客对公交线路选择做出决策。由于固定公交线路的到站时间只受自身运营条件影响，因此这项数据的获取只需输入对象车辆检测得到的数据。

3) 预计到达交叉口时间(AIT)。

属于拓展的公交信息服务，将常规路网节点也作为决策点，因此可在具有电子地图显示功能的手机应用软件和交通枢纽触摸屏上显示。一方面可以为乘客提供与私人小汽车对比的出行时耗，另一方面可以为交通管理者提供交叉口交通控制决策的输入数据。这项数据的获取主要依赖于车内 GPS 检测数据。

4) 到达车站时车内乘客数(SN)。

作为公交服务质量或舒适度的一个体现，影响乘客的乘车选择，可在电子站牌上显示。这项数据由视频、手机、IC 卡数据融合分析得到，可进一步转化成人均道路使用面积，为交通管理者提供公交优先管制的实施依据与评价指标。

5) 公交延误信息(TD)。

作为公交运营状态的评价指标或者突发事件的判别指标，可选择通过广播播报以便使所有道路使用者获知，由公交运营的时间、速度等指标分析得到。

以上所述公交信息挖掘软件不仅能够用于获知所发布的信息，也能够拓展他用，例如识别常发型与偶发型交通拥堵事件。具体实现过程如下：通过 GPS 数据得知公交车辆是否长时间在道路上停驶等交通状态，推断

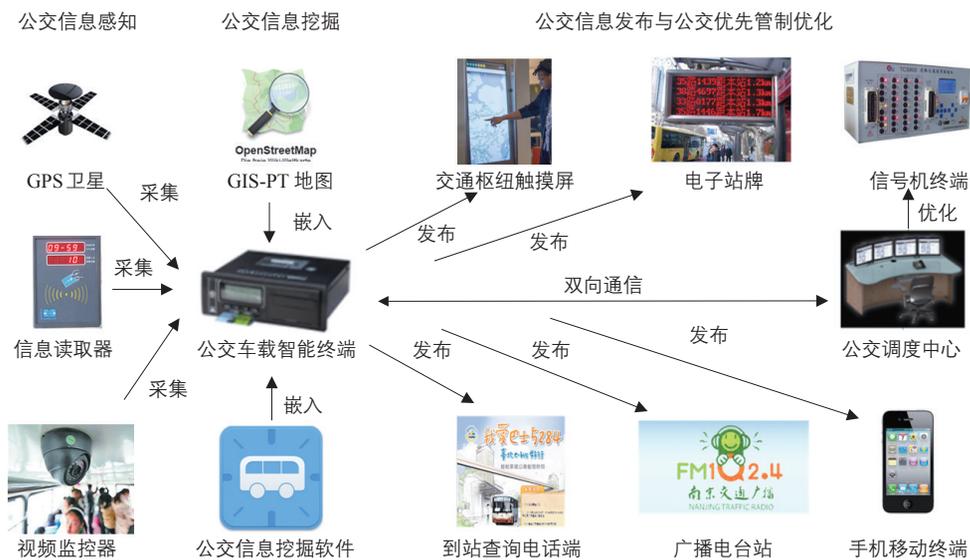


图2 公交信息感知与挖掘系统实施流程

Fig.2 Implementation process of the bus information perception and mining system

对应路段是否发生交通事故或者公交车辆抛锚等偶发事件。

3 结语

文献[6]将智能交通技术的发展分成四代, 公交信息感知与挖掘系统能为中国大数据交通背景下的第四代智能交通技术的发展提供公交营运方面的设备和软件产品支撑。公交车载智能终端作为信息处理器, 提高了公交信息实时处理与发布能力, 由此设计了基于分布式处理技术的公交信息感知与挖掘系统, 为易产生大规模公交数据的中国城市提供智能公交体系建设的核心技术。中国已将智能交通建设纳入城市交通发展的长期战略, 相关法律法规配套出台, 为该系统的实施提供了保障。然而, 国家尚未出台智能公交设备应用的国家及行业标准, 有必要事先拟定公交信息感知与挖掘系统的地方或企业标准, 为分布式处理技术的应用提供便利接口。根据示范效果具体情况修订地方或企业标准内容, 进一步上升至国家及行业标准, 为推广本系统做好技术标准服务。

参考文献:

References:

- [1] 张泉, 黄富民, 杨涛. 公交优先[M]. 中国建筑工业出版社, 2010.
Zhang Quan, Yang Tao. Public Transport Priority[M]. China Architecture & Building Press, 2010.
- [2] 戴霄, 陈学武. 基于公交IC信息的公交数据分析方法研究[D]. 南京: 东南大学, 2006.
- [3] 北京市交通委员会. 北京市“十二五”时期交通发展建设规划[R]. 北京: 北京市交通委员会, 2012.
- [4] 上海市交通运输和港口管理局. 上海市城市公共交通“十二五”规划[R]. 上海: 上海市交通运输和港口管理局, 2012.
- [5] 东南大学. 广靖锡澄高速公路数据中心设计[R]. 南京: 东南大学, 2013.
- [6] Ran B, Jin P J, Boyce D, et al. Perspectives on Future Transportation Research: Impact of Intelligent Transportation System Technologies on Next-generation Transportation Modeling[J]. Journal of Intelligent Transportation Systems, 2012, 16(4): 226-242.

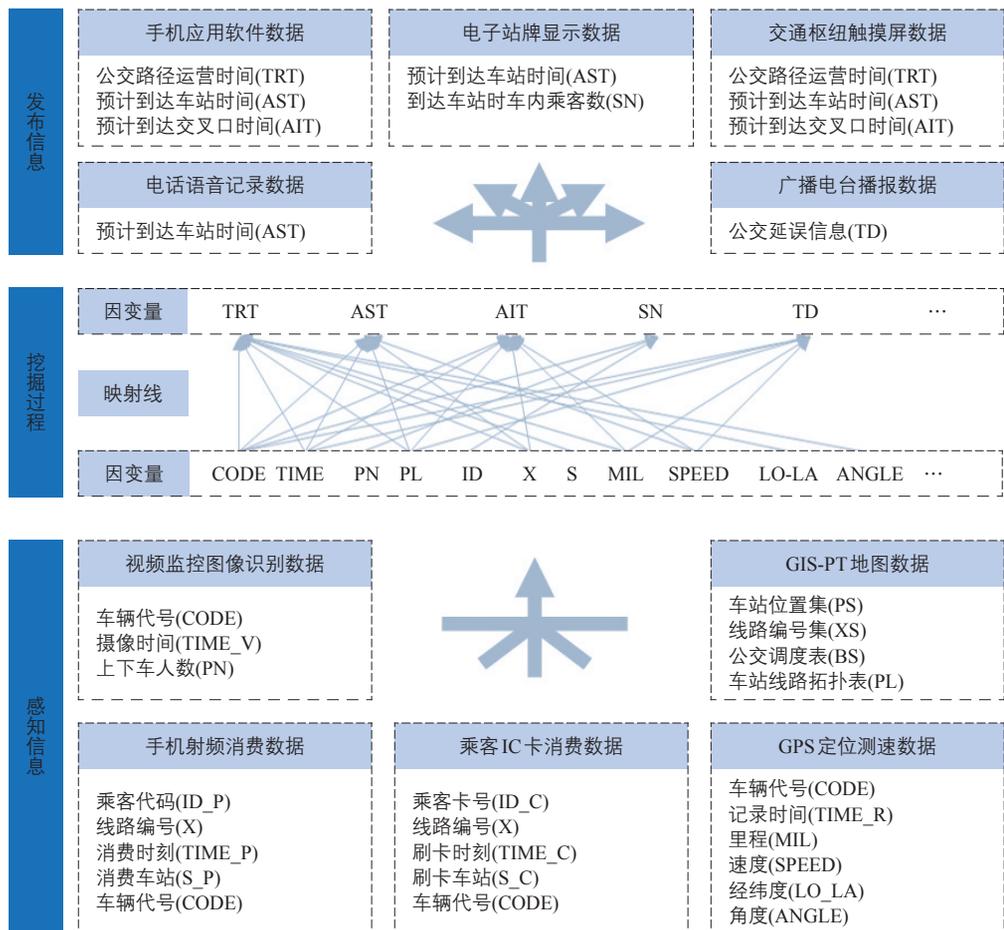


图3 车载智能终端公交信息挖掘软件的信息映射模型

Fig.3 Information mapping model of the bus information mining software at the onboard intelligent terminal