

对基于大数据道路交通管理的冷思考：“小”即是美

李翔敏 戴帅

【摘要】本文针对近年来备受追捧的基于大数据的道路交通管理潮流进行了回顾和冷思考，以期促进道路交通管理者及业界对大数据研究和应用的正确认识。面对大数据热潮，本文提倡在研究和应用大数据技术的过程中，要回归交通管理信息化发展的初衷，科学理性决策，既要看到大数据技术的价值，也要看到它背后的问题和挑战。为了应对近年来道路交通信息化管理所面临的新形势，既要积极稳健、科学合理地发展大数据技术，也要善于运用手头的“小数据”，科学深入运用现有的技术和资源，努力发掘“小数据”中的价值，注重实践，夯实科学交通管理的基础。

【关键词】交通管理；大数据；交通信息化管理

1 引言

自大数据（Big data 或 Mega data）的概念提出以来，西风东渐，英国人维克托·迈尔—舍恩伯格和肯尼斯·库克耶所写的《大数据》^[1]在国内出版后，更是一时间洛阳纸贵。与信息技术和网络技术相伴而生的大数据处理技术更是被当作解决各种疑难杂症的灵丹妙药，各种研究和应用不断涌现，各行各业都争相追捧，道路交通管理行业也不例外。

在近年来智能交通不断发展的背景下，大数据作为很多智能交通解决方案的基础，获得了前所未有的关注。很多道路交通管理部门以及相关企事业单位都开始大规模投入，在规划设计、软硬件配置、软件研发、数据采集处理、系统运维等方面都进行了大量的探索和尝试，形成了颇多成果，也不乏很多相对成功的实际应用^[2, 3, 4]。在成功案例的激励和各种绚丽的图形化表现形式的吸引下，越来越多的道路交通管理部门和相关企事业单位对大数据概念越发狂热，争相投入，加紧上马各种项目。与此同时，业界也有一些针对大数据应用实际效果的质疑^[5]。

学术领域有关道路交通管理行业大数据的研究也颇为活跃，从中国知网的相关文献检索结果来看（以“交通”和“大数据”作为关键词检索），相关论文数量逐年上升，研究角度逐步多样化，涵盖规划设计、平台搭建、系统架构、技术应用等方面。但是，相关研究成果主要集中在理论研究和系统设计方面，实证研究和案例研究较少，“重理论，轻实践”现象突出。

从目前业界一些示范性的大数据应用项目来看，依然是新品装旧酒，用的还是传统的数理统计分析和商业智能（BI, Business Intelligence）的方法，只是冠以大数据的帽子而已。因此，在我国交通管理领域大数据狂热的背后，存在诸多的潜在问题与风险，应该适时对基于大数据的道路交通管理热潮进行冷思考，确保大数据在道路交通管理领域中的恰当运用，消除盲目追捧，回归交通信息化管理的初衷，真正提高交通信息化管理的水平。

为了应对近年来道路交通信息化管理所面临的新形势，本文在反思基于大数据的道路交通管理的基础上，认为道路交通管理信息化的发展路径应该遵循既要发展大数据、又要善用“小数据”的原则。在追逐大数据技术潮流的同时，也要强调“小即是美”的理念，充分利用手头已有的“小数据”和技术，努力夯实科学交通管理的基础。

2 正确认识交通管理领域中的大数据

美国著名的信息技术研究和分析公司高德纳（Gartner）将大数据定义为：“是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产”。美国国际商业机器公司（IBM）则归纳出了大数据的“4V”特征，即数据量大（Volume），多样性丰富（Variety），商业价值高（Value）和处理速度快（Velocity）。虽然不同学者或机构对大数据的具体定义有所不同，但是都基本上提及了上述的4项特征。大数据给科学研究和技术应用带来了思维方式的改变，即（1）使用全样本数据而不是抽样样本；（2）关注效率而非精确度；（3）关注事物的相关性而非因果关系^[1]。大数据技术不仅能够分析结构化的数据，更倾向于对由视频摄像头、传感器、RFID、GPS等采集的非结构化的海量多样数据进行更好地分析。

大数据技术由于其强大的预测能力以及借助商业智能（Business Intelligence）手段而展现出的绚丽的图形化表现形式，常常被人神话和敬畏。其实，大数据只是信息技术和网络技术发展到一个新阶段的产物，并非“高大上”，所用到的数学原理和统计学知识也并不复杂，只是数据的处理量和处理方法上有新的改变。道路交通管理领域的大数据应用，也只是基于海量的交通数据，根据交通工程学、运筹学、数学、统计学等学科的知识，并配合新的海量数据处理方法来解决道路交通管理领域的实际问题罢了。

道路交通领域存在各种类型的多源海量交通数据，涵盖了交通流特征参数数据、“车驾管”数据、交通基础设施数据、交通管理设施设备数据等，既有静态数据又有实时数据。交通大数据的意义并非在于收集和存储海量的数据信息，而在于对这些隐含交通意义的数据进行专业化处理、分析和挖掘，将海量数据潜在的价值挖掘出来，提高数据的附加值，用以辅

助交通管理，减少交通拥堵和交通事故所带来的损失。大数据技术固然是先进的、符合时代技术潮流、能够提高交通管理水平的有效手段，但是大数据也带来了许多新的挑战，特别是对于那些缺乏分析技术、信息技术投入和与大数据相适应的管理文化的机构和部门[6]。因此，在引进和运用大数据技术的过程中，就需要科学理性，在看到大数据优势的同时，既不能夸大其作用，也应该看到它的劣势[7]。

2.1 大数据不是万能的

与所有的技术一样，大数据技术显然不是万能的。海量数据所具有的价值只是潜在且不确定的，业界经验认为海量数据库中只有约 10%-20%的数据才是真正有价值的。大数据技术的成功运用，受制于包括数据基础、算法、软硬件等各种因素。例如，数据信息采集技术的限制、采集方法的缺陷以及样本量过小等因素，都会极大影响后期的分析处理，使得大数据技术成为无本之源。

我国的道路交通管理信息化进程的起步相比发达国家较晚，在基础数据的规划设计、采集处理、管理应用等方面存在着严重不足，具体表现在：（1）各类交通信息采集设备缺乏科学统一的选型、安装和运维，致使设备无法返回有效的、高质量的数据；（2）大量交通信息采集设备功能没有发挥，例如大量视频摄像头没有安装流量检测软件而导致无法采集相应的数字信息，只能用做视频监控，造成资源浪费；（3）数据碎片化，没有形成统一的、相互关联的数据库；（4）软硬件技术水平和专业人员配备落后，数据蕴含的价值无法被挖掘，造成数据资源的闲置和浪费。

鉴于我国道路交通管理领域信息化建设尚处在早期的发展阶段，所积累的数据量有限，数据质量相对不高，很多道路交通管理部门在应用大数据技术时存在“冷启动”的挑战。因此，就需要道路交通管理部门在大数据的热潮中，能够冷静思考大数据技术的限制和应用前提，特别是在自身底子薄、零基础的情况下，需要慎之又慎，切忌为了政绩和相关利益企业的游说而盲目上马大数据项目。

2.2 大数据的是有成本的

海量的大数据与其他众多消费品一样，它的采集、存储和处理是有成本的。除了一些人工的数据采集方法，海量大数据的采集主要还是依赖于技术和资本密集型的交通检测器，而大数据的多样性又依赖于安装在不同位置的不同类型的交通检测器，这些交通检测器的购置和运维的成本较大。此外，交通大数据一般无法用单台的普通计算机存储和处理，需要依赖分布式处理、分布式数据库和云存储、虚拟化技术，而这些技术的开发应用，也需要大量资源。

因此，在追赶大数据潮流上马交通大数据项目的时候，道路交通管理部门需要首先衡量自身的人力和物力基础，在成本效益分析的基础上做考虑。在资金不足或者效益远低于成本的情况下，要果断放弃，首先做好基本的交通管理工作。在资金充裕和效益大于成本的情况下，要避免“不差钱”的想法，认真考虑自身的专业人员配备是否到位、是否具有系统运维能力等问题，避免交通大数据系统建设成为烂尾工程，或建设完毕后因使用不当而成为摆设，造成资源闲置和浪费。

2.3 隐私问题是众多大数据应用的隐患

交通是人类的基本需求之一，人们的日常出行必然会产生一系列与个人相关的交通数据。在移动互联网、卫星定位导航、车辆识别技术快速发展的今天，人们的交通出行所产生的交通信息数据更容易被采集和存储，而交通大数据技术的发展也正是顺应这一潮流而生。在法制化深入和公民个人意识觉醒的今天，公众对于个人隐私问题较为关切，交通大数据特别是与个人隐私相关的交通大数据问题就显得颇为重要。

目前，我国还未出台专门的法律法规来界定和规范信息时代下的个人隐私，一些涉嫌侵犯个人隐私的案件多采用其他相关法律法规来裁判。因此，现阶段我国交通大数据发展过程中还未对个人隐私及数据安全给予足够重视。例如，成熟的车辆号牌视频识别技术已经被交通管理部门和相关企业广泛运用，车主的号牌信息及出行信息在未被告知的情况下被记录和存储，个别交通管理部门和企业数据的管理和使用过程中不对车主的号牌信息进行相应隔离，一旦泄露或者被违法买卖，风险极大。相比之下，在我国香港地区，由于法律限制，交通管理部门或者相关企业在未经车主允许的情况下，不能够采集车辆的车牌信息，以确保个人隐私。在这种情况下，以车辆号牌视频识别技术为基础的大数据应用就受到限制。在此情况下，如果仍然想要跟踪车辆的形式轨迹而不涉及隐私问题，就需要对车辆视频识别技术进行改进，由根据号牌识别车辆改为根据外观特征识别车辆，以确保车主的个人隐私。

因此，交通管理部门及相关企业在搜集可能会涉及个人隐私的交通大数据的过程中，要有前瞻性，提前做好准备，依法依规地采集、处理、存储、应用和传播数据。虽然交管部门的交通信息数据有的较为敏感，有的涉及公共安全和公民隐私，但是并不代表交通信息数据不能公开。交通管理部门要在保障公共安全和公民隐私的前提下，有勇气推进信息公开，对涉密信息进行脱敏处理，积极稳健地逐步开放和共享数据，提供给更有能力的个人和机构进行分析并加以利用，以便最大限度挖掘交通数据信息的潜在价值，助力交通管理。比如，2014年，贵州省与阿里巴巴集团共同宣布面向全球启动“智慧交通算法大挑战”。贵阳市向参赛

者开放了海量的脱敏交通数据，包括公交车 GPS 数据、出租车 GPS 数据、高德导航数据，逐步启动大数据商业开发，为交通数据信息开放起到了很好示范作用^[3]。

2.4 发展大数据应用需要坚实基础，不能操之过急

国外发达国家的道路交通信息化管理走过了一条漫长而坚实的道路，因其先进的管理体制、完善的交通管理设施、发展相对较早的信息技术和网络技术，在过去的交通管理工作中，逐步积累了大量交通数据。在此基础上，国外同行开发了各种交通模型、应用和系统，并将它们成功应用到实际中，持续且有步骤地推进交通管理的信息化建设。近年来，由于大数据技术的发展，原本积累的很难有效处理的海量交通数据的就被充分挖掘和利用起来。

与国外发达国家相比，我国的交通管理领域的信息化建设则呈现跳跃式发展。我国的交通信息化管理起步较晚，科技、设施、人才、经济基础相对较差，以往积累的信息数据种类和数量有限，数据质量不高，利用率较低。在这样的基础上，面对大数据浪潮的袭来，业界反应不一，有的在现有基础上适当发展大数据应用迎头赶上，有的没有基础也要强上马大数据；有的积极完善数据采集、研发各种交通模型充分利用大数据技术，有的则在追逐大数据的同时却做不好基本的交叉口配时应用。

因此，发展大数据技术不能操之过急，要首先判断自身的科技、设施、人才、经济、数据的积累，做好科学系统的规划后再逐步引进和应用大数据技术。此外，在应用大数据技术的同时，不能只关注到成功的案例，还要多关注失败的案例，从失败的案例中汲取教训。

3 既要发展大数据，更要善用“小数据”：“小”即是美

基于大数据的交通管理本质上还是交通信息化管理的一种途径或者应用。目前，我国的道路管理面临新形势，对交通信息化管理提出了新的挑战 and 更高的要求。在新型城镇化的背景下，我国的机动化、城镇化进程加速，交通供需矛盾进一步加剧，运用信息化技术手段缓解交通拥堵成为必由之路。此外，在我国经济增长呈现新常态、国家预算法的出台、地方政府债务管理体制的变革的背景下，大大减少了道路交通管理部门投入的随意性，规范了投入的力度，更加注重科学规范。

为了应对道路交通信息化管理所面临的新形势，在反思基于大数据的道路交通管理的基础上，交通信息化管理的发展路径应该遵循既要发展大数据、又要善用“小数据”的原则。这里的“小数据”是指相对大数据而言数据量较小、种类较少的样本数据，大都根据传统的数理统计学原理抽样得来。基于“小数据”的交通管理，即是运用有限样本量的抽样数据，在充分研究事物因果关系的基础上，开发、标定、运用各种交通模型以辅助科学交通管理。

由于我国交通信息化管理起步较晚，各方面基础相对不强，因此，站在大数据时代的门口回望，依然会发现我国基于“小数据”的交通管理水平还有很大的提升空间。因此，最重要的是首先“盘活存量”，善用手头已有的“小数据”，努力夯实科学交通管理的基础。

(1) 充分利用“小数据”，深化现有交通管理技术的科学运用。要重视对现有交通数据的充分利用和挖掘，加强专业人员和配套资金的投入。例如，目前很多交叉路口的信号灯配时仍然不科学合理，很多是由于缺乏科学的路口交通流量调查基础、缺乏规范动态的配时调整机制、缺乏专业技术的人才以及对配时软件和模型的不正确使用或降低难度的简单使用。在缺乏科学合理的交叉路口信号配的情况下，即使引进先进的大数据技术，交通管理的效果也难以有显著提升。如果道路交通管理部门能够科学地进行路口流量调查，充分利用好“小数据”，科学深入地应用信号配时技术，才能够真正夯实科学交通管理的基础。

(2) 切勿盲目追求技术的升级改造。大数据潮流的背后更多的是投资人、技术和服务供应商、咨询公司以及新闻媒体的热炒，但实际上真正运用大数据的政府管理部门和企业却颇为迷惑。业界很快发现，很多所谓大数据技术解决方案其实是新瓶装旧酒，并未比之前的交通信息化管理带来革命性变化。传统的数据库技术和算法模型已经能够基本满足需要，并非一定需要运用诸如 MapReduce, Hadoop 等大数据技术。此外，在很多交管部门缺乏基础交通数据和专业人员的情况下，即使上马大数据项目，也很难充分发挥其作用。因此，与其投入大量的资源来上马不实用大数据技术，倒不如踏踏实实地用好已有的资源，切勿盲目追求技术的升级改造。

(3) 循序渐进提升交通信息化管理水平。海量的非结构化的大数据中确实蕴含着巨大的潜在价值，不断更新的大数据相关技术也是进行数据挖掘、辅助交通管理决策的利器，但这并不意味着就要从基础弱、底子薄的状态一下跃变到追逐技术潮流的巅峰。交通管理技术的变革是一个渐进的过程，应当稳健推行，认真研究交通参与者的切实需求，充分运用已有的交通管理技术和方法，切勿急功近利。在交通信息化管理的基础和条件仍然不够完善的情况下，就需要先用好手头的“小数据”，在充分运用好传统交通信息化管理技术手段的基础上，循序渐进，逐渐引入大数据技术。不论大数据概念如何火爆，交通信息化管理的发展都应该回归理性，以解决问题的有效性和高效性为出发点，恰当运用各种技术。

4 结语

大数据技术的发展确实能够提升道路交通管理的科学性和有效性，进一步提高道路交通管理的信息化水平。得益于大数据技术，过去很难进行处理的海量的非结构化、半结构化的

多源交通数据变得易于处理，而且处理能力和速度有了极大的提升，可以实现很多以前实现不了的功能，蕴藏在海量交通数据中的潜在价值得以被更好地发掘。但是，技术的发展和应用是一个需要坚实基础和循序渐进的过程，大数据热潮下，需要回归理性，正确认识大数据，稳健推进交通信息化管理的发展。

正确认识大数据，即是要科学、理性、全面地了解大数据的局限、成本、风险和应用基础，而非人云亦云、贪大求全地跟风盲目上马大数据项目。在自身尚未完全准备好全面拥抱大数据时代时，首先要做好的就是秉承“小即是美”的理念，善用“小数据”。善用“小数据”，即是重视当下，加强对现有交通数据信息的充分利用和价值挖掘，改进和加强科学技术对交通管理工作的辅助作用。

【参考文献】

- [1] [英] 维克托·迈尔-舍恩伯格, [英] 肯尼思·库克耶. 大数据时代[M]. 杭州: 浙江人民出版社, 2013.
- [2] 徐炜. 交通管理已进入大数据时代[J]. 道路交通安全管理, 2013, 11: 36-37.
- [3] 李佳师. 贵州用大数据解决交通拥堵[N]. 贵州日报, 2015-01-06 (011).
- [4] Yu Huijun, Wang Zhigang, Liu Xiaoyan, et al. A Big Data Application in Intelligent Transport Systems [J]. Applied Mechanics and Materials, 2015, 734: 365-368.
- [5] 马继华. 解决交通拥堵, 大数据应用靠谱吗? [N]. 人民邮电, 2014-05-05 (004).
- [6] Buckley Stephen, Lightman, Deborah. Ready or Not, Big Data Is Coming to a City (Transportation Agency) Near You [A]. TRB 94th Annual Meeting Compendium of Papers [C]. Washington DC: Transportation Research Board, 2015.
- [7] 刘涛. 大数据不是万能的[N]. 中国国防报, 2014-10-29 (004).

【作者简介】

李翔敏, 男, 硕士, 公安部道路交通安全研究中心, 交通政策规划研究室工程师, 研究实习员职称。电子信箱: robin_lixiangmin@163.com

戴帅, 女, 博士, 公安部道路交通安全研究中心, 交通政策规划研究室副主任, 副研究员职称。电子信箱: blue80520@163.com