

·前沿与热点·

大数据在公共交通中的应用

陈 美

(武汉大学信息管理学院 湖北武汉 430072)

摘 要:及时、高效、准确的交通数据获取是分析交通管理机制,构建合理城市交通管理体系的前提,而这一难题可以通过大数据管理得到解决。美国、英国和我国的深圳对大数据在交通管理中的应用进行了实践与探索。通过大数据进行公共交通管理将面临着如何开放公共交通数据、如何保护个人隐私、如何进行交通数据的存取等问题,可通过注意保护个人隐私信息、提高交通数据存取的多样性、提高交通数据质量等途径去解决。

关键词:大数据 智能交通 公共交通 数据存取

中图分类号:G250.7

文献标识码:A

文章编号:1003-6938(2012)06-0022-07

Application of Big data in Public Transportation

Abstract There are many public traffic problems in the field of public service. In order to solve these problems, this paper generalizes long tail characteristics of big data. Based on these, this paper discusses the affects of big data on traffic problems, and analyzes the elements in the process of solving this problem, then makes some suggestions about big data management in order to perfect the construction of intelligent transportation.

Key words big data; intelligent transportation; public transportation; data access

近年来,各国日益关注“大数据”,力图通过扩大其在国内的应用范围,进一步释放数据所蕴含的潜在价值。2012年3月29日,奥巴马政府公布“大数据研发计划”,旨在改进现有人们从海量和复杂的数据中获取知识的能力,从而加速美国在科学与工程领域发明的步伐,增强国家安全,转变现有的教学和学习方式^[1]。我国学届亦于2012年7月22日在北京大学举行“首届中国大数据应用论坛”,主要议题包括大数据的发展趋势、不同场景的大数据应用、云计算与大数据、大数据与商业智能等^[2],旨在共同讨论大数据的应用价值。同时,交通拥堵、交通污染日益严重,交通事故频繁发生,这些都是各大城市亟待解决的问题,科学分析交通管理体系成为改善城市交通的关键所在。为此,及时、高效、准确获取交通数据是分析交通管理机制,构建合理城市交通管理体系的前提,而这一难题可以通过大数据管理得到解决。

1 大数据概述

1.1 大数据的概念

互联网等信息技术的发展导致了信息大爆炸,使得

我们进入了“大数据”时代,尽管目前对大数据还没有统一的和更为权威的定义,但对其实质的认识基本趋于统一。如IBM实体分析的首席科学家杰夫·乔纳斯认为,大数据是这样的一堆数据:它是如此之大,并且利用起来如此之好,以至于它能够做出大体上更好的预测^[3];维基百科认为,在信息技术中,大数据是一个数据集的集合,这个集合是如此大而复杂,以至于它很难通过现有数据库管理工具来进行处理^[4]。以上两个定义的共同点是:大数据是大量数据的集合。它们也揭示了其管理的难度与利用的价值。大数据更多的特点可以用“4V”(Volume、Velocity、Variety、Virtual)^[5]来描述(对其特点的提炼见表1)。

表1 大数据的特征

体积大 (Volume)	大型数据集规模一般为TB级左右,而大数据一般是PB级至EB级。
速度快 (Velocity)	大数据不需要消耗仓储,且大数据的数据挖掘技术与传统有所不同,因而其处理速度非常快。
种类多 (Variety)	大数据类型很多,不再是传统结构化数据,更多是非结构化、分布式和单调模式,如日常所使用的日志、视频、图片等。
实时性 (Virtual)	尽管大数据的容量非常大,但一旦用户提交数据需求,大数据就可自动且及时提取相关部分。

收稿日期:2012-11-28,责任编辑:魏志鹏

1.2 大数据已有的应用领域

世界各地应用大数据解决问题的范围广泛,既有形成“智能电网”来促进输电网络现代化,也有诸如沃尔玛这样的企业形成库存管理系统来解决运营问题。整体看来,云时代的到来促进了大数据的广泛应用:在天文学领域,大型望远镜和先进的光探测器促使宇宙界的数据快速增加,天文学研究人员由此能在几分钟内从大数据中提取相关内容,以鉴定所有的类星体^[6];在生物医学领域,生物医学的大数据有助于促进智能健康服务系统的形成^[7]。在气候学领域,一家名为气候公司(Climate Corporation)的企业将大数据、气候学、农业经济学相结合,通过产生与天气相关的大数据来保护农业^[8]。大数据除应用以上多科学研究领域外,现已扩展到公共问题领域,例如交通管理和控制。世界各地政府也都纷纷将交通运输数据由纸质型转向数字方式储存,建立智能交通系统,查看交通流量计数并依据车辆行程和路况拥挤程度进行电子收费,从而对交通堵塞和交通污染排放进行隐形控制。

2 大数据管理与智能交通

2.1 大数据:改变传统公共交通管理的路径

社会经济的快速发展促使城市机动车辆的数量大幅增加,城镇化的加速打破了城市道路系统的均衡状态,传统的交通管理信息系统难以满足当前复杂的交通需求,交通堵塞成为棘手问题。而长尾理论^①告诉我们,大数据将海量数据聚合在一起,将离散的数据需求聚合能形成数据长尾,从而满足传统中难以实现的需求,例如交通需求。即:个体对某一道路路段的特定信息需求,尽管人数不多、份额较少,但聚合效应加长了数据的长尾,这些个体的需求数据会在数据的长尾下形成庞大的需求市场,为满足人群的需求以及个性化服务奠定了数据基础。可见,用大数据管理交通是交通管理模式的变革,与此同时也变革了公共交通市场管理的整个内涵,阻碍传统交通的瓶颈可通过大数据解决。其原因在于:

第一,大数据可以跨越行政区域的限制。行政区域的划分是国家为了有效统治和管理,而将一个国家划分不同行政区域。这个划分在促进各个行政区域自治的同时,也导致各个地方政府追求各自辖区利益的最大化,而对

地方政府之间边界区的公共交通基础设施建设、过境交通线路等漠不关心。交通大数据的虚拟性,有利于其信息跨越区域管理,只要多方共同遵照相关的信息共享原则,就能在已有的行政区域下解决跨域管理问题。

第二,大数据具有信息集成优势和组合效率。我国大部分城市的各类交通运输管理主体分散在不同主管部门,呈现出条块分割的现象。这种分散造成公共交通管理的碎片化,如交通信息分散、信息内容单一等问题。大数据有助于建立综合性立体的交通信息体系,将用户可能利用的各种交通数据纳入系统,构建公共交通信息集成利用模式,发挥整体性交通功能,通过在大数据中进行集成检索、利用和分析来提取相关信息,满足各种交通需求,以解决实时交通障碍。

第三,大数据能较好的配置公共交通信息资源。传统的交通部门权责界定未厘清,专业分工的细化也促使公共交通管理部门职能重叠,因而在运营上浪费大量人力、物力。大数据能辅助人们制定出较好的统筹与协调解决方案,在各个交通部门之间合理配置交通职能,针对有关道路问题进行合理信息资源配置。

第四,大数据能促进公共交通均衡性发展。用传统的思维来改善交通拥堵,一般是加大基础设施投入,即加宽道路、增加道路里程来提高交通通行能力,但这种做法又会受到土地资源的限制,而且这种解决模式不利于交通发展、城市空间发展以及土地利用发展这三者之间的整合。大数据解决方案则将技术决定论与制度理论相结合,将信息技术应用于公共交通,从制度层面提高信息资本利用率,减少对诸如土地等外部资源的依赖。

2.2 大数据在解决公共交通问题上的优势

2.2.1 提高交通运转效率

公共交通的改善所涉及工程量较大,而大数据的大体积特性有助于解决这种困境。在大数据中,随着数据库摄入更多数据,所消耗的计算工作量则递减。换言之,在对公共交通的车辆进行配置过程中,配置成本会随着大数据的聚合而减小。这种高效配置能提高车辆的有效路段里程,进而提高交通运输效率,例如,根据美国洛杉矶研究所的研究,在车辆运营效率增加的情况下,减少46%~84%的车辆运输就可以提供相同或更好的运输服务^[9]。伦

^①长尾理论是《连线》杂志主编克里斯·安德森旨在解释经济领域的现象而提出的,即“我们的文化和经济重心正在加速转移,从需求曲线头部的少数的大热门(主流产品)转向需求曲线尾部的大量利基产品。在一个没有货架空间的限制和其他供应瓶颈的时代,面向特定小群体的产品和服务可以和主流热点具有同样的经济吸引力。”参见【美】克里斯·安德森·乔江涛、石晓燕译:长尾理论[M]北京:中信出版社,2009:51。

敦市利用大数据来减少交通拥堵时间,提高运转效率。当车辆即将进入拥堵地段,传感器可告知驾驶员最佳解决方案,例如帮助驾驶员 15 分钟内找到免费停车位,这大大减少了行车的经济成本。

2.2.2 有利于促进交通的智能化管理

大数据的实时性,使处于静态闲置的数据一旦被处理和需要利用时,即刻可被智能化利用,面向用户的智能软件应用程序还可以将那些浩瀚数字转换成可理解的图形化界面。公共交通的智能化管理表现在:一旦某个路段发生问题,能立刻从大数据中调出有用信息,确保交通的连贯性和持续性;另一方面,大数据具有较高预测能力,可降低误报和漏报的概率,可随时针对公共交通的动态性给予实时监控。因此,在驾驶员无法预知交通的拥堵可能性时,大数据亦可帮助用户预先了解。例如,在驾驶员出发前,大数据管理系统会依据前方路线中导致交通拥堵的天气因素,判断避开拥堵的备用路线,并通过智能手机告知驾驶员^[9]。

2.2.3 节约资金

交通运输是美国仅次于房屋的第二大消费成本,美国司机一年只有 4% 的时间在开车,但却要每年为车辆支付 8000 美元^[9]。在智能交通管理下,尽管引入处理大数据的超级计算机需要耗费一定资金,每年对其的维护也需耗费一定财力,但是从长远来看,其经济效益更大。如在新泽西州引入大数据处理交通堵塞问题之前,其主要依赖交通摄影机和耗资 2 万美元的路边传感器,但这些信息仅覆盖整个州道路的 5%。引入 INRIX 大数据管理系统^①之后,尽管新泽西州每年耗费在 INRIX 系统上约 45 万美元,但其覆盖面更广,信息准确性更高,而且给人们减少的时间成本都是无法计量的^[10]。为解决交通问题,过去城市至少要投资数百万资金来升级公路基础设施,如扩阔道路和兴建更多停车场等等,受制于资金限制很多城市就难以解决这一民生问题。用大数据管理系统解决交通拥堵,不仅可以降低管理成本、提高功效,而且还有益于城市交通管理的规范化。

2.2.4 适于海量数据处理

大数据的智能交通管理系统特别适于处理大型数据,该系统的设计是基于云计算、云管理和云操作系统的,因此不仅能满足海量数据处理和实时分析的要求,还

能 24 小时覆盖所有网络,对交通堵塞检测和报警可跨区域信息共享。这种数据处理能力是突破交通问题瓶颈的动力。目前,IBM 正计划创建预测交通的智能管理系统,分析相关信息并提出回应事故和其他异常现象的最佳解决方案。

现今,人们已经具备了利用大数据管理系统的基础。在硬件上,移动手机和无线设备的迅速扩张有助于诸如 INRIX 这样的大数据系统更方便地开发、储存和分析数据;在软件上,一些针对交通问题的应用程序得到较高认可,如谷歌 Android 平台的地图应用程序就得到 5000 万的下载量。此外,许多政府、企业、高校针对大数据在交通领域的应用研究也提供了指引作用,例如,IBM 研究中心与加利福尼亚州运输部以及加利福尼亚大学伯克利分校的加州创新运输中心(CCIT)针对交通的大数据管理这一主题进行合作,旨在预测上班族的交通条件;像谷歌、苹果和超级技术的企业也使用大数据来解决交通问题。

2.3 应用大数据解决交通问题的基础

在交通问题解决过程中,基于大数据的智能交通数据处理体系流程依次为:(1)输入交通数据。这些数据包括静态数据和动态数据,前者指道路环境、车辆信息等长时间不会改变的数据,这类数据通过线圈(类似于磁性检测器)和摄像机(交通视频)进行搜集,后者指在交通运行中而产生的实时数据(如车辆行驶速度),这类数据通过 GPS 全球定位技术、手机网络信号来搜集。(2)数据中心对实时交通流数据进行提取,同时规定统一的数据格式,促进数据交换中心之间对数据进行交换和处理。(3)通过基于云计算的云存储来对数据进行储存,将大数据集成起来。(4)控制中心将这些大数据在电脑地图上以不同色彩来呈现,分别以不同颜色注明各个路段的拥堵情况(见图 1)。在这一体系中,为了真正利用好大数据,必须要处理好如下问题:高速连接、大数据管理、开放数据等。

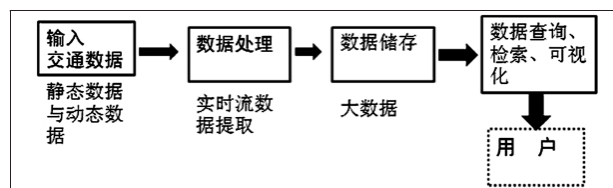


图 1 智能交通的数据处理体系

2.3.1 高速连接

①Inrix 是美国实时交通数据的提供商,其提供的 Inrix 系统是世界上最大的一个全球交通智能平台,致力于解决交通问题,其提供的 INRIX 交通及 INRIX 行程收音机等应用程序能让用户获取交通速度信息以及路况信息。

在“大数据”时代,人、设备之间都依靠数字网络进行连接,数据是推动全球经济和促进创新的重要动力。在交通领域,公共交通数据的获取途径有两种:免费连接WIFI、付费连接蜂窝数据。获取数据的途径是,通过车辆连通互联网来上传数据或下载数据,或是车辆与车辆之间进行数据共享。这些都有赖于整个地区网络速度的提升,以提高数据的传送速度,减少所有车辆的宽带费用。还有一种可能的解决方案是,在每辆车内置WIFI以便车辆之间可以共享数据,每辆车也可以从电视广播中获取数据。

2.3.2 大数据管理

大数据管理系统是一种操纵和管理大数据的数据库大型软件,用于建立、使用和维护该数据库,它对数据库进行统一的管理和控制,以保证数据库的安全性和完整性,它可使多个应用程序和用户用不同的方法在同时或不同时刻去访问数据库。但仅依靠技术是难以解决带来高效益的。需要通过高效的管理,使该管理系统可连续、实时地采集相关数据,将这些数据整合在大数据管理系统中,以便进行科学分析,并及时将智能分析后的用户个性化需求结果转交用户,使用户、交通线路以及拥堵影响因素有机聚合为相互关联的、物联化与智能化的交通运输体系。

2.3.3 开放数据

维基百科指出“开放数据具有这样一种思想:对任何人来说,他们都能如自己所愿那样免费使用和再次发布某些数据,而不用受版权、专利或其它控制机制的限制。”^[11]开放数据所提倡的是一种积极的信息公开理念,可公开的数据必须公开,以使其发挥更大的作用。只有将与交通相关的所有信息进行公开,才有助于大数据管理系统科学的分析问题,提供被分析的数据越多越及时,所做出的结论就可能越精确。若没有与交通相关的大量公共信息的公开,大数据管理将是无源之水。

3 利用大数据解决交通问题的实践与相关问题

3.1 国内外相关实践

美国与英国是利用大数据管理交通较早与有代表性的国家,我国深圳也在探索中。

3.1.1 美国的智能交通管理

美国许多州都积极运用大数据管理理念,以实现智能交通管理目标^[12],其主要应用有:

①应用大数据减小交通堵塞。美国新泽西州安装了INRIX计算机系统,可对手机和GPS信号进行分析,因为它们是最有前途的数据源,可以保持较高的准确性。新泽西州将它们分析之后,转化为一张完整的道路交通状况地图,并在地图上以不同颜色标示各个路段的运行现状,以确定造成交通堵塞的地点。例如,如果某个位置变成红色的“蛇形”和黑色的线条,这就意味着这一段交通已停歇。

②应用大数据处理恶劣天气的道路状况。美国俄亥俄州运输部(ODOT)充分利用INRIX的云计算分析以及所提供的交通信息,帮助俄亥俄州利用大型数据,从而实现在暴风雪淹没了其400多个关键路线后,在三个小时内实现清理道路状况的目标。为了尽快将交通恢复至正常状况,ODOT使用来自气象信息站和INRIX交通高速数据的信息,以评估对全州关键路线的道路进行清理所要耗费的时间,从而提高处理道路状况的效率。这种大数据应用,减少了冬季连环撞车事故发生概率,通过提高公共安全来确保商业正常运行和日常生活有序。

③应用大数据评估路况。美国俄亥俄州运输部(ODOT)计划使用INRIX交通的高速数据和分析,以评估出关键路段的行驶可靠性。此外,为了能从高速公路改善工程中最大受益,ODOT通过INRIX确定哪些路段的行驶速度低于最高限制速度,以评估出公路改善的实施战略和须改善的路段位置。

④定位拥挤路段。波士顿城市计划推出名为“Street Bump(路拱)”的手机应用程序,该程序可以利用类似于重力系统的原理来确定城市道路中的拥挤路段。重力系统原理指的是,当智能手机屏幕处于倾斜时,通过重力来转换智能手机的方向。而与该原理有所差别的是,这款应用程序能通过检测手机中的加速度计的微小变化来确定不同路段的拥挤程度。这个应用为路段改善指明了方向,因为过去波士顿为改善路段花费大量资金,例如,波士顿城市每年除了投入8万美元来开发减速带之外,还花费约20万美元来使用传统技术测量整个波士顿道路系统。

3.1.2 英国的“智能城市”

英国除了投入1亿英镑在10个城区创建100Mb/s的全市网络之外,还投入5000万英镑用于改善其他10个城市的互联网接入,目的在于开展“连接城市”项目,以使用大数据减少交通堵塞^[13]。具体讲来,其原理就是以大数据的不间断分析为基础,利用超高速连接和强健的数据管理,将各城市之间互相连接起来,从而实现更好地控

制市区城市的目标。“连接城市”能通过大数据推进各个部门对基础设施的高度共享,并能更便捷地实现跨部门合作,从而完善交通和道路系统。在教育方面,伦敦教育信息资源管理局(LGfL)已经将2500所学校连接至高速光纤网络,从而彻底改变了首都的教育。

此外,在举办奥运会期间,英国政府也利用免费INRIX应用软件和在线服务确保交通顺畅。INRIX的大数据实时路况平台为消费者提供了一些有用的应用程序,使用户能了解何处正在拥堵,从而以选择出行的最佳模式。这些程序包括:

(1)INRIX 交通(包含 iPhone、Android 系统、Windows Phone 和黑莓的不同版本)。INRIX 交通是一个免费的软件,该软件在同行业中排名第一,这与其较强实用性、便利性有关。这个软件能根据所有驾车人的数据,帮助用户确定速度最快的路线,避免出现不必要的延误。此外,那些使用 iPhone 版本的 INRIX 交通的用户只需点击两次,就可通过文本或电子邮件将实际抵达时间告知正在路上的同伴,从而能准时他们会和,以减少时间成本损耗。

(2)INRIX 旅行收音机。英国驾驶者大多通过商业广播电台报告获取实时交通情况,但这些信息并不能满足各个车主的个性交通信息需求。然而,INRIX 旅行收音机能为用户提供按需获取整个运输 24/7 模式^①的实时行程状况,这些运输涵盖公路、铁路、海运、空运等领域。此外,INRIX 旅行收音机亦会对全国范围内的最新旅游情况进行实时更新,以保证数据准确性、及时性。对于用户而言,其只需通过下载 INRIX 旅行收音机应用程序,就可在线收听当前交通状况。

3.1.3 深圳的智能交通

深圳是我国较早设立智能交通中心的城市,该市针对交通运输问题进行了信息化建设。在智能交通方面,深圳于2000年设立智能交通指挥中心,该中心依托包括智能交通信号控制、闭路电视监控、智能交通违章管理等多个系统的交通管理网络,成为集信息、监控、指挥于一身的交通管理中枢^[14]。在交通信息化方面,智能交通指挥中心以及智能交通处是负责交通运输信息化建设工作的部门,对全市智能交通进行整体统筹,汇总城市交通信息并加强这些信息的共享,实现公共交通的动态化管理。此外,为了推进全市智能交通的发展,包括8家企业的发起

单位于2007年成立了深圳市智能交通行业协会,该协会旨在为智能交通行业建立良好沟通发展平台。

智能交通指挥中心的信息采集渠道包括两种:一种利用公共交通基础设施,诸如监控交通状况的闭路电视监控以及各种车辆检测器;另一种是人为补充,诸如110交通报警信息以及路面民警、市民所反馈的交通信息。为了对交通信息整合并共享,深圳市于2010年开始投资10亿元运行“智能交通1+6”项目,“1”即指构建一个资源共享平台,将整合交通、交警、规划等各部门各方面交通运输信息;“6”指依靠这个平台为交通的如下信息服务提供支撑:交通监测信息、交通管理信息、道路交通调控信息、公共出行信息、交通指挥应急信息、交通管理决策信息。此外,为了更好利用大数据,深圳市近年来也开始加快综合交通运行指挥中心的分中心建设,力图能聚合更多公共交通的海量数据,以多种方式实现交通信息共享,为深圳“智慧交通城市”提供支持。

3.2 利用大数据进行交通管理带来的问题

3.2.1 如何开放公共交通数据

智能交通的最大驱动力是开放数据,然而在国外的公共交通领域,大多城市是在私人数据库中管理它们的交通和运输数据,且仅由市政工作人员监视系统性能以及实施改善措施^[15]。这种对数据进行的封闭式管理不会促进信息的增值,要使交通数据真正产生价值,唯一的方法就是对其进行开放,因为开放数据对政府、企业和公民都有益处,既能为政府提高公信力和合法性,也能为企业带来商业利益,还有助于公民参与决策制定。因此,要使交通数据不至于闲置,数据的开放方式成为关键所在。

3.2.2 个人隐私问题

数据是信息产品的基础,数据能产生新的应用程序和信息服务,从而带来商业利益和社会效益。但大数据也导致了数据泛滥,这种泛滥催生出个人隐私的忧虑。在过去,个人信息可以通过匿名化、密钥编码等手段来得到保护。但如今,大数据扩大了信息范围,加快了信息传递和共享速度,若不加以严格控制,其所含的商业信息或私密信息就可能泄露,例如个人所在位置、个人出行习惯以及用户最喜欢的主路线等。一旦个人察觉到这些私密信息有泄露,就会抵制大数据管理系统的广泛应用。

3.3.3 交通数据的存取方式

^①24意指一天24小时,7意指一周7天,24/7模式也就是全天候服务。

各地机构都具有交通数据并能被大数据管理系统应用,但很多车辆计数(计算交通车辆数目)的数据都以静态格式(如PDF)存储,使得系统所具备的计数特性无法被除人之外的事物进行检索,这种传统“人对物”的互联网连接方式不符合物联网的“物对物”特性。交通数据的物联化是通过智能手机、传感器和机载车辆等硬件,不间断收集、通信和处理诸如交通条件、天气状况等移动数据。

4 创建我国基于大数据的智能交通建议

4.1 广泛开放公共交通数据

尽管大数据包含较大容量信息,但是其利用率却不高,例如,交通部门可能拥有各种所需信息,但并不总是知道以何种正确方式来筛选它。通过开放交通数据,有助于形成良好的公私伙伴关系,促使交通主管部门利用私营机构的专业信息技术知识,通过专业开发人员以及第三方应用程序开发者来完善交通信息服务,根据具体人群需求来定制个性交通服务。因此,为了更大程度开放交通数据,交通主管部门应建立诸如 Transportation Information Group^[16]的开放交通运输数据的门户网站,尽可能以XML、Text/CSV、KML/KMZ、Feeds、XLS等多种格式开放交通运输数据,提高机器可读性;同时,在门户网站上配备数据挖掘和抽取工具,促进用户根据个人喜欢来获取数据。此外,政府应在阐明数据开放原则、知识产权规则的前提下,制定促进交通数据共享的奖励措施,推动交通信息的开放和整合。如此,这种公私部门之间的合作机制有利于形成双赢局面,即开发人员可以在信息领域数占据利润丰厚的市场份额,城市则挖掘出大量创造力,加速城市化发展进程。

4.2 保护个人私密信息

为避免个人私密信息非法泄露,政府应制定一部完整的数据隐私法,对个人数据的定义、数据可发布范围、数据发布的基本原则、数据可利用的范畴等方面进行规范。交通主管部门在遵守这部法律的基础上,进一步细化可发布的交通信息,并开展数据隐私、安全的教育项目,加大用户对隐私规则的了解。除了加大法律监管之外,为了防止数据的过度开发和过度保护,交通主管把握的主要原则是:数据的商业性开发、公益性利用能够与个人隐私权之间相平衡,政府在赋予企业更大程度利用数据和获得潜在商业利润的同时,要减少公民对个人隐私和数据安全的担忧。落实这一原则的一种可行模式是:交通主

管部门将数据的开发决定权交给个人,由当事人决定对这些个人数据进行开发利用或开放限制;数据开发商则由主动向消费者提供信息服务的“推”模式,逆转为在得到当事人授权同意的前提下,根据消费者的个性化信息需求而被动提供信息服务的“拉”模式。

4.3 提高交通数据存取多样性

交通部门应加快交通信息的数字化步伐,并积极以动态方式提供存取,从而减少往交通数据中心插入数据所花的人力和时间。因此,为了满足用户的特定数据结构要求,交通部门必须聚合各种交通数据,一方面要重视数字化交通数据,另一方面要对重要核心交通数据进行纸质保存,这样可以通过资源共享的方式来丰富整个智能交通的数据长尾。此外,为了真正实现公共交通的智能化,可以加大交通数据中心的自动化程度,让用户能自动收发交通数据。

4.4 提高交通数据质量

不同交通机构会使用不同数据格式,而且数据的完整性和质量各有不同。因此,为了提供数据质量,必须在以下几方面进行改进:第一,交通数据中心要确定一个数据质量标准,确保以透明的信息发布环境来呈现原始数据,提高数据的精确性、完整性、客观性,减少用户对数据的怀疑和忧虑。第二,交通数据中心要形成一套严格的数据质量控制流程,以内部检查和外部评估相结合的方式,即数据中心要检查所采集的数据是否关乎交通机构职能行驶和公共利益的实现,而对于涉及隐私、安全的交通数据,则交由外部同行专家来进行评审。第三,数据的发布要形成良好用户反馈机制。用户要被赋予维护数据质量的权力,使其能要求信息发布机构对错误的信息进行更正,而且如果用户的回应没有得到及时反馈或反馈结果不如人意,用户可以进行投诉。交通部门要具有较高数据质量意识,倾听用户的数据需求,对数据进行及时更新,确保这些数据对于用户而言有用。如果政府在数据采集和处理方面处于劣势的话,政府也可以进行公共服务市场化,将这项服务外包给数据提供商,自身只起着数据质量评估的角色。

5 结语

在当前大数据时代,数据充斥所带来的影响远远超出了企业领域,其不仅能带来商业价值,亦能产生社会价值。随着信息通讯技术的发展,交通运输从数据贫乏的困

境转向数据丰富的环境,而面对众多的交通数据,如何从中根据用户需求提取有效数据成为关键所在。可见,大数据管理是一项巨大挑战,一方面要及时提取交通数据以满足用户需求,另一方面须在数据的潜在价值与个人隐私之间进行平衡。为了更好迎接这项挑战并实现大数据在公共交通中的应用价值,本文在探讨大数据内涵、概念、应用领域的基础上,简明阐述了大数据在解决公共交通问题所具有的优势及完善措施,但要真正利用大数据构建一体化的公共交通运输管理体制,还需要对交通数据采集、处理等方面进行梳理,需要对智能交通系统的构建以及用户界面的完善做进一步研究。

参考文献:

- [1]美国政府出台大数据研发计划[EB/OL].[2012-11-12].
http://www.most.gov.cn/gnwkjdt/201204/t20120424_93877.htm.
- [2]“首届中国大数据应用论坛”即将开幕[EB/OL].
[2012-11-12].<http://zt.ciotimes.com/20120620/>.
- [3]Jeff Jonas. Big Data Q&A for the Data Protection Law and Policy Newsletter [EB/OL]. [2012-09-15].http://jeffjonas.typepad.com/jeff_jonas/2012/04/big-data-qa-for-the-data-protection-law-and-policy-newsletter.html.
- [4]Big data[EB/OL].[2012-09-16].http://en.wikipedia.org/wiki/Big_data.
- [5] Lambert, Courtney. Big Data: It's awkward [EB/OL]. [2012-09-16].<http://www.slideshare.net/cjlambert/big-data-its-awkward>.
- [6]Ross Andersen. How Big Data Is Changing Astronomy (Again) [EB/OL]. [2012-10-16].<http://www.theatlantic.com/technology/archive/2012/04/how-big-data-is-changing-astronomy-again/255917/>.
- [7]Katharine Miller. Big Data Analytics In Biomedical Research [EB/OL]. [2012-10-16].<http://biomedicalcomputationreview.org/content/big-data-analytics-biomedical-research>.
- [8]Siraj Khaliq. Big Data Meets Big Weather [EB/OL]. [2012-10-16]. <http://strataconf.com/strata2012/public/schedule/detail/22511>.
- [9]Greg Rucks and Alisha Kuzma. How big data drives intelligent transportation [EB/OL]. [2012-09-16].<http://thinkprogress.org/climate/2012/08/30/776631/how-big-data-drives-intelligent-transportation/?mobile=nc>.
- [10]Marta Cobo Gago. Big Data being used to reduce traffic jams in 'smarter' cities [EB/OL]. [2012-09-16].<http://news.techworld.com/applications/3374804/big-data-being-used-to-reduce-traffic-jams-in-smarter-cities/>.
- [11]Open data [EB/OL]. [2012-09-17].http://en.wikipedia.org/wiki/Open_Data.
- [12]SHIRA OVIDE. Tapping 'Big Data' to Fill Potholes [EB/OL]. [2012-09-16].<http://online.wsj.com/article/SB10001424052702303444204577460552615646874.html>.
- [13]Big Data Analytics Applied to Traffic Issues in New Jersey [EB/OL]. [2012-09-17].<http://nyconvergence.com/2012/06/big-data-analytics-applied-traffic-issues-new-jersey.html>.
- [14]董海龙.深圳智能交通系统应用现状[EB/OL]. [2012-10-17].http://www.tranbbs.com/news/cnnews/ITS/news_64980_2.shtml.
- [15]Greg Rucks, Alisha Kuzma. How big data drives intelligent transportation [EB/OL]. [2012-10-25].<http://www.greenbiz.com/blog/2012/08/15/how-big-data-drives-intelligent-transportation?page=0%2C0>.
- [16]Alaska Department of Transportation & Public Facilities—Transportation Information Group [EB/OL]. [2012-11-07].<http://www.dot.state.ak.us/stwdplng/transdata/>.

作者简介:陈美(1987-)男,武汉大学管理学院博士研究生。