



浙江工业大学

本科毕业设计论文

# 开题报告

题目：交通大数据分析网站的  
设计与开发

作者姓名 朱鑫栋

指导教师 李永强

专业班级 电气 1401 班

学 院 信息工程学院

提交日期 2018 年 3 月 22 日

# 交通大数据分析网站的设计与开发

## 一、 研究背景与意义

### 1.1 研究背景

现代大城市的交通拥堵是一个世界性难题，即使是在高度现代化的发达国家，城市交通的拥堵问题也没有得到根本的解决，各个主干道依旧拥堵不堪。交通问题是一个大城市经济发展的命脉，在现代都市，随着运输系统的快速发展，交通已经成为人类生活的重要组成部分并且对人们的生活质量产生了深远的影响，甚至成为了城市发展的瓶颈。调查显示，大约平均有 40% 的人每天至少需要花费 1 小时来通勤。为了改善交通状况，人们不断尝试着各种方法，其中，基于大数据的智能交通系统极具前景。<sup>[1]</sup>

近年来，“大数据”一词被越来越多的提及，其中叫包括交通大数据。交通大数据具有其鲜明的特点：<sup>[2][3][4]</sup>

- 1) 数据量巨大，采集渠道广泛，且部分数据需要长期储存。 不同于其它系统，交通大数据系统的数据来源非常广泛，现代城市道路系统布满了各种传感器、监控探头，它们每天都会采集海量的数据，而且数据种类多样，且需要长期储存。
- 2) 对数据实时性和处理速度要求极高 交通大数据网站的作用之一就是实时分析城市道路的交通状况，并为用户提供出行建议和指导，因此需要快速处理数据并且实时响应。
- 3) 数据模态多样化 由于交通大数据来源广泛，因此其类型也极其多样化，这对数据分析系统的兼容性提出了很高的要求
- 4) 有可视化需求 不同于传统的报表形式，交通大数据的分析结果往往需要以可视化的形式展现给用户。

因此，由于交通大数据特点鲜明，合理的运用将会带来巨大的优势。包括但不限于如下的应用：

- **交通情况的视觉监控** 通过实时监控数据（例如，隧道或交叉路口的视频监控），实时交通状况，可以观察和跟踪综合事件，以了解和分析交通堵塞的原因和机制。

- **情境感知探索和预测** 交通大数据系统都能够探索和解释交通状况，例如，实时查询城市中的出租车行程以及预测城市中大量汽车的轨迹。
- **路线规划和建议** 交通规则和路线建议是交通大数据网站的重要组成部分之一。由大数据驱动流量控制和路径规划已经取得令人满意的结果。展望未来，将人工智能纳入分析过程可以进一步提高效率。

## 1.2 研究意义

由大数据作为支撑的智慧交通系统能够有效的解决现代城市的交通问题，世界范围内，各个国家也都已经开展了相应的研究。

日本从 1990 年开始交通信息服务系统（VICS，Vehicle Information and Communication System）的研究和建设，已覆盖东京等大城市及主要高速公路。自 2002 年 VICS 中心可通过手机终端、掌上电脑、个人电脑、车载终端和电视接收器等多种途径提供交通拥堵、交通事故、道路施工、广域最优路径建议、天气状况及停车场信息等多样化的信息服务<sup>[5]</sup>。新加坡则拥有出行者信息服务系统，能够为用户提供准确实时的地铁、公交等服务信息。除了车辆的发车时间和预计通勤时间，用户还可查询到任意两地间的最少周转、最低票价或最快抵达的推荐交通路线和相应票价信息。

我国的交通信息服务系统建设以北京、上海为典型代表。

2006 年，北京市交通委组织实施交通部公众出行信息服务系统示范工程建设，开通北京公众出行网（[www.bjjt.cn](http://www.bjjt.cn)），在整合交通行业信息资源的基础上，为公交乘客和自驾车出行者提供实时、动态和综合性的交通信息服务。2007 年，在示范工程基础上开展北京市公众交通信息服务系统一期工程建设，为 2008 年奥运会提供公众交通信息服务奠定基础<sup>[6]</sup>。

2006 年，上海率先建设市级交通综合信息平台，全面、实时整合和处理全市道路交通、公共交通、对外交通领域车流、客流、交通设施等基础信息资源。在此基础上，为保障 2010 年世博交通高效有序运行，建设和完善世博交通决策支持信息服务系统、世博交通网、世博交通指南、电台电视台、世博交通信息咨询热线、路侧可变信息标志、手机与车载导航、自助查询触摸屏等多种信息服务方式。为世博交通管理者、决策者及广大公众提供全面、实时、准确的世博交通信息服务，对世博交通安全保障起到关键性作用<sup>[7]</sup>。

目前交通大数据的应用主要集中在三个方面：交通管理方面、智慧交通方面和交通事故的分析处理方面。同时，交通数据采集行业也发展迅速。

在大数据深入人心的这个时代，数据采集作为大数据的一个入口点，同时也是数据库、机器学习等方面的交汇点，已经成为满足个性化网络数据需求的最佳实践。大量的应用场景都需要数据采集工作，比如市场预测、机器语言翻译，医疗诊断领域，当然也包括交通大数据领域。显然，交通大数据网站的发展必定离不开数据采集技术的进步，未来的发展应该是二者共同繁荣的局面。<sup>[8][9][10]</sup>

## 二、 研究目标与内容

### 2.1 研究目标

中国的一些特大都市已经初步建立并使用了智慧交通系统，但是随着社会经济的快速发展，这些城市的交通管理仍然存在着许多问题，这些问题主要体现在技术和管理两大方面。

目前来看，在技术方面主要存在以下几个方面的问题：

#### 1) 设备故障问题

毫无疑问，交通大数据的发展离不开支撑它的众多数据采集设备，包括距离传感器，红外探头等。由于交通系统极其复杂和庞大，这些传感器的数量也非常巨大，随之而来的故障及检修就成为了当代交通系统的一大难题。中国的道路网往往异常的庞大和拥挤，给设备的检修和更换带来了极大的不便，与此同时，设备的检修维护费用也非常昂贵，这些都极大阻碍了现代交通系统的发展。

#### 2) 数据源的准确性问题

由于道路交通系统设备生产商的产品质量良莠不齐，数据的准确性和质量无法得到保障，这对大数据分析平台造成了极大的困扰。大量异常数据涌入系统，轻则会对分析结果造成错误影响，重则会让分析平台和系统崩溃。目前在建或者已经建成的交通大数据平台，由于系统的健壮性不足，难以对错误数据进行判断剔除或者纠正，从而使得交通诱导系统、路径规划系统不能发挥预期作用，极大的影响了交通大数据系统的使用体验。

### 3) 信息安全问题

交通大数据在产生, 传播, 存储等方面都面临着安全风险, 由于数据量庞大, 隐私问题处理起来较为棘手。目前来说存在以下问题:

用户位置信息的隐私数据暴露严重、企业和团队使用过程中的安全问题、缺乏相关的法律法规保障。<sup>[11]</sup>

而管理方面的问题主要体现在:

#### 1) 政府监管体系的不完善

就目前来说, 政府没有专门设立相应的机构来开发、管理大数据智慧交通系。行业内自由竞争, 缺乏一个领导者来指定相应的通用标准。缺乏通用标准的后果是各种传感器数据混乱, 没有统一的接口, 数据冗余问题严重, 各家的交通数据未来难以实现共享使用。

#### 2) 大部分地区交通基础设施建设落后

除了一些大都市, 国内大部分地区的交通基础设施建设较为落后, 缺乏相应的数据采集源, 不足以支撑大数据平台的分析。

#### 3) 市场问题

目前大众对于大数据交通系统的认知普遍不足, 缺乏了解, 因此目前智慧交通的市场还很小。但这是痛点同时也是一个契机, 长远来看, 无论是国内市场还是国外市场, 都非常广泛。

交通大数据具有“6V”特征, 即: “Value”特征, 蕴含了众多的信息, 因此会产生安全隐私问题; “Veracity”特征, 由此带来了去伪存真等一系列数据冗余问题; “Volume”、“Velocity”、“Visualization”特征则要求网络通信需要快速、低延迟、大带宽; “Velocity”特征要求智慧交通系统具备较高的计算性能。“Volume”特征也给数据存储带来了巨大的压力, 由于每天都会产生海量的新数据, 目前存储技术的发展远远赶不上数据的增长速度, 大量的存储服务器提高了智慧交通系统的运行维护成本。这些都是国内外大数据发展过程中遇到的问题。<sup>[13][14]</sup>

## 2.2 研究内容

### 2.2.1 单车道绿灯浪费时间估计算法

单车道绿灯浪费时间估计算法是数据驱动信号配时排查算法的基础。其基本

思想是：利用停车线过车数据点的密度来估计绿灯浪费时间。算法如下：

输入：感兴趣时间段内，车道停车线处的过车数据；

输出：车道的绿灯浪费时间；

可调参数：绿灯时间网格化的网格长度，数据点密度阈值；

步骤：

a. 将所属相位的绿灯时间以固定长度等分，得到如图 2 所示的红色矩形 (高为，宽为感兴趣的时段长度)；

b. 计算每个红色矩形中的数据点个数，其中对应绿灯结束时的第一个矩形，以此类推；

c. 若  $\rho_1 < \bar{\rho} \wedge \rho_2 < \bar{\rho} \wedge \dots \wedge \rho_i < \bar{\rho} \wedge \rho_{i+1} > \bar{\rho}$ ，则绿灯浪费时间为。

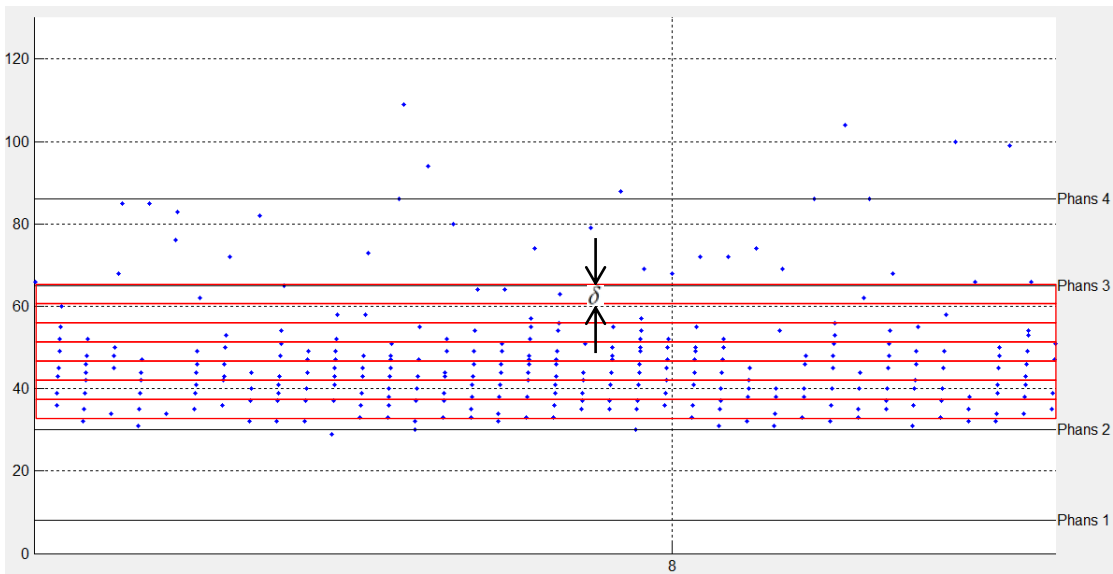


图 1: 单车道绿灯浪费时间估计算法原理

- 若绿灯浪费时间等于  $\delta$ ，则合理；
- 若绿灯浪费时间等于 0，则不合理，建议增加绿灯时间；
- 若绿灯浪费时间大于  $\delta$ ，则不合理，建议减少绿灯时间  $(i-1) \cdot \delta$ 。

### 2.2.2 单车道绿灯不足时间估计算法

function $[\hat{\rho}, \hat{\tau}] = \text{estPtTt}(\rho, \tau, T, C, g, \Delta g)$	
输入:	某车道的过车时间数据 $\rho \in \mathbb{R}^N$ , $N$ 为数据点个数 某车道的旅行时间数据 $\tau \in \mathbb{R}^N$ , $N$ 为数据点个数 获得数据时的周期 $C \in \mathbb{R}$ 获得数据时的绿灯时间 $g \in \mathbb{R}^2$ , $g(1)$ 为绿灯开始时刻, $g(2)$ 为绿灯结束时刻 增加的绿灯时间 $\Delta g \in \mathbb{R}$ 标称旅行时间 $T \in \mathbb{R}$
输出:	新配时方案下, 过车时间的预测值 $\hat{\rho} \in \mathbb{R}^N$ 新配时方案下, 旅行时间的预测值 $\hat{\tau} \in \mathbb{R}^N$
<pre> n:=0 % 过车时间提前的车辆数 f<sub>lastNew</sub>:=0 % 上一辆车的轨迹是否是新产生的 for i:=1:N     <math>\hat{\rho}(i) := \rho(i-n)</math> % 更新过车时间     <math>\hat{\tau}(i) := \tau(i) - (\rho(i) - \hat{\rho}(i))</math> % 更新旅行时间      while <math>\hat{\tau}(i) &lt; T</math> and <math>n &gt; 0</math> % 旅行时间必须大于等于标称旅行时间         n:=n-1         <math>\hat{\rho}(i) := \rho(i-n)</math> % 更新过车时间         <math>\hat{\tau}(i) := \tau(i) - (\rho(i) - \hat{\rho}(i))</math> % 更新旅行时间     end      % 判断更新后的过车时间能否在增加的绿灯时间内通过     if f<sub>lastNew</sub> = 1         <math>\hat{\rho}^{\text{temp}} := \hat{\rho}(i-1) + \rho(i) - \rho(i-1)</math>         <math>\hat{\tau}^{\text{temp}} := \tau(i) - (\rho(i) - \hat{\rho}^{\text{temp}})</math>     else         <math>\hat{\rho}^{\text{temp}} := \hat{\rho}(i) - (C - g(2) + g(1))</math>         <math>\hat{\tau}^{\text{temp}} := \hat{\tau}(i) - (C - g(2) + g(1))</math>     end      if <math>\hat{\tau}^{\text{temp}} \geq T</math> and <math>\hat{\rho}^{\text{temp}} \geq \hat{\rho}(i-1)</math> and <math>g(1) \leq \text{mod}(\hat{\rho}^{\text{temp}}, C) \leq g(2) + \Delta g</math>         <math>\hat{\rho}(i) := \hat{\rho}^{\text{temp}}</math>         <math>\hat{\tau}(i) := \hat{\tau}^{\text{temp}}</math>         n:=n+1         f<sub>lastNew</sub>:=1     else         f<sub>lastNew</sub>:=0     end end end </pre>	

表 1 过车时间/旅行时间预测算法(绿灯时间增加, 周期不变)

## 三、研究方案

### 2.2.1 研究方案简介

本方案的系统利用城市交通流大数据，实现道路交通信号配时方案智能化、精细化管理。通过对路网产生的海量交通流大数据进行自动挖掘和分析，对信号配时方案进行自动排查、优化、评价和监控，做到辖区内道路路口全覆盖与时间全覆盖，准确、持续发现信号控制管理问题并给出改进方案，以提高道路交叉口的通行能力和通行效率，减少交通延误和资源浪费，有效缓解城市交通拥堵。

### 2.2.2 系统运行环境

Windows10 操作系统

mysql5.7

acodana3

python3.6

pycharm

chrome 浏览器

python 库：

acodanna3 基本库

django1.11.4 以上

pymysql 0.7.11 以上

### 2.2.3 应用领域

用于帮助交警实现城市道路交通信号灯配时控制的自动化、持续化和智能化建设与管理，协助交警自动找到拥堵路口、拥堵路线和拥堵时间段，找到拥堵原因，依据实时、历史交通流数据分析获得路口控制的不合理问题，自动得出优化信号灯配时方案和路口管理重点，评价信号配时方案控制和管理的效果，做到从时间上（24 小时）和空间上（辖区内所有道路路口）实现信号灯配时的精细化管理与控制，做到交通大数据治堵的真正落地。

### 2.2.4 主要功能模块

该系统主要应用于挖掘海量交通流数据中的有效信息并将其作用于交通信号灯的优化，主要包括：



- 大数据驱动周期、绿灯时间、相位序列评价算法，可自动评价周期、绿灯时间、相位序列是否合理，并自动优化周期、绿灯时间、相位序列。
- 大数据驱动相位差评价算法，可自动评价相位差是否合理，并自动优化相位差。

### 三、 进度安排

- 2017.12.20 - 2018.03.04  
收集材料，阅读相关文献完成外文翻译和文献综述。
- 2018.03.05 - 2018.03.19  
完成开题报告，参加开题交流，初步确定开发方案。
- 2018.03.20 - 2018.04.30  
初步完成前端页面开发，接受中期检查。
- 2018.05.01 - 2018.05.31  
完成全部页面设计和功能开发，进行大量单元测试，撰写论文初稿。
- 2018.06.01 - 2018.06.30  
论文修改，毕业答辩，提交相关文档资料。

## 参考文献

- [1] J. Zhang et al., "Data-driven intelligent transportation systems: A survey,"IEEE Trans. Intell. Transp. Syst., vol. 12, no. 4, pp. 1624–1639,Dec. 2011.
- [2] N. Andrienko, G. Andrienko, H. Stange, T. Liebig, and D. Hecker, "Visualanalytics for understanding spatial situations from episodic movementdata," KI-Künstl. Intell., vol. 26, no. 3, pp. 241–251, Aug. 2012.
- [3] N. Ferreira, J. Poco, H.T.Vo, J.Freire, and C. T.Silva, "Visual explorationof big spatio-temporal urban data: A study of New York City taxi trips,"IEEE Trans. Vis. Comput. Graphics, vol. 19, no. 12, pp. 2149–2158,Dec. 2013.
- [4] H. Liu et al., "Visual analysis of route diversity," in Proc. IEEE Conf.Visual Anal. Sci. Technol., 2011, pp. 171–180.
- [5] 舒采焘, 张孜 . 新型城市化背景下的先进交通信息服务体系构建 [J]. 交通科技与经济, 2016, 18 (5): 21-25.
- [6] 关积珍 . 对北京奥运公共交通信息服务的探讨 [J]. 交通运输系统
- [7] 朱昊, 王磊, 张会娜 . 世博交通决策支持信息服务系统研究 [J]. 城市交通, 2010, 8 (5): 84-88
- [8] 陆化普,孙智源,屈闻聪.大数据及其在城市智能交通系统中的应用综述[J].交通运输系统工程与信息,2015,15(05):45-52.
- [9] 陈阳.大数据在智能交通系统中的应用研究[J].信息通信,2016(07):142-143.
- [10] 苏刚,王坚,凌卫青.基于大数据的智能交通分析系统的设计与实现[J].电脑知识与技术,2015,11(36):44-46.
- [11] 卢彪,李悦,张万礼.基于大数据技术的智能交通数据分析平台系统的研究与设计[J].湖北科技学院学报,2016,36(05):6-9.
- [12] 顾承华,张扬,翟希.交通大数据关键技术研究 [J]. 交通与运输(学术版),2015(02):49-53.
- [13] 于硕,李泽宇.交通大数据及应用技术研究 [J]. 中国高新技术企业,2017(04):90-91.
- [14] 王庆纲.基于大数据的智慧枢纽交通信息服务系统框架研究[J].中国市政工程,2017(06):94-97+116-117.

