

大数据驱动下智慧城市建设的若干思考

马新强^{1,2,3}, 刘勇², 范婧², 黄羿^{1,3}, 吴茂念⁴, 张明义¹

1. 贵州大学计算机科学与技术学院, 贵阳 550025
2. 浙江大学智能系统与控制研究所, 杭州 310007
3. 重庆文理学院机器视觉与智能信息系统重点实验室, 重庆 412060
4. 湖州师范学院信息工程学院, 湖州 313000

摘要 在当前的智慧城市建设中,随着互联网、物联网、通信网及大规模智能化应用服务的深入发展,产生了海量、异构、多源的城市时空大数据。在大数据时代,智慧城市的建设值得人们更深层次的思考。如何应对这些大规模、复杂的城市数据,尤其是对数据的感知、采集、存储、管理、分析、挖掘、计算、可视化及高效的服务应用都面临着巨大挑战。本文主要从技术、立法、管理3个视角重点阐述了大数据对智慧城市发展驱动作用。在技术上从纵向和横向两个视角分析城市时空大数据的特征及应用;从数据资产的角度强调了立法在智慧城市中的作用;从政策管理、奖惩与市场机制等方面分析了管理在智慧城市建设中的重要性;从综合视角思考大数据驱动下智慧城市建设中的若干问题。

关键词 智慧城市;大数据;城市计算;数据挖掘;可视化分析

近年来,随着城市化进程的深入发展,以传感技术为代表的物联网、智能移动设备为代表的互联网、通信网络和云计算技术,被广泛地应用于智慧城市的建设中。信息化、数字化、智能化的城市发展赋予人们现代化生活,不仅改变了社会及经济生活方式,同时也促进了人们对物理及认知空间、认知能力新的看法。然而,智慧城市带来便利的同时,也带来了许多挑战及问题,例如现代化城市中普遍存在的环境污染、交通堵塞、能源紧张、房价高涨、城市规划落后、人口老龄化、教育及医疗资源紧张等。这些城市发展过程中存在的问题严重影响了人们生活的宜居环境和城市的可持续性发展。目前,在大数据背景下,面对这些城市环境中极其复杂的挑战,如何解决这些必须解决的问题,更值得深度思考。

当前,伴随着信息技术的深度发展(尤其是移动互联网、智能感知、云计算、人工智能、深度学习等)和广泛应用,巨大信息流背后产生的海量、异构、多源城市数据已成为亟待提取、管理、分析、挖掘及有效利用的宝藏。这为解决现代化城市中的问题和挑战提供了机遇。政府、企业、社会和个人都希望利用深度发展的信息技术让城市变得更加智慧化,使人们的生活、经济和人文等得到全方位的提升。在大数据时代,利用高效、实时、融合的数据驱动模式不仅能发现城市环境中存在的问题,还能解决智慧城市建设中的挑战。近年

来,一批相关学者从不同角度进行了智慧城市的分析及相关研究。Hall、Li、王家耀等从智慧城市概念及大数据时代下的智慧城市构建方面进行了分析^[1-9];Zheng、徐静等从城市计算和智慧城市模型构建角度探讨了城市大数据的感知、采集、存储、管理、挖掘、计算及海量多源异构数据的融合^[10-12];Song、潘纲等利用人类在城市中的轨迹进行行为预测^[13-15];也有相关学者利用可视化分析技术进行了城市交通方面研究^[16-18]。以上相关研究大都从技术角度进行分析,而在大数据驱动下的智慧城市建设不仅依靠信息技术,还应依靠社会层面及其他领域的共同发展;应从多个角度综合使用数据,使其在智慧城市的建设过程发挥更大价值。

本文主要从城市中存在的大规模、异构、多源时空数据出发,从技术上讨论如何感知、计算、处理及应用数据,立法上思考数据的法理属性,管理上探讨数据的政策机制体制;通过探索大数据驱动下智慧城市建设中的若干问题及解决思路,为形成大数据驱动下的“技术、立法、管理”有益互补的综合性智慧城市服务体系提供决策支持。

1 相关概念及内涵

在讨论大数据驱动的智慧城市建设过程中,应明确以下几个相关概念及关系。

收稿日期:2017-03-23;修回日期:2017-04-20

基金项目:国家自然科学基金项目(61562011);浙江大学工业控制技术国家重点实验室开放课题(ICT170330);浙江省自然科学基金项目(LY16F020015);重庆市基础科学与前沿技术研究项目(cstc2016jcyjA0568, cstc2015jcyjA40026)

作者简介:马新强,副教授,研究方向为城市计算、健康医疗大数据,电子信箱:xinqma@163.com;刘勇(通信作者),教授,研究方向为机器学习、机器人,电子信箱:yongliu@iipc.zju.edu.cn

引用格式:马新强,刘勇,范婧,等.大数据驱动下智慧城市建设的若干思考[J].科技导报,2017,35(21):131-137;doi:10.3981/j.issn.1000-7857.2017.21.016

1.1 智慧城市

2000年, Hall等^[1]在第二届国际生命延长技术研讨会上最早提出了“Smart Cities”建设愿景, 2008年IBM公司首先提出智慧城市的理念, 2009年美国提出要推出智慧基础设施建设, 期望对社会经济发展带来新能量, 标志着智慧化城市建设新阶段的开始。此后中国、欧洲各国、日本等国家也提出智慧城市、智慧地球建设等概念。在学术界, 国外相关学者主要从城市规划、大数据技术、智慧城市整体架构、云计算、物联网、数据隐私安全保护、城市管理及政策等方面进行了智慧城市研究^[1-5], 在中国由中国科学院院士、中国工程院院士、国际欧亚科学院院士李德仁等为代表的一批学者相继提出关于智慧城市基础建设的相关研究^[6-8], 他提出了智慧城市的基础架构, 如图1所示, 并认为智慧城市 = 数字城市 + 物联网 + 云计算, 即在数字城市的基础上有机地融合物联网和云计算技术, 以实现对现实城市中人和物的自动控制和智能服务, 同时也从技术视角分析了智慧城市中的大数据如何建设问题。中国工程院院士王家耀认为大数据时代的智慧城市建设, 首先应从城市信息化入手, 围绕数字时代的智慧城市建设; 其次回答了什么是“智慧城市”的概念, 以及为什么要建、建一个什么样和怎样建设智慧城市等问题^[9]。王家耀强调智慧城市是城市信息化的高级阶段。人们应明确智慧城市与智能城市的区别。智慧是针对复杂、综合的大系统, 而智能、数字化是针对某项技术、某种高级仪器设备的。显然, 现代化城市是一个大系统, 是人、自然、社会、经济、人文等的综合体。

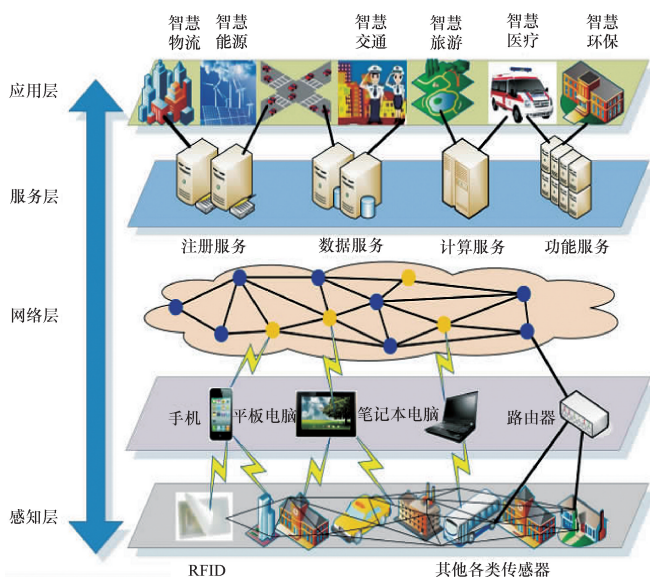


图1 智慧城市的总体架构

Fig. 1 Architecture of smart city

智慧城市建设中不能目中无“人”^[10], 人是城市的主体, 正是城市中各种人类活动的轨迹和人作为各种数据源的传感载体, 赋予了智慧城市基础数据真正价值。业界专家也指出, 若从“人-物-环境-自然”城市一体化角度看, 目前的智慧城市建设更重建设轻标准、规范和管理, 即智慧城市的标准、

规范及管理体系没有建立, 缺乏完善的行业、智慧化城市标准规范。在大数据时代, 为了更好地促进智慧城市建设发展、有限制地开放数据、保护各种数据安全和个人隐私, 需要真正的标准规范、法律法规及有效的管理机制。

1.2 大数据及其对智慧城市的影响

众所周知, 城市是人类活动最为密集的区域, 大规模的人类活动与城市中各个行业的运行数据不可避免地在城市当中爆发。城市数据具有典型的大规模(volume)、数据类型多样化(variety)、产生速度快(velocity), 同时由于城市时空数据本身的稀疏性导致数据价值密度低(veracity), 可以说城市数据是一种典型大数据样本, 完全符合大数据的4V特性。传统的技术处理方法和思维模式, 已经无法适应和处理当前的海量、异构、快速的城市大数据了。而大数据、数据活化、数据挖掘等数据管理、应用与分析技术在智慧城市建设当中却具有核心作用^[20]。

正是“大数据时代”的到来进一步推动和支撑了“智慧城市”, 改变了人们对城市信息化建设的认识, 加速了由“数字城市”到“智慧城市”的转变。在“大数据时代”, 任何人都必须用数据来说话, 因为大数据的本质是要用“大数据思维”去发掘“大数据”潜在价值。当前, 最重要的是学会驾驭大数据, 这涉及智能感知技术、分布式存储技术、智能统计分析和数据挖掘技术、智能化实时动态可视化技术、云计算技术及基于网络的智能服务技术等, 而这些也正是“智慧城市”建设涉及的关键技术。智慧城市建设, 必然产生大数据, 大数据的应用必将推进智慧城市, 大数据时代的到来和智慧城市的兴起, 是全球信息化发展到高级阶段的必然趋势^[9]。

1.3 智慧城市与城市计算的关系

在2003年就提出了城市计算这个概念, 但是并没有引起人们的重视, 正是由于大数据驱动下智慧城市建设的发展, 近年来城市计算逐渐成为研究热点^[20-24]。微软亚洲研究院主管研究员郑宇研究团队认为: 城市计算是一个通过不断获取、整合和分析城市中多种异构大数据来解决城市所面临的挑战(如环境恶化、交通拥堵、能耗增加、规划落后等)的过程, 同时提出了一个城市计算框架, 如图2所示^[10-11]。城市计算将无处不在的感知技术、高效的数据管理和分析算法, 以及新颖的可视化技术相结合, 致力于提高人们的生活品质、保护环境和促进城市运转效率。城市计算帮助人们理解各种城市现象的本质, 甚至预测城市的未来。

在大数据时代下的智慧城市建设, 不是单纯地凭经验对各种数据源进行简单的信息化、数字化, 而是面对大规模、异构、多源的数据, 利用有效的技术手段进行智能“计算”, 目的是为了更有效地对数据存储、管理、分析和决策支持。从智能计算技术的角度看, 图1中智慧城市的总体架构分为感知层、网络层、服务层、应用层, 这4层——对应城市计算中的感知计算、存储与管理、分析与挖掘、服务与应用。城市计算是智慧城市建设的必备核心计算技术, 智慧城市系统是城市计算技术实施的有效载体。在大数据时代, 城市计算技术将起

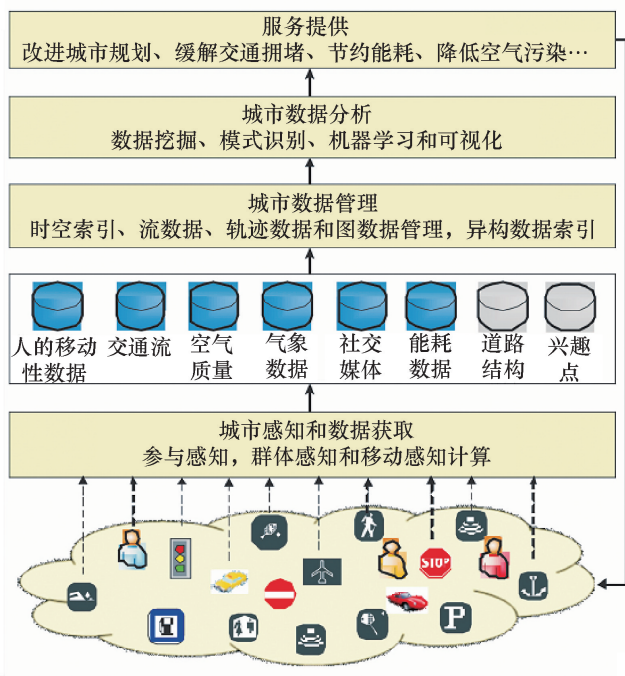


图2 城市计算框架

Fig. 2 Framework of urban computing

到核心作用,因为传统的单一处理技术已经无法有效处理当前的城市数据了,二者紧密关联、互为依存。

2 大数据驱动下的智慧城市技术支撑

中国科学院院士、西安交通大学教授徐宗本在2016年3月15日《人民日报(理论版)》发表的《用好大数据须有大智慧》一文中指出:实施国家大数据战略,必须理性认识大数据、准确把握其带来的机遇,科学应对其带来挑战,用大智慧实现大数据的大价值;突出强调了大数据的价值主要通过大数据技术来实现^[25]。在本文第1节中提到的智慧城市、城市计算概念中都强调了大数据技术的重要性。国内外企业、学者等都都对大数据开源技术及趋势做了相关研发^[2,11,26-27]。下面主要从大数据处理技术的采集、存储管理、分析挖掘、可视化展示及应用的纵向角度,以及从城市数据的属性及类型等横向角度做探讨。

2.1 纵向角度分析城市大数据处理技术

徐宗本院士认为:大数据技术是基础性信息技术,它刻画了新一代信息技术中机器与机器、人与机器之间信息交换的内容特征,构成了现代信息技术的基本信息处理模式^[25]。本节针对城市中各种海量、异构、多源的数据并结合目前先进的大数据处理前沿技术,探索和设计了一种大数据驱动下的智慧城市技术框架,如图3所示。

本技术框架的总体思路:采用目前先进的分布式并行计算模式 Lambda 混合架构,即同时进行历史数据批处理和实时数据处理的混搭模式。实时数据进入一个流处理系统进行检测分析,同时也进入 Hadoop 集群,全量数据放在 Hadoop



图3 大数据驱动的智慧城市技术框架

Fig. 3 Big data-driven for technology framework of smart city

中的分布式文件存储系统(HDFS)里。一是对历史数据进行分析,二是对采集的实时数据进行分析,做到两者数据的融合处理,然后应用程序利用这个“大数据库”做城市中各个行业主题库(例如各种轨迹大数据、交通领域大数据、城市居民的消费大数据、健康医疗大数据、社交媒体大数据、气象环境和空气质量大数据、城市地图和地理信息大数据、制造业工业化大数据等领域)的分析、挖掘、智能计算和可视化展示等,做到了同时处理实时数据和历史数据。纵向处理数据方式遵循“数据采集→数据存储→数据管理→数据分析和挖掘→数据可视化及应用”的多层次技术路线。

本框架既能按照某个行业领域主题构建智慧城市下的“行业私有云”大数据处理技术平台,又能创建某个城市多种数据源公共大数据技术平台。目前更倾向于某个行业下的大数据技术平台,这样更有利于行业大数据的存储管理和垂直主题的挖掘、分析;既能有效保护行业数据资产又能提供精细化、深层次的应用服务。在此,大数据驱动下的智慧城市技术框架中的具体技术细节不再阐述,但整个技术框架做到了以海量、异构、多源大数据为基础,先进的分布式并行计算及前沿技术为支撑,应用为导向,监测和安全为保护手段;同时也会考虑标准化的行业规范、法律、法规和有效管理机制作为保障。

2.2 横向角度分析城市大数据处理技术

大数据驱动下整个智慧城市建设中,应充分考虑到各个不同行业领域数据、数据类型、数据属性。例如针对城市中各种移动轨迹数据进行预测研究,就非常复杂,如图4所示^[15]。若按照城市数据类别、对象分为:各种用户个体、城市用户群体和由城市居民形成的区域性社会数据等;按照城市数据属性又分为时间、空间、时空数据或者离散型和连续型的数值数据等。

针对不同的城市数据类别、对象和属性,采用的处理技术会有异同,有些应用需求强调“时效性”。例如针对支持快速实时查询与处理的智能交通监控数据,可采取分布式的并



图4 轨迹预测框架^[15]

Fig. 4 Framework of trajectory prediction

行计算、Spark 内存计算和实时流数据处理模式 Spark Streaming 等。

面对城市数据源的多元化、异构性、弱关联性等特点,需要能够处理各种数据源的融合计算技术。例如针对城市中某地区环境污染治理预测,不仅要考虑城市中的气候、气象数据、环境污染数据,同时也应考虑该地区的建筑构成、兴趣点(POI)分布、交通工具流量、城市路况、人流量的突变等数据融合计算^[24]。

面向城市中的各种视频、音频及图形图像的非结构化数据,需要更加智能的计算处理技术。例如城市中公共安全视频监控、运营商的电话呼叫语音数据、健康医疗行业影像大数据等,都需要利用先进的人工智能算法、深度学习和新的机器学习算法来处理。例如在 2015 年 12 月出版的《Science》^[28]上,Chen 等提出类人概念学习,并引起了机器学习领域的广泛关注^[29-30]。

面向城市时空数据的高维、稀疏性以及传统方式下对决策支持的评价尺度单一、缺乏深度等问题,需要新的大数据技术来有效处理;这里强调自动化的可视化技术(automated analytical approaches)^[18]。该技术的实质是将数据以形象直观的方式展现,以迭代求精的方式将数据的复杂度降低到人类和计算机能处理的范围,以便以视觉理解的方式获取数据中蕴含的信息,发现非结构化、非几何的抽象数据背后的本质问题。可视化分析展示及应用对于决策支持的评价是非常有效、直观的工具。目前,已有部分相关工作,例如:可视化分析在城市计算中的应用研究^[18],大城市地铁网络客流变化及微博的可视化探索^[31],基于电信数据在城市人类活动共现的可视化分析^[32]等。

在城市大数据处理技术中,也会遇到一些问题和挑战。例如:城市中异常现象的检测,如何预防交通拥堵,自然灾害及公众事件引起的人流突变及预测;如何在高密度公共场所,例如地铁、广场,估计人群数量和密度,同时检测人群过

密、异常聚集、滞留、逆行、混乱等多种异常现象。又如:实现重大活动、重要区域的人流统计与控制,如何提供实时报警功能;在城市交通视频、图形图像处理过程中如何适应各种光照变化、人群遮挡等复杂环境等。这些问题都需要更有效的大数据处理及智能化分析计算技术作为支撑。

3 数据立法与管理保障分析

3.1 立法角度探讨城市数据

数据安全性和隐私保护问题一直是智慧城市建设过程中不可避免的重大问题:大数据时代个人、用户、企业、政府等有没有隐私,数据的所有权及使用权问题,即谁有权来处理 and 拥有数据。照片的肖像权或产生电子底片数据图像应归谁所有。另外,政府机构数据泄露事件后果非常严重,例如 2015 年 2 月 1 日德国援引美国国家安全局前承包商雇员爱德华·斯诺登爆料文件报道,美国情报机构正致力于准备网络战争,此前被曝光的大规模监听活动只是其中一部分。

关于数据安全性及隐私保护问题,也有相关学者进行了探讨,例如:个人数据的大数据应用需要构建溯源机制^[33],关于大数据交易核心法律问题^[34],医疗大健康及大数据的应用及其隐私保护分析^[35],大数据安全与隐私保护^[36],数据访问控制及用户授权在云存储中的研究^[37]等。其中中国信息通信研究院互联网法律中心高级工程师王融认为:原始或底层的个人数据,所有权归用户本人所有;而在原始数据基础上,经过充分匿名化获得的数据集,企业享有限制性的所有权。她强调了有“限制性”的开放数据。

大数据安全性及隐私保护问题,目前的研究存在一些问题和难点。

1) 数据产权归属问题界定不清晰。当前,对去除了个人身份属性(数据脱敏、匿名化)后的数据,在数据交易中到底谁是数据主体,是产生数据的个人,还是记录数据的企业拥有数据的所有权。目前,在法律层面上并没有给出明确的界定,当数据使用或交易过程产生分歧后,无法进行有效判属。这样的后果导致用户、企业、政府不愿意开放数据,对数据的来源造成很大困难。

2) 法律层面上,对数据本身应有的权利属性宣传不到位,导致公民的数据维权意识淡薄。例如:数据的可携带权、数据的被遗忘权,大多数用户对其含义不清楚。

3) 数据保护意识淡薄。例如:在城市的社交媒体领域,个人用户可能泄露了自己各方面信息,但是自己并没有意识到。数据企业厂商利用大数据分析挖掘技术根据用户在社交平台上“点赞的行为”,判断用户的性别、年龄、职业、感情状况、兴趣取向等,如果该信息被不法分子利用,会造成非常严重的后果。同样,在城市时空数据的各个行业领域中,也存在用户的私密信息泄露、篡改等,例如健康医疗数据、个人身份信息、企业商业数据、政府部门相关保密数据被攻击、被偷窃、损坏等。

4) 技术保护手段单一,缺乏有效的保护方式,导致大数

据时代、云环境下的数据安全技术相对滞后。

笔者认为,针对智慧城市中的大数据安全性及隐私保护,不仅仅是一个单纯的技术问题,更是一个法制、道德和标准化的问题,应当引起立法、城市管理部门及用户的高度关注。提出以下一些建议及措施。

1) 大数据相关的立法及制度制定时,不应只有法律界人士,还应有大数据相关技术人员和城市数据产生的相关行业领域人员参与。

2) 加大数据相关法规宣传力度,明确数据法理属性概念,增强公民的数据维权意识。例如,数据的可携带权是指公民享有要求服务商提供个人数据迁移便利的权力;数据的被遗忘权是指用户有权要求搜索引擎删除过时的和自己有关的搜索结果。

3) 利用大数据分析技术进行预防犯罪,可针对群体进行防范,但不宜针对个人,避免个人隐私信息泄露。

4) 明晰数据产权归属。只有具备了清晰的产权归属才是数据交易的前提与基础,个人对原始数据具有所有权(数据拥有权)、享有数据可携带权、数据被遗忘权,相关数据企业在经过充分匿名化的数据集上享有有限的所有权和使用权(数据收集的限制、数据使用的限制)。在智慧城市的建设中,政府应加强立法和执法监管力度,保护城市数据安全和用户隐私。

5) 在充分理解数据的基础上应用大数据相关技术和各种信息安全技术进行数据保护。例如城市数据私有云存储安全机制、数据安全访问控制技术、数据匿名技术、数据安全加密级别、密钥分配机理、能效的加密算法、可接入性认证、可信性认证、数据完整性认证、隐私增强技术、身份管理等。

3.2 管理角度探讨城市数据

在整个智慧城市建设过程中主体应是市场机制驱动下的各个行业及城市居民,政府、城市的各个管理部门及城市规划者应是出台各种标准化行为准则规范、完善市场机制的客体。在不违背政府部门基本的公共安全和国家安全的前提下,由政府管理者、立法专家及大数据技术专家共同制定规范化、标准化的大数据市场奖惩机制,这种奖惩机制应充分考虑“互联网世界、电子虚拟世界”的城市数字化、智慧化的特殊需求,均衡各方的利益分配,尽最大可能开放城市时空数据,目的是提高城市的智慧化管理水平。针对城市的行业管理部门或城市规划者,为了更有效地管理和使用相关大数据,需高效、快速地挖掘和分析相关数据价值。笔者认为构建行业主题的“数据资产私有云”或“行业云”,这样对于数据资产的有效管理更加精细化、对于发挥垂直行业内数据价值的最大化更有益。同时,也提供各类应用服务技术接口(API)及灵活多变的对接方式等,方便其他相关行业数据的远程调用和融合,提升各类数据的增值服务,有效应对城市的各种社会问题。

总之,明确政府、企业及个人对数据的政策管理及奖惩措施,形成积极有效的数据驱动下的保障机制,目的是为了

使政府、社会、企业和个人更好地开放和使用数据,同时保护数据安全和公民隐私,为政府和管理部门提供高效、科学、合理的决策支持服务。

4 讨论

智慧城市的建设是一个极其复杂、系统性的工作,当前由于各种信息化技术的深度发展产生了海量、异构、多源的城市大数据。面对这些城市大数据,如何更好地利用、深度挖掘与分析,解决城市中存在的各种社会问题,同时为城市发展及市民提供更加智能化便利生活服务是政府、城市规划者、学术界、工业界及每个市民值得深思的问题。

在大数据时代,笔者提出数据驱动下智慧城市建设的几点思考或建议如下。

1) 确立数据驱动思维模式,以大数据为基础,先进前沿技术为支撑,应用为导向,规范的立法和高效的管理机制为保障,形成智慧城市建设中“技术、立法、管理”三位一体的综合服务体系,为解决城市中的各种问题开辟新的思路。

2) 明确构建城市行业私有云,遵循行业标准化、规范化建立行业大数据基础平台,并充分利用人工智能、深度学习、机器学习、数据挖掘等前沿技术挖掘、分析行业垂直领域数据价值,形成城市行业数据资产库。

3) 从法律、法规层面明确城市大数据的法理属性,正如在第3.1节中给予的建议及措施;只有界定了城市数据的各种所有权及使用权,政府、企业、社会和个人才能透明化地利用数据、知晓数据,减少各种不必要的数据纠纷、违法行为和非法事件等。

4) 数据驱动技术和市场奖惩机制有益结合,从政府和管理部门层面不仅对数据的有效使用、挖掘数据价值上给予奖励机制,同时更要加大违法使用数据、滥用数据的监管和惩罚力度。

5) 目前,智慧城市中数据还面临一系列挑战:如多源、爆炸式的数据如何快速有效地采集、存储与高效管理,复杂数据本身的价值如何分析、挖掘及应用,数据以何种形式展示更有利于用户理解,大数据时代的隐私问题,大数据的所有权及使用权问题等。

在信息化深度发展的机遇及挑战面前,利用大数据驱动的思维模式来解决智慧城市建设过程中一系列的各种问题,可能会是一种好的解决方案。

5 结论

从数据的技术支撑、数据相关立法、数据的政策管理及奖惩机制等方面,探讨了大数据驱动下智慧城市建设中若干问题,提出相关建议,同时给出了一些解决思路。

目前从智慧城市建设的全面性角度看,要达到“人-物-环境-自然”城市一体化、智慧化共赢的结果,仅仅依靠技术支撑还远远不够,还需要社会化的立法、高效的管理作为有效补充。为了使政府、社会、企业和个人更好地开放和使用

数据,同时保护数据安全和公民隐私等,应从单一的“技术驱动”向多元化的“数据驱动”转变,利用信息技术获取、处理和挖掘分析数据,为更好的城市化服务提供数据基础;从法理性角度明确数据界定,尤其是数据所有权、使用权的归属性问题;从管理的视角明确政府、企业及个人对数据的政策管理及奖惩机制。

参考文献 (References)

- [1] Hall R E, Bowerman B, Braverman J, et al. The vision of a smart city[C/OL]. [2017-04-10]. <https://ntl.bts.gov/lib/14000/14800/14834/DE2001-773961.pdf>.
- [2] Batty M. Big data, smart cities and city planning[J]. *Dialogues in Human Geography*, 2013, 3(3): 274-279.
- [3] Martinez-Balleste A, Perez-Martinez P A, Solanas A. The pursuit of citizens' privacy: A privacy-aware smart city is possible[J]. *IEEE Communications Magazine*, 2013, 51(6): 136-141.
- [4] Petrolo R, Loscrì V, Mitton N. Towards a smart city based on cloud of things, a survey on the smart city vision and paradigms[J]. *Transactions on Emerging Telecommunications Technologies*, 2015, 28(1), doi: 10.1002/ett.
- [5] Nam T, Pardo T A. Smart city as urban innovation: focusing on management, policy, and context[C]/Proceedings of the International Conference on Theory and Practice of Electronic Governance. New York: ACM, 2011: 185-194.
- [6] Li D R, Cao J J, Yao Y. Big data in smart cities[J]. *Science China Information Sciences*, 2015, 58(10): 1-12.
- [7] Li D R, Yao Y, Shao Z F, et al. From digital Earth to smart Earth[J]. *Science Bulletin*, 2014, 59(8): 722-733.
- [8] Li D R, Shao Z F, Zhou X R, et al. Geomatics for smart cities—concept, key techniques, and applications[J]. *Geospatial Information Science*, 2013, 16(1): 13-24.
- [9] 王家耀, 邓国臣. 大数据时代的智慧城市[J]. *测绘科学*, 2014, 39(5): 3-7.
Wang Jiayao, Deng Guochen. Smart city in the era of big data[J]. *Science of Surveying and Mapping*, 2014, 39(5): 3-7.
- [10] Zheng Y, Capra L, Wolfson O, et al. Urban computing: Concepts, methodologies, and applications[J]. *ACM Transactions on Intelligent Systems & Technology*, 2014, 5(3), doi: 10.1145/2629592.
- [11] Zheng Y. Introduction to urban computing [J]. *Geomatics & Information Science of Wuhan University*, 2015, 40(1): 1-13.
- [12] 徐静, 陈秀万. 基于信息空间理论的智慧城市模型构建[J]. *科技导报*, 2013, 31(22): 56-59.
Xu Jing, Chen Xiukai. Construction of smart city model based on information space theory [J]. *Science & Technology Review*, 2013, 31(22): 56-59.
- [13] Song C M, Qu Zehui, Blumm N, et al. Limits of predictability in human mobility[J]. *Science*, 2010, 327(5968): 1018-1021.
- [14] 潘纲, 李石坚, 齐观德, 等. 移动轨迹数据分析与智慧城市[J]. *中国计算机学会通讯*, 2012, 8(5): 31-37.
Pan Gang, Li Shijian, Qi Guande, et al. Moving trajectory data analysis and smart city[J]. *Communications of the CCF*, 2012, 8(5): 31-37.
- [15] 孙未未, 毛江云. 轨迹预测技术及其应用——从上海外滩踩踏事件说起[J]. *科技导报*, 2016, 36(9): 48-54.
Sun Weiwei, Mao Jiangyun. Trajectory prediction technology and its application from the the bund stampede in Shangha[J]. *Science & Technology Review*, 2016, 36(9): 48-54.
- [16] Zheng X H, Chen W, Wang P, et al. Big Data for social transportation [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 17(3): 620-630.
- [17] Chen W, Guo F Z, Wang F Y. A survey of traffic data visualization [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(6): 1-15.
- [18] Zheng Y X, Wu W C, Chen Y Z, et al. Visual analytics in urban computing: An overview [J]. *IEEE Transactions on Big Data*, 2016, 2(3): 276-296.
- [19] 眉间尺. 智慧城市建设不能目中无“人”[N]. *科技日报*, 2016-04-15.
Mei Jianchi. Construction of smart city unable to people[N]. *Science and Technology Daily*, 2016-04-15.
- [20] 王静远, 李超, 熊璋, 等. 以数据为中心的智慧城市研究综述[J]. *计算机研究与发展*, 2014, 51(2): 239-259.
Wang Jingyuan, Li Chao, Xiong Zhang, et al. Survey of data-centric smart city[J]. *Journal of Computer Research and Development*, 2014, 51(2): 239-259.
- [21] Paulos E, Goodman E. The familiar stranger: Anxiety, comfort, and play in public places[C]/Conference on Human Factors in Computing Systems, CHI 2004. New York: ACM, 2004: 223-230.
- [22] Zhang J B, Zheng Y, Qi D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C/OL]. [2017-04-10]. <https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/download/14501/13964>
- [23] Zheng Y. Trajectory data mining: An overview [J]. *ACM Transactions on Intelligent Systems & Technology*, 2015, 6(3): 1-29.
- [24] Zheng Y. Methodologies for cross-domain data fusion: An overview [J]. *IEEE Transactions on Big Data*, 2015, 1(1): 16-34.
- [25] 徐宗本. 用好大数据须有大智慧[N]. *人民日报*, 2016-03-15.
Xu Zongben. Good use of big data must have great wisdom[N]. *People's Daily*, 2016-03-15.
- [26] 程学旗, 靳小龙, 杨婧, 等. 大数据技术进展与发展趋势[J]. *科技导报*, 2016, 34(14): 49-59.
Cheng Xueqi, Ji Xiaolong, Yang Jing, et al. Technological progress and trends of big data[J]. *Science & Technology Review*, 2016, 34(14): 49-59.
- [27] Chen C L P, Zhang C Y. Data-intensive applications, challenges, techniques and technologies: A survey on big data[J]. *Information Sciences*, 2014, 275(11): 314-347.
- [28] Lake B M, Salakhutdinov R, Tenenbaum J B. Human-level concept learning through probabilistic program induction [J]. *Science*, 2015, 350(6266): 1332-1338.
- [29] 刘勇, 廖依伊. 类人概念学习: 机器学习下一个飞跃?[J]. *科技导报*, 2016, 34(7): 80-81.
Liu Yong, Liao Yiyi. Human-level concept learning: the next leap in machine learning[J]. *Science & Technology Review*, 2016, 34(7): 80-81.
- [30] 陈孝良. 小样本的类人概念学习与大数据的深度强化学习[J]. *科技导报*, 2016, 34(7): 82-84.
Chen Xiaoliang. Small sample of human-level concept learning and deep reinforcement learning of big data[J]. *Science & Technology Review*, 2016, 34(7): 82-84.
- [31] Itoh M, Yokoyama D, Toyoda M, et al. Visual exploration of changes in passenger flows and tweets on Mega-City Metro Network[J]. *IEEE Transactions on Big Data*, 2016, 2(1): 85-99.
- [32] Wu W C, Xu J Y, Zeng H P, et al. TelCoVis: Visual exploration of co-

- occurrence in urban human mobility based on telco data[J]. IEEE Transactions on Visualization & Computer Graphics, 2016, 22(1): 935-944.
- [33] 王忠. 个人数据的大数据应用需要构建溯源机制[J]. 科技导报, 2014, 32(4): 12.
Wang Zhong. Construct of traceability mechanism for big data application of personal data[J]. Science & Technology Review, 2014, 32(4): 12.
- [34] 王融. 关于大数据交易核心法律问题——数据所有权的探讨[J]. 大数据, 2015, 1(2): 49-55.
Wang Rong. Discussion on the legal core question of the data ownership in big data trade[J]. Big Data Research, 2015, 1(2): 49-55.
- [35] 王黎洲. 医疗大健康及大数据的应用及其隐私保护分析[J]. 中国卫生产业, 2016, 13(13): 120-122.
Wang Lizhou. Analysis of the application of medical one health and big data and their privacy protection [J]. China Health Industry, 2016, 13(13): 120-122.
- [36] 冯登国, 张敏, 李昊. 大数据安全与隐私保护[J]. 计算机学报, 2014, 37(1): 246-258.
Feng Dengguo, Zhang Min, Li Hao. Big data security and privacy protection[J]. Chinese Journal of Computers, 2014, 37(1): 246-258.
- [37] Yang K, Jia X H. Expressive, efficient, and revocable data access control for multi-authority cloud storage[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 25(7): 1735-1744.

Big data-driven smart city: Some thoughts

MA Xinqiang^{1,2,3}, LIU Yong², FAN Jing², HUANG Yi^{1,3}, WU Maonian⁴, ZHANG Mingyi¹

1. College of Computer Science and Technology, Guizhou University, Guiyang 550025, China
2. Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310007, China
3. Key Laboratory of Machine Vision & Intelligent Information System, Chongqing University of Arts and Sciences, Chongqing 412060, China
4. School of Information Engineering, Huzhou University, Huzhou 313000, China

Abstract In the construction of smart city, the depth development of Internet, Internet of things, communication networks and large-scale intelligent services will result in massive, heterogeneous, multi-source big spatio-temporal urban data. In the era of big data, we should think deeply about the construction of smart city. How to deal with these large-scale and complicated urban data is facing enormous challenges, including data collection, sensing, storage, management, mining, visualization and efficient services. In this paper we mainly discuss the driving effects of big data on smart city in the views of technology, legislation and management. We analyze the features and applications of big spatio-temporal urban data in two directions, point out the importance of legislation from the perspective of data assets, and summarize the management of smart city in terms of policy management, rewards and punishments, and market mechanisms. Finally, we put forward some thoughts on the construction of smart city driven by big data.

Keywords smart city; big data; urban computing; data mining; visual analysis

(责任编辑 刘志远)