

# 如何利用“大数据”加强拥堵管理

卡尔森·哈格罗夫, 何大娜, 丹尼尔·康利, 皮特·纽曼  
澳大利亚 珀斯 科廷大学可持续发展政策研究所

**摘要:** 交通拥堵是世界各地交通规划者和管理者面临的一个关键问题, 许多人现在都在问是否有任何有前途的技术可以提供新的解决方案。在美国, 2012 年的拥堵成本为 1210 亿美元, 仅在 2015 年, 澳大利亚首都的拥堵成本就估计为 160 亿美元, 预计到 2030 年将增加到 370 亿美元。随着数据的可用性和分析大型数据集的能力的快速增长, 本文研究了“大数据在帮助拥堵管理方面可以发挥什么作用?” 人们对“大数据”产生了极大的兴趣和炒作, 本文总结了对其帮助缓解拥堵价值的调查。本文探讨了新兴类型的大型数据集, 考虑了未来车辆和交通基础设施将如何获取和共享数据, 并探讨了一些相关的挑战。尽管大数据的机遇尚未完全实现, 但很明显, 它为世界各地的交通规划者和管理者提供了一个重要的工具, 可以帮助他们管理拥堵。这项研究基于可持续建筑环境国家研究中心 (SBEnc) 的研究。

**关键词:** 大数据; 预测性拥塞管理; 技术支持的交通

## How Can “Big Data” Be Harnessed to Enhance Congestion Management

Karlson Hargroves, Daena Ho, Daniel Conley, Peter Newman  
Curtin University Sustainability Policy Institute, Curtin University, Perth, Australia

**Abstract:** Traffic congestion is a key issue facing transport planners and managers around the world with many now asking if there are any promising technologies offering new solutions. In the US, the cost of congestion was \$121 billion in 2012 and in 2015 alone Australia's capital cities were estimated to have a combined congestion cost of \$16 billion, expected increase to \$37 billion by 2030. With the rapidly growing availability of data and the ability to analyse large data sets this paper investigates the question “What role can ‘Big Data’ play to assist with congestion management?” There is great interest and hype around ‘Big Data’ and this paper provides a summary of an investigation into its value to assist in relieving congestion. The paper explores the emerging types of large data sets, considers how data will be sourced and shared by vehicles and transport infrastructure in the future, and explores some of the associated challenges. Despite the opportunities of Big Data not being fully realised it is already clear that it presents a significant tool for transport planners and managers around the world to assist in managing congestion. The research is based on research undertaken with the Sustainable Built Environment National Research Centre (SBEnc).

**Keywords:** Big data; Predictive congestion management; Technology enabled transport

### 1. 引言

大数据有多种定义。最常见的情况是, 该术语用于广泛描述数据集, 因为数据集太大, 无法通过传统的数据存储和处理方法进行存储和分析。现在可以从越来越多的来源获得大量数据, 然而, 这只是其复杂性的一个方面。接收数据的速度和可用信息的多样性增加了创造价值的挑战。此外, 数据现在以多种格式、语言和软件配置生成, 具体取决于数据的来源。

正是这三个特征(被称为三个 V——体积 (Volume)、速度 (Velocity) 和多样性 (Variety)) 将“大数据”与其他形式的数据区分开来。如此庞大而复杂的数据集的出现主要是由于传感和观测技术成本的降低, 以及全球系统

和过程的大规模数字化。结合全球分布式档案网络中的大量可用信息, 来自大规模传感器网络和计算机模拟的数据现在正在创造一个巨大的资源。

数据不仅可以用来观察直接的交通相关现象(如简化交通信号计时和匝道计量), 而且对传统和非传统数据流的询问提供了一种独特的潜力, 可以识别以前看似不相关的数据与交通相关活动之间的联系。由于产生的信息量非常大, “大数据”需要分析才能产生有意义的结果。“大分析”一词用于描述对多个海量数据集的处理, 以提取有用的算法和信息, 这些算法和信息可以可视化, 例如交通控制中心。

大数据有望超越所谓的“小数据”, 即交通计数、

平均速度、温度条件、红绿灯信号持续时间等数据,包括考虑数百个数据源,这些数据源可以为拥堵管理工作提供信息,例如直接从车辆流式传输的数据、停车场数据,有关公共交通的数据、可能影响交通的社会事件的数据、气象数据等,甚至以前认为与拥堵管理无关的数据。

考虑到拥堵的经济影响,使用新的数据流为交通管理和交通规划提供信息至关重要。减少拥堵有经济、环境和健康方面的好处,可以通过利用新的数据集大大增强这些好处。首先,减少高峰时段的拥堵推迟了所需的资本投资:如果高峰时段的交通不再是一个影响因素,就不必修建额外的道路或高速公路。此外,道路拥堵也有财务成本。在美国,2012年拥堵成本为1210亿美元,相当于每位通勤者每年818美元。<sup>[1]</sup>根据BITRE的一项研究,据估计,2015年澳大利亚首都城市的综合拥堵成本为160亿澳元,预计到2030年将增加到370亿澳元。<sup>[2]</sup>

此外,减少交通堵塞中的车辆等待时间可以减少车辆尾气,从而减少碳排放和空气污染。仅在美国,每年就有2500万吨二氧化碳被困在拥堵的道路上。<sup>[3]</sup>此外,长时间吸入汽车尾气也与脑细胞损伤等人类健康问题有关。<sup>[4]</sup>这些负外部性都表明了管理道路拥堵的需求增加,而数据的不断增加可能会提供部分解决方案。

## 2. 大数据和运输系统

### 2.1 “大数据”为交通系统创造的价值

利用数据为运输提供信息并不是一个新现象;交通系统长期以来一直产生观测和感官信息流,但在今天的术语中,这通常被称为“小数据”。2014年,澳大利亚政府基础设施、交通和区域经济局(BITRE)的一项研究确定了一些可用的数据收集技术,并得出结论:“最近和新兴的技术为收集更多关于个人旅行活动和道路使用的信息、更具成本效益提供了重要机会,这可以更好地为日常网络管理、长期基础设施规划和道路用户出行选择提供信息。”<sup>[5]</sup>

除此之外,越来越多的证据表明,交通管理人员正在利用越来越大的数据规模,从而实现更高水平的效率,从而节省成本,减少能源需求,更好地提供服务,提高生活质量,减少环境影响。在本文中,交通是指包括车辆、道路、铁路、地铁、公共汽车、出租车、自行车、渡轮和共享乘车在内的流动系统。每种交通方式在城市的流动性中都发挥着至关重要的作用,如果利用得当,可以以合理的成本安全高效地将人员和产品运送到目的地。

通过向公共交通模式转变和更好地管理道路网络来减少拥堵的努力带来了许多好处,例如:

- a. 由于在拥堵中损失的时间减少,城市的宜居性增强,
- b. 更快、更便宜的旅程,减少车辆和道路网络的磨损,
- c. 通过提供更好、更高效的流动性来吸引企业进入城市,

- d. 减少空气污染和温室气体排放等环境影响,以及
- e. 缓解城市交通预算压力,最大限度地提高昂贵交通资产的效益。

根据澳大利亚交通和资源经济局的数据,到2020年,澳大利亚“可避免的交通拥堵社会成本将上升至约200亿美元”。<sup>[6]</sup>

### 2.2. 收集大数据

数据存储选项(内部和远程)容量的快速增长,以及计算能力的提高,意味着利用现有和新数据源的额外价值具有巨大潜力,这些价值可以为现代城市的工作,特别是其交通系统提供信息。来自运输系统的数据种类繁多,分为三大类:

1. 高度结构化的数据集,源于为解决明确定义的问题而实施的技术(例如,用于处理收费公路付款的自动收费公路付款转发器的数据,或来自交叉口传感器的关于交通流量和道路网络使用时间的数据),这些数据集可以被视为“小数据”,直到其数量上升到无法使用传统方法进行分析。

2. 道路使用者与数字基础设施之间的任何互动产生的非结构化数据集。鉴于移动电话、个人电脑、传感器、摄像头和设备的使用激增,利用这些数据流来通知拥堵的潜力巨大(但在很大程度上尚未开发),而拥堵现在正进入“大数据”领域

3. 来自看似不相关的来源的数据,可以深入了解交通系统的行为和功能,例如特定公共停车场的停车价格、非法停车的罚款水平、步行超过1公里前往公共交通工具的人数、天气状况等,这些数据通常不包括在小数据技术中,或者需要由交通控制中心运营商从其他来源手动输入。

目前,对数据的访问并不令人担忧,因为有大量可用的数据源可以产生丰富的信息(然而,可能的情况是,目前不可用的额外数据集对拥堵管理可能比目前公开可用的数据集更有价值)。面临的挑战是通过在更高层次的趋势和情景以及与交通基础设施日常管理相关的较低层次上处理和解释数据来利用这些数据。高数据量意味着处理需要时间,并且需要先进的计算技术来提高响应时间。<sup>[3]</sup>目前,数据用于通知行程时间和路线选择;然而,大数据可以用于预测分析和高级用户信息平台的开发。

这种分析需要从多个数据流和历史数据库中提取价值的程序和技术,这些数据流和数据库中包含的数据似乎与运输脱节,但显示出的相关性在其他情况下会被隐藏。它是对可用数据流的综合“大分析”处理,其中存在大数据的真正潜在价值。<sup>[7]</sup>

### 2.3. 数据分析和分析

由于现代数据流和历史数据库的数量、速度和多样性,在分析和利用这些信息方面存在固有的挑战。特别地,存储数据的不同数据格式和语言可能导致使用数据挖掘算法进行处理的困难。<sup>[8]</sup>然而,潜在的回报是令人

印象深刻的。“大数据”的可用性提供了对实际乘客和道路使用行为的洞察，而不是报告的可能无法呈现全貌的行为和偏好。<sup>[9]</sup>

大数据的多层次性也使数据挖掘程序能够在监控摄像头、智能手机和地铁卡（智能卡）使用以及传感器等多个平台上找到相关性和趋同的旅行者偏好。<sup>[10]</sup> 这有助于制定考虑多种交通方式（私家车、公共交通、自行车等）的交通需求预测，形成对预期出行模式的全面概述，为长期交通规划和基础设施投资提供信息。

将“大数据”的价值传达给道路使用者、规划者和运营商对于改善交通网络至关重要。为了在拥堵变得严重之前减少道路拥堵，当实时拥堵预光标与严重拥堵事件的历史信息相匹配时，大型分析算法必须能够与红绿灯和其他交通控制系统进行通信。<sup>[11]</sup>

需求预测可以为规划者的交通基础设施投资提供信息，并确保实施的项目专门满足客户需求。就公共交通而言，提供实时交通状况和准确的等待时间估计大大提高了客户对服务有效性的感知。<sup>[12]</sup> 因此，必须权衡使用大数据的好处与大分析技术的有效性和效率，以及使用这些分析程序所产生的成本。

### 3. 利用数据通知拥堵管理的选项

有三种主要途径可以利用大数据为拥堵管理提供信息，即：

1. 通过实时数据使用缓解现有交通拥堵：来自各种来源（如交通信号灯、车辆计数器、闭路电视流等）的实时信息正被世界各地的许多城市用作拥堵管理的一种形式。目前，这项技术正在扩展，包括与车辆直接通信以发送和接收数据的技术。

2. 通过预测策略避免交通堵塞：利用多个数据流并与历史数据库进行比较，可以超越对小数据流的实时数据分析，使预测算法能够将实时数据与通勤习惯和首选路线的历史数据集相结合，从而实现预测性交通管理。通过选择指示符数据流和存储来增强这种平台的预测能力，然而这些并不总是显而易见的。

3. 创建复杂的公共交通路线：通过访问多个数据流和数据源，数字平台可以生成高分辨率信息，用于构建公共交通需求图，从而更好地分配公共交通资源。这一领域的研究仍然相对较新，然而，与旅行者为什么使用交通服务（如购物时间、公共活动、气候条件）相关的大量数据流和存储的融合潜力，可以使人们对影响乘客的因素有更深入的了解。

#### 3.1. 实时拥塞管理

目前，许多完善的拥堵管理策略都使用实时数据，尽管数据来源有限。存在多种类型的软件，可以快速响应交通量、交通移动需求和行驶方向的实时变化。在澳大利亚，目前有三种主要类型的软件用于通知交通控制系统，即 SCATS、STREAMS 和 InSync。<sup>[13]</sup> SCATS 在大多数首都城市使用，代表“悉尼协调自适应交通系统”，

该系统监测实时交通信号和车辆流量，以协调相邻的交通信号，减少交通拥堵并优化交通流量（可选择由控制系统操作员进行用户干预）。SCATS 的使用已被证明与减少总体行驶时间、车辆停车、油耗和红色交通信号灯等待时间相对应。<sup>[14]</sup> STREAMS 以 SCATS 系统为基础，已在昆士兰实施，取得了可喜的成果。

InSync 是另一种自适应交通控制系统，它使用安装在道路路口的摄像头来检测和管理交通状况。InSync 与 SCATS 的不同之处在于，它不使用循环长度、分裂和偏移的概念，而是使用有限状态机的概念，该概念由交叉点内的所有可能状态组成。这意味着，在任何给定时刻，都可以识别出特定的状态，这将导致适当的信号转换。西澳大利亚州主干道最近开发了一种名为 NetPRoS（网络性能报告系统）的工具，该工具集成了来自多个来源的数据，并根据多个指标报告道路网络性能。该工具目前仅限于历史性能，但预计将扩展到实时性能分析。

在数据可视化和报告方面，实时拥堵报告是如此普遍，以至于它变得无处不在，也令人期待。这方面最好的例子是“谷歌地图”方向应用程序，它以几乎所有第一世界道路用户都知道的易于访问的形式整合了实时拥堵信息。虽然实时拥堵缓解技术已在多个城市和国家广泛实施，但任何实时策略在改善交通状况方面的范围都有限。这主要是因为一旦观察到拥堵，就已经为时已晚。实时缓解策略通常基于通过交通信号或响应性道路收费来阻止额外的交通向该地区移动，但这两种策略都没有消除现有的拥堵。因此，人们现在非常感兴趣的是预测策略，这些策略试图在交通拥堵开始之前就遏制交通拥堵，以及数据流和存储如何为交通基础设施规划和投资提供信息。<sup>[15]</sup>

#### 3.2. 预测性拥塞管理

现在有人提出了一个问题，即“大数据”的使用如何有效地预测交通状况，简短的回答是“是的，但不是完全的，现在也不是”。长期的答案是，虽然从数百万个来源收集的高分辨率数据可能包含预测旅行模式和确定问题区域所需的信息，但挑战在于处理这些信息，以提取有用的信息和相关性，并与历史数据进行比较，从而对未来情况做出明智的预测。<sup>[16]</sup>

即使有这样的过程，考虑到交通条件也会受到非线性因素的影响，例如由交通动态的突然变化或非理性的人类反应引起的车辆碰撞，预测也不会是完美的。即使是在完美的几何和环境条件下完全安全的车辆，也可能由于道路动力学的突然变化、正常交通流的中断或车内的中断而发生碰撞。为了更好地改进拥堵预测模型，可以用实时数据代替聚集的交通流量变量（例如，假设车速等于限速），以创建更真实的模型。<sup>[17]</sup>

目前有许多“大型分析”流量预测系统正在开发中。这一领域的早期推动者是全球公司 HERE，该公司每天处理从 20 多亿个交通探测器收集的信息，并使用算法

将其与 2011 年以来的历史数据进行比较,以生成道路交通拥堵问题的预测。<sup>[18]</sup>微软还开发了各种预测交通状况的软件,一些软件平台考虑到了意想不到的交通状况,并取得了可喜的成果。<sup>[19]</sup>

### 3.3 公共交通规划与部署

公共交通系统越来越多地配备了自动化数据收集系统,这些系统可以与其他数据流一起使用,以深入了解乘客需求,并确定最佳的公共交通网络、路线和连接。<sup>[12]</sup>对这些数据的分析可以提供有关乘客需求和行为的信息,并提供对系统性能和实时条件的评估。此外,这种数据分析可以使道路和运输组织量化服务不足的成本。至关重要的是,量化还允许对可能的解决方案进行公平和基于模拟的评估(例如,时间表同步),允许根据其成本效益和用户体验效益对每个解决方案进行排名。<sup>[12]</sup>

公共交通网络对两个关键的大数据来源感兴趣:1)由手机提供的自动车辆定位(AVL)数据,以及2)由智能卡(地铁卡)、监控系统(即摄像机)、Wi-Fi和蓝牙跟踪器以及连接到资产、信号和开关的传感器提供的自动乘客计数(APC)数据。目前,AVL和APC数据都用于系统性能评估。然而,这两个数据源都没有在系统规划和开发中得到广泛使用,这使得这些数据在很大程度上没有得到充分利用。如果加以利用,这些形式的数据可以为客运量的预测提供信息,这对有效预测未来需求至关重要,并可以起到促进交通网络设计和优化的作用。<sup>[12]</sup>

此外,对AVL和APC数据等数据源的分析可以取代对旅行习惯和既定偏好的大规模、昂贵且经常被夸大的调查。算法可以根据观察到的出行模式直接构建出行需求,并为公共交通规划提供基础,如战术规划(涉及服务频率、时间表、车辆和人员调度的中期计划)和战略规划(一项长期计划,涉及整体网络和服务设计,如站点定位、线路拓扑和容量)。<sup>[20, 21]</sup>使用大数据不仅降低了交通调查的成本,还提供了更详细、更高分辨率的信息,如季节性影响、日内和日常需求变化,这些对时间表至关重要。<sup>[22]</sup>

人们对挖掘公共交通网络拓扑结构的方法非常感兴趣,这导致了支持这一点的工具的开发,例如“密度共识聚类”方法。<sup>[23]</sup>这种方法试图通过GPS时间序列数据推导公共交通网络的静态知识。所提出的方法能够生成静态数据,管理数据变化,并实时检查突然的弯路。该方法的创建者表示,使用该方法收集数据所需的基础设施规模小且成本低,包括一台主服务器和每辆车的车载单元。

使用大数据为公共交通系统提供信息需要收集和处理多个数据流,以输入预测算法。通过应用机器学习技术分析交通数据和公共交通车辆数据来生成信息,机器学习技术可以利用大量数据来揭示复杂的模式。服务中断也可以通过严重中断期间乘客行为的离线分析来缓

解,这使得运输线路可以调整到高需求地区。<sup>[12]</sup>最终,基于大数据构建的预测算法可以用于优化未来的交通网络,以及实时监控、调度和管理中断。

### 4. 技术赋能交通的未来

预计在不久的将来,车辆将越来越多地充当移动计算机,产生、处理和响应源源不断的输入数据,包括其他汽车和物体的位置、真实和预期的交通状况、停车场的开放时间、公共交通替代方案的客流量和出行时间水平,以及前往给定目的地的最佳路线。车辆将能够利用来自各种来源的数据,如其他车辆、道路状况和道路基础设施等。

具有这些功能的车辆被称为“联网车辆”,需要广泛接入多种通信服务才能使此类车辆有效运行。<sup>[24]</sup>世界各地的汽车制造商都在竞相将更高水平的技术嵌入汽车,目的是最终提供所谓的“自动驾驶汽车”,即不需要驾驶员的汽车。无论这一目标是否实现,或者事实上更可取的是,考虑到这将导致拥堵加剧,无论司机是否把手从方向盘上移开,都可以获得大部分好处。

联网车辆提供了实时利用高速车辆生成的信息流和来自相关基础设施的数据的潜力。为了澄清,有三种类型的车辆相关数据传输:车辆到车辆(V2V)、车辆到基础设施(V2I)和车辆到一切(V2X)。

#### 4.1. 车对车(V2V)

车对车通信允许车辆之间的信息传输,创造了组织和合作的潜力,以防止碰撞,并绘制网络之间的车辆信息地图,以减少拥堵。预计这种联网车辆可以根据从其他车辆和交通基础设施接收到的信息进行干预,以防止事故发生,例如在碰撞发生前制动,这可能是由驾驶员造成的,也可能是由于其他车辆的行为造成的,驾驶员在为时已晚之前无法看到。

许多汽车制造商目前正在测试五氧化二钒原型车,丰田公司于2016年宣布,将在密歇根州安阿伯市将其启用二氧化钒的车队测试规模增加到5000辆。<sup>[25]</sup>然而,由于梅赛德斯E级设计用于与其他E级车型通信,数据兼容性对行业来说是一个挑战。为了简化工作,美国运输部建议要求2021以后的所有新车都采用V2V技术,使用标准V2V频率,以实现交叉制造和模式通信。<sup>[26]</sup>

#### 4.2. 车辆到基础设施(V2I)

车辆到基础设施通信的一个例子是,当数据从乘坐车辆的移动电话发送到私人公司(如果特定软件处于活动状态),以进行聚合,从而提供交通出行时间的估计。然而,这仅仅是V2I概念能够实现的开始。在不久的将来,车辆本身将直接向交通基础设施发送包括多个变量的复杂数据流,用于拥堵管理、应急响应和预测分析。

反之亦然,基础设施可以直接与车辆通信以提供警报,指定最佳速度以减少拥堵和出行时间,并在应急车辆和公共交通车辆周围建立开放走廊。例如,奥迪在2017款奥迪Q7和A4车型上发布的红绿灯信息系统将

通知驾驶员红绿灯变绿之前的时间。<sup>[27]</sup> 第三方应用程序, 如“EnLighten”应用程序, 为驾驶者提供类似的信息, 利用交通信号计时来提供行驶速度建议, 目前正在宝马汽车上安装。<sup>[28]</sup>

车辆到一切 (V2E)

车辆到一切 (V2X) 结合了 V2V 和 V2I, 同时还与行人、设备和网络进行通信, 本质上允许车辆与道路网络上可能影响它的所有周围元素进行通信。V2X 的应用接近“大数据”比例, 并利用广泛的数据流和输入, 为道路用户提供信息, 以创建更安全、更高效的道路网络。

互联系统的关键点之一是一致性, 因为互联车辆应该具有兼容的技术, 以便道路上的车辆能够与基础设施和彼此通信, 无论其品牌和型号如何。在这种情况下, 国际技术标准协调的进展推迟了澳大利亚联网汽车的实施和采用。<sup>[29]</sup> 澳大利亚通信和媒体管理局正在与行业协商, 以制定“合作智能交通系统”(C-ITS)的授权制度。<sup>[30]</sup>

澳大利亚交通认证局和澳大利亚国家公路局也在开发一个系统, 以确保澳大利亚 C-ITS 系统的安全性、稳健性和可信度。<sup>[31]</sup> 新南威尔士州也已开始对联网车辆进行测试, 该州已在伊拉瓦拉为 60 辆装有 V2V 和 V2I 技术的参与重型车辆实施了 C-ITS 试验台, 该技术在 5.9 GHz 无线电频谱上进行广播。<sup>[32]</sup> 在一份内部报告中, 西澳州主干道还建议分阶段安装具有连接能力的路侧装置, 尤其是在规划的道路开发中。<sup>[33]</sup>

## 5. 隐私问题

移动相关数据的指数级增长将引发运输行业的重大变化, 同时人们对确保隐私的法规的充分性越来越担忧。<sup>[34]</sup> 即使是被称为“匿名”的数据, 如果与其他相关数据来源交叉引用, 也可能与特定的个人来源相关联, 特别是因为目前大部分数据都与私营公司共享, 几乎没有责任。交通管理中心不仅要解决这个问题, 还必须决定数据是否足够可靠。目前, 这些数据中的大部分都必须通过其他数据源进行验证, 如传感器和摄像机镜头或静态镜头。此外, 公司可能需要迁移到非关系型 (NoSQL) 数据库, 以容纳和处理大型非结构化数据集。这些 NoSQL 数据库通常使用外部安全强制机制; 因此, 为了减少安全漏洞, 公司必须使用额外的安全软件, 审查操作系统和 NoSQL 数据库之间的“中间件”的安全策略, 同时加强 NoSQL 数据库本身, 使其与对应的关系数据库相匹配。<sup>[35]</sup>

大数据的多层性质意味着事务日志存储在多层介质中。在较小的数据集中, IT 经理可以手动在层之间移动数据, 从而为他们提供一定程度的控制; 然而, 随着数据集呈指数级增长, 自动分层对于大数据存储管理可能变得越来越必要。由于自动分层无法跟踪数据的存储位置, 因此不太容易检测到对数据存储的未经授权的访问, 可能会发生数据泄露。因此, 必须开发新的机制来防止数据被盗并保持全天候可用性。<sup>[36]</sup>

在澳大利亚, 《隐私法》规范和保护个人信息, 包括《澳大利亚隐私原则》, 该原则定义了处理和评估个人信息的标准、权利和义务。大数据改变了关键隐私原则的应用方式, 包括数据收集、最大限度地减少数据保留和使用限制。然而, 由于应用程序在技术上是中立的, 企业和其他组织可以调整其大数据处理政策, 以保护个人信息, 同时最大限度地利用大数据分析得出的信息。<sup>[37]</sup>

根据《隐私法》, 组织必须采取合理措施实施保护个人信息的实践、程序和系统。这些组织还必须能够处理来自个人的与隐私相关的投诉。必须使用系统的风险管理方法, 根据实体的规模、资源和运营的复杂性确定合理的步骤。与处理大数据分析结果的实体相比, 直接处理大数据的组织必须采取更严格、更详细的隐私保护程序。<sup>[38]</sup>

确保隐私的一种可能技术是使用去标识来删除个人标识符, 如地址和出生日期, 以及任何其他独特的个人特征。这意味着《隐私法》不再适用。<sup>[38]</sup> 然而, 这项技术并非万无一失, 如果去识别的数据集与其他数据集或其他信息相匹配, 则有可能重新识别个体。<sup>[39]</sup>

## 6. 结论

随着大量数据在一系列平台上不断爆炸, 大数据在提供高分辨率信息以优化交通系统方面越来越重要。由于大多数交通管理系统都利用实时、专用的“小数据”, 大数据的令人兴奋的含义尚未完全实现。大数据在防止拥堵方面具有很高的潜力, 并取得了可观的回报, 但其实现并非没有挑战。尽管大数据可以提供关键信息来评估、规划和改进交通系统, 但其利用的关键挑战是大量信息需要多种数据分析和处理模式。由于有这么多信息可用, 必须开发软件和程序, 筛选出不相关的信息, 并将重点放在数据的关键特征上, 为交通预测模式提供必要的输入。

然而, 由于数据的规模、数据的多样性和快速频繁的变化, 集成、可视化、分析和响应查询是一项具有挑战性的任务。当前的数据分析系统提供的分析能力有限, 响应时间很长, 只有几分钟, 这是实时数据分析的障碍。最近, 已经发现内存计算技术可以实现显著更高的效率, 处理速度大约为 1 秒。多家 IT 公司正在这一领域积极开展工作, 研究人员目前正在研究提高处理器速度和响应能力的新方法。

此外, 在使用大数据进行未来运输量预测时, 必须使用高度专业化和精确校准的数据挖掘程序, 以制定准确和稳健的预测, 因为可用的信息量太大, 难以进行分析。使用大数据开发的算法和预测也必须根据实际运输量场景进行适当校准, 以确保预测的系统性能足够准确。

未来必须克服这些挑战, 才能准确、有效、高效地利用大数据进行拥堵管理和应急响应。然而, 总的来说, 交通规划者的胜利远大于他们的损失, 特别是加强交通系统管理系统, 及时预测何时会出现瓶颈, 使交通规划

者能够设计出防止这些地区拥堵的方法,从而有效地推迟了交通系统扩张的资本投资。

#### 参考文献

- [1] Texas A&M Transportation Institute 2013, 'As traffic jams worsen, commuters allow extra time for urgent trips', Texas A&M University, February 5.
- [2] Bureau of Infrastructure, Transport and Regional Economics (BITRE) 2015, Traffic and congestion cost trends for Australian capital cities, Commonwealth of Australia, Canberra. (ISBN 978-1-925216-99-8)
- [3] Mullich, J. 2013, 'Drivers avoid traffic jams with Big Data and Analytics', Bloomberg L. P., New York.
- [4] Hotz, R. L. 2011, 'The hidden toll of traffic jams', The Wall Street Journal, November 8.
- [5] BITRE 2014, 'New traffic data sources - An overview', Australian Department of Infrastructure and Regional Development, Canberra.
- [6] Transmax 2015, 'Streams ITS', Transmax Pty Ltd 2015.
- [7] International Transport Forum 2015, 'Big Data and Transport', International Transport Forum.
- [8] Hassaniien, A. E., Azar, A. T., Snasael, V., Kacprzyk, J., Abawajy, J. H. (ed) 2015, op. cit. ISBN: 978-3-319-11056-1.
- [9] van Oort, N. & Cats, O. 2015, 'Improving public transport decision making, planning and operations by using Big Data: Cases from Sweden and the Netherlands', IEEE 18th International Conference on Intelligent Transportation Systems. DOI 10.1109/ITSC.2015.1.
- [10] Carter, K. B. 2014, Actionable Intelligence: A guide to delivering business results with Big Data fast!, John Wiley & Sons, Singapore. (ISBN: 1118920651)
- [11] Sawyers, P. 2015, 'How Microsoft's using big data to predict traffic jams up to an hour in advance, Venturebest, April.
- [12] van Oort, N. and Cats, O. 2015, op. cit. DOI 10.1109/ITSC.2015.1.
- [13] Fernando, B., Gray, E., Kellner, J. 2013, 'A review of current traffic congestion management in the City of Sydney', Infrastructure Australia, Canberra.
- [14] Main Roads WA 2014, 'SCATS', Government of Western Australia, Perth.
- [15] Lu, H. P., Sun, Z. Y., Qu, W. C. 2014, 'Big Data-Driven Based Real-Time Traffic Flow State Identification and Prediction', Discrete Dynamics in Nature and Society, vol 1. (DOI: 10.1155/2015/284906)
- [16] Jie Xu et al. 2015, 'Mining the Situation: Spatiotemporal Traffic Prediction with Big Data', IEEE Journal of Selected Topics in Signal Processing, vol. 9, no. 4.
- [17] Hossain, M. 2012, 'A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways', Accident Analysis & Prevention, vol. 45. (<http://dx.doi.org/10.1016/j.aap.2011.08.004>)
- [18] Highway Engineering Australia 2015, 'Big data: The key to unlocking the future of traffic, transport and infrastructure', HEA, vol. 47, no. 2, pp. 40-41.
- [19] Horvitz, E. 2011, 'Predictive Analytics for Traffic', Microsoft Research, Sept 26.
- [20] Ma, X., Wu, Y. J., Wang, Y., Chen, F. and Liu, J. 2013, 'Mining smart card data for transit riders' travel patterns', Transportation Research Part C: Emerging Technologies, Vol. 36, pp. 1-12, <http://dx.doi.org/10.1016/j.trc.2013.07.010>.
- [21] Zhao, J., Rahbee A. and Wilson, N. H. M. 2007, 'Estimating rail passenger trip origin-destination matrix using automatic data collection systems', Computer Aided Civil and Infra. Eng., Vol. 22, pp. 376-387. (DOI: 10.1111/j.1467-8667.2007.00494.x)
- [22] Berkow, M., El-Geneidy, A. M., Bertini, R. L. and Crout, D. 2009, 'Beyond generating transit performance measures', Transportation Research Record, Vol. 2111, pp. 158-168. (DOI: 10.3141/2111-18)
- [23] Fiori, A., Mignone, A., Rospo, G. 2016, 'DeCoClu: Density consensus clustering approach for public transport data', Information Sciences, vol. 328, no. 1, pp. 378-388.
- [24] Weeratunga, K. and Somers, A. 2015, Connected vehicles: Are we ready?, Main Roads Western Australia internal report, Perth.
- [25] Asian Development Bank 2016, 'Safety and IntelligentTransport Systems Development in the People's Republic of China', Asian Development Bank, Section B. 8.4.
- [26] National Highway Traffic Safety Administration 2016, 'US DOT advances deployment of connected vehicle technology to prevent hundreds of thousands of crashes', United States Department of Transportation, Washington DC.
- [27] Herndon, Virginia. Audi USA Homepage, Press Release (August 15, 2016) "Audi announces the first vehicle to infrastructure (V2I) service - the new Traffic light information system".
- [28] Zurschmeide, J. 2015, 'Stop wasting gas in the city with the app that knows when traffic lights will change', Digital Trends.
- [29] ACMA 2016, 'Proposed regulatory measures for the introduction of C-ITS in Australia', Australian Communications and Media Authority, Australian

Government, accessed 25 Jan 2017.

[30] ACMA 2016, ‘Proposed regulatory measures for the introduction of C-ITS in Australia’ , op. cit.

[31] TCA 2017, ‘Cooperative Intelligent Transport Systems (CITS)’ , Transport Certification Australia, Australian Government, accessed 25 Jan 2017.

[32] Centre for Road Safety 2016, ‘Cooperative Intelligent Transport Initiative’ , NSW Government, accessed 25 Jan 2015.

[33] Weeratunga, K. and Somers, A. 2015, Connected vehicles: Are we ready?, op. cit.

[34] International Transport Forum 2015, ‘Big Data and Transport’ , op. cit.

[35] CSA 2012, Top ten Big Data Security and Privacy Challenges, Cloud Security Alliance, Rolling Meadows, Illinois.

[36] CSA 2012, Top ten Big Data Security and Privacy Challenges, op. cit.

[37] OAIC 2015, Consultation draft: Guide to Big Data and the Australian Privacy Principles, Office of the Australian Information Commissioner, Canberra.

[38] OAIC 2015, Consultation draft: Guide to Big Data and the Australian Privacy Principles, op. cit.

[39] OAIC 2016, Privacy business resource 4: De-identification of data and information, Office of the Australian Information Commissioner, Canberra.