项目简介：

随着我国社会经济的快速发展和城市化进程的加快，道路交通己成为制约城市发展的重要问题，解决道路交通瓶颈己成为各级政府部门需要解决的首要任务。多年来，国内外实践经验证明，在解决城市交通问题方面，除了要有相应的宏观交通政策予以支持，现代化的交通管理手段也是十分重要的一环。与此同时计算机技术、电子技术等高新技术的发展，也为交通管理提供了解决交通问题的新思路和新手段，为最大限度地发挥现有道路系统的工作效率提供了技术支持。发达国家交通发展实践表明，采用智能交通系统是解决交通拥堵减少交通事故、防止交通污染、提高交通管理水平的最有效的方法和手段。

本项目的总体目标是运用计算机技术、电子技术等高新技术的发展对城市的交通数据进行分析（称为城市计算），为城市的交通管理提供新思路和新手段，为最大限度地发挥现有道路系统的工作效率提供了技术支持。基本方法是采集数据并通过统计物理的方法对城市数据进行分析研究，建立合理的数学或物理模型，最终能够提供完善的相关城市规划改进方案。

指导教师承担科研课题情况：

国家自然科学基金：

国家自然科学基金"面上项目"：偶氮高分子薄膜对非均匀光场响应性质的理论研究  
2018-01-01--2021-12-31，67万元（直接经费），主持

国家自然科学基金"面上项目"：蠕虫链高分子模型的高斯涨落理论及应用  
2016-01-01--2019-12-31，65万元（直接经费），主持

国家自然科学基金"青年基金"：受限条件下半刚性高分子体系的相行为的理论研究  
2014-01-01--2016-12-31，25.0万元，主持

国家自然科学基金委员会-广东省政府联合基金（第二期）超级计算科学应用研究专项资助（天河II号）  
2016-01-01，主持

国家自然科学基金"面上项目"：高效净化氯代恶臭废气微粒分子调控制备、性能及应用  
2014-01-01--2017-12-31，90.0万元，参加（第一参加人）

北京市自然科学基金：

北京市自然科学基金"面上项目"：手性对称性在高分子体系的各级尺度间传递的理论研究  
2018-01-01--2020-12-31，20万元，主持

教育部基金：

教育部留学回国人员基金：蠕虫状链体系的各向同性、各向异性转变  
2014-01-01--2015-12-31，3.0万元，主持

国家重点实验室开放基金：

超分子结构与材料国家重点实验室(吉林大学)，开放课题 ：形变对柔性高分子光电子器件光电性能的影响  
2013.1.1-2013.12.31  主持

北京交通大学基金与人才项目：

北京交通大学青年英才计划，2015-01-01--2017-12-31，主持

北京交通大学人才启动基金：受限高分子系统中的对称性破缺，2013-11-01--2014-11-01，3.0万元，主持

基本科研业务费：氯代恶臭废气治理的新型材料的研究，2013-01-01--2014-12-31，7.0万元，主持

基本科研业务费：半刚性高分子体系中纳米颗粒的扩散行为，2015-01-01--2016-12-31，7.0万元，主持

基本科研业务费：单轴对称约束对嵌段共聚物自组装结构的调控，2018-04-01--2020-03-31，6.0万元，主持

国有大型企业横向项目：

SF6/N2混合气体状态分布计算机仿真研究，2017-11-01--2017-12-31，22.0万元，主持

指导教师对本项目的支持情况

能够调配实验室的相关软硬件资源为项目提供支持，能够提供相关的指导。

立项背景：

随着我国社会经济的快速发展和城市化进程的加快，道路交通己成为制约城市发展的重要问题，解决道路交通瓶颈己成为各级政府部门需要解决的首要任务。多年来，国内外实践经验证明，在解决城市交通问题方面，除了要有相应的宏观交通政策予以支持，现代化的交通管理手段也是十分重要的一环。与此同时计算机技术、电子技术等高新技术的发展，也为交通管理提供了解决交通问题的新思路和新手段，为最大限度地发挥现有道路系统的工作效率提供了技术支持。发达国家交通发展实践表明，采用智能交通系统是解决交通拥堵减少交通事故、防止交通污染、提高交通管理水平的最有效的方法和手段。

The taxi plays a significant role in transportation systems in the urban areas across the world, providing flexible and demand-responsive transport service with widely spatiotemporal coverage. The taxi industry is often the largest employer in an urban community; therefore it will touch the lives of a significant proportion of a city population (Cooper et al., 2010). However, inefficient regulations, intense external competitions as well soaring costs have caused constant concern among taxi drivers in many cities around the world, thereby inducing strikes, protests or demonstrations over decades (BBC News, 2013; Branigan, 2008; The Irish Times, 2009; Untapped Cities, 2015). These actions can be mainly attributed to concerns about low incomes.

To tackle the tough issue, many researchers have examined the economic consequences of regulatory restraints, thereby providing information for government decision-making regarding specific regulations (Yang et al., 2002). In addition to such regulator-side studies, an increasing number of driver-side studies emerged in the wake of available taxi GPS data (Cao et al., 2005; Huang et al., 2015; Li et al., 2009; Veloso et al., 2011; Yuan et al., 2013; Zheng et al., 2011, 2009). Some of them specifically focused attentions on the low-income problem and intended to seek underlying crowd intelligence from GPS patterns, in which several factors affecting incomes were explored respectively, including duration and distance (Liu et al., 2010;

Machine Learning to Predict Taxi Fare —Part One : Exploratory Analysis

Data Cleaning, Visualisation, Feature Engineering

Machine Learning to Predict Taxi Fare — Part Two: Predictive Modelling

Model Evaluation, Model Tuning

Urban Computing with Taxicabs

Yu Zheng1, Yanchi Liu1,2, Jing Yuan1, Xing Xie1

1Microsoft Research Asia, Beijing, China

2University of Science and Technology Beijing, Beijing, China {yuzheng, v-jinyua, xingx,}

Towards mobile intelligence: Learning from GPS history data for collaborative recommendation

Vincent W. Zheng a,∗, Yu Zheng b, Xing Xie b, Qiang Yang

Mining factors affecting taxi drivers’ incomes using GPS trajectories

Guoyang Qin a, Tienan Li a, Bin Yu b, Yunpeng Wang b, Zhenhua Huang c, Jian Sun a,

研究内容：

以北京市为研究平台，运用新兴的城市计算方法，对北京城市数据进行整合分析，建立合理的模型，为城市规划提供支持。

项目方案：

本项目是基于城市产生的庞大数据上进行研究，这些数据来自于实验室合作的官方机构，真实可靠，所以数据的获取问题迎刃而解。在确定研究目标，进行任务需求分析后。首先对得到的大量数据进行除噪处理，剔除冗余的杂乱数据，为了保证数据的可靠性和准确性，这个过程将会持续一段时间。并且我们组将会针对特定的问题提出相应个性化的筛选数据标准，以更加适应日新月异和庞大的城市变化。然后我们将对数据进行可视化处理，初步建立模型并进行验证，在这个过程中，我们组的来自应用物理专业和数学专业的同学将会充分运用专业知识，从不同的角度进行分析。接下来我们组的软件学院的成员将会通过数值分析和编写机器学习算法，为我们的结果可解性提供了保障。在取得初步结果后，我们还会不断测试并完善我们的模型。最后我们将会在实际的城市规划中运用我们的模型，提出科学而有效的城市规划优化方案，证明我们成果的有效性。在完成这些后，我们将会将我们的成果总结并撰写相关论文发表。

该项目具备极高的研究价值和推广应用价值，能够在多个领域内的常见场景中得到应用。项目的其他相关补充内容已经通过国内外的其他项目得到实际验证，技术风险可控。

在整个过程中，我们来自不同学院不同专业的同学将根据实际情况，明确分工，通力合作，发挥多学科多背景的优势，共同承担项目的模型建立、代码开发、实验设计与数据分析、相关文献调研、论文撰写等工作。

通过对出租车承载行为的分析,提出了基于载客热点的出租车调度方法,来提高城市出租车调度管理系统的科学、有效、合理性,提升市民出行的便捷度及出租车公司的运营效益;降低城市出租车的空驶率及运营成本;减少城市道路交通负荷及因乘客滞留导致的公共安全危害。通过对出租车服务特性、承载行为特性和调度特性的分析,了解其承载行为在时间上的特性指标有出行次数、单次出行耗时和时间空驶率,在空间上的特性指标是出行需求分布、出行距离分布、出行路径偏好。通过统计出租车回传至信息中心的出租车位置、速度、载客状态等数据,进一步分析路网运行状况和出租车服务情况;并对数据采集误差分析,通过SQL数据库剔除冗余数据和错误数据;对电子地图进行拓扑处理,如清除微短线、弧段重叠坐标,删除悬挂弧段,坐标转换等;由于GPS数据和GIS地图数据都存在误差,釆用最短距离算法将出租车行驶轨迹数据匹配到GIS路网数据上。