

交通规划行业中的数据应用现状及思考

综合交通所

邹海翔

2019 年 3 月





大纲

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

① 大数据时代下的交通规划

② 数据“菜谱”

③ 数据分析的“武器库”

④ 数据在业务中的应用案例

⑤ 再认识与展望



目录

交通规划行业中
的数据应用现状
及思考

3

① 大数据时代下的交通规划

大数据时代的
交通规划
交通规划业务
什么是大数据
大数据对规划编制
的影响
数据“菜谱”
数据分析的“武
器库”
数据在业务中的
应用案例
再认识与展望

② 数据“菜谱”

③ 数据分析的“武器库”

④ 数据在业务中的应用案例

⑤ 再认识与展望

52



交通规划业务

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

4

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望



图：1948 年旧金山路网规划图



图：城市公交站点交通设计图



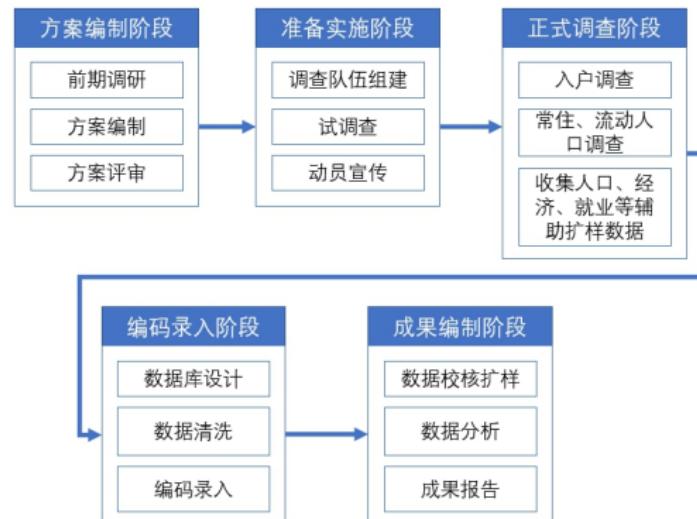
交通规划业务

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

5

- **交通调查**是交通规划业务最主要的数据来源，并以此为依据建立分析模型推断规划方案
- 国内一般**5-10年**进行一次城市居民出行调查，每次调查的时间长达数月甚至一年



图：交通调查的一般流程

52



什么是大数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

6

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

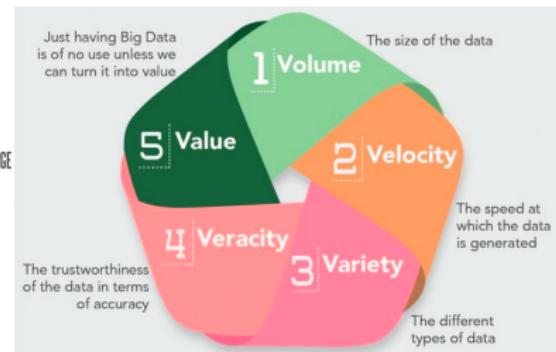
数据在业务中的
应用案例

再认识与展望



IBM 公司对大数据的定义

- ① Volume: 海量的数据规模
- ② Velocity: 快速的数据流转和动态的数据体系
- ③ Variety: 多样的数据类型
- ④ Value: 巨大的数据价值
- ⑤ Veracity: 数据的准确性和可信赖度



图：大数据的 5V 定义



什么是大数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

7

- 与传统数据相比，不仅体现在数据量巨大，更重要的是可以覆盖业务的近乎全部数据
- 数据爆炸时代的必然产物
- 商业公司进行的一场成功的营销

52



什么是大数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 与传统数据相比，不仅体现在数据量巨大，更重要的是可以覆盖业务的近乎全部数据
- 数据爆炸时代的必然产物
- 商业公司进行的一场成功的营销

7



什么是大数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 与传统数据相比，不仅体现在数据量巨大，更重要的是可以覆盖业务的近乎全部数据
- 数据爆炸时代的必然产物
- 商业公司进行的一场成功的营销

7



什么是大数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

7



- 与传统数据相比，不仅体现在数据量巨大，更重要的是可以覆盖业务的近乎全部数据
- 数据爆炸时代的必然产物
- 商业公司进行的一场成功的营销

大数据是一场思维的颠覆：放弃对因
果关系（为什么）的渴求，而取而代之关注
相关关系（是什么）。

—维克托·迈尔·舍恩伯格



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

8

● 外部条件与时代背景

技术进步

随着计算机技术的发展，尤其是云计算和人工智能技术的进步，使得数据获取变得容易了许多，而且处理和分析的能力越来越强

经济转型

当前中国的经济正处于重要的转型时期，经济的发展模式、发展要素、发展路径等等都亟需转变，以适应现在的大环境

社会转型

虽然我们的经济实现了快速发展，但社会矛盾出现越来越尖锐化的趋势

52



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

8

● 外部条件与时代背景

技术进步

随着计算机技术的发展，尤其是云计算和人工智能技术的进步，使得数据获取变得容易了许多，而且处理和分析的能力越来越强

经济转型

当前中国的经济正处于重要的转型时期，经济的发展模式、发展要素、发展路径等等都亟需转变，以适应现在的大环境

社会转型

虽然我们的经济实现了快速发展，但社会矛盾出现越来越尖锐化的趋势

52



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

8

● 外部条件与时代背景

技术进步

随着计算机技术的发展，尤其是云计算和人工智能技术的进步，使得数据获取变得容易了许多，而且处理和分析的能力越来越强

经济转型

当前中国的经济正处于重要的转型时期，经济的发展模式、发展要素、发展路径等等都亟需转变，以适应现在的大环境

社会转型

虽然我们的经济实现了快速发展，但社会矛盾出现越来越尖锐化的趋势

52



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

8

● 外部条件与时代背景

技术进步

随着计算机技术的发展，尤其是云计算和人工智能技术的进步，使得数据获取变得容易了许多，而且处理和分析的能力越来越强

经济转型

当前中国的经济正处于重要的转型时期，经济的发展模式、发展要素、发展路径等等都亟需转变，以适应现在的大环境

社会转型

虽然我们的经济实现了快速发展，但社会矛盾出现越来越尖锐化的趋势

52



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

9

● 传统城市空间规划面临方法转型

规划转型

- 传统时空间概念被重新定义，以空间研究和布局为主要内容的城市空间规划面临着研究范式的转型和规划编制方法上的革新
- 随着国内城市化进程的发展，规划面临的更多是城市的存量式发展，由粗放向集约进行转型，对规划的精细化、定量化管理提出了要求

城市规划体系转型

- “自上而下”与“自下而上”结合的规划理念
- 基于大数据应用的城乡规划方法体系创新
- 基于GIS的城乡规划系统整合
- ...

新的规划方法和技术

- 数据获取与处理技术
- 现场调研手段
- 方案编制新方法
- 规划公众参与方式

图：南京大学规划专业试行的学科教学改革方案

52



大数据对规划编制的影响

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

交通规划业务

什么是大数据

大数据对规划编制
的影响

数据“菜谱”

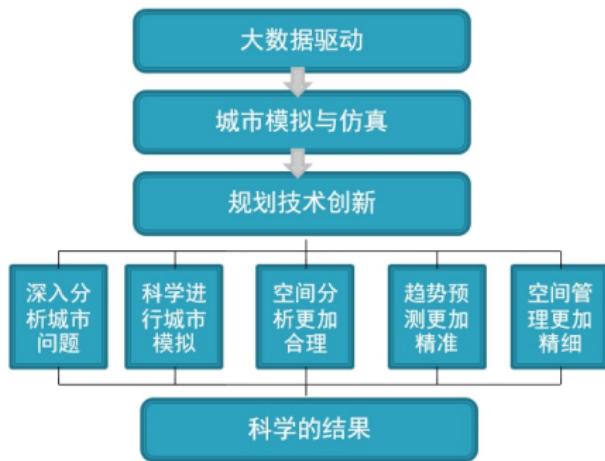
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

10

- 大数据提供了认识和分析城市问题新的思维和技术方法
- 大数据时代到来，可以让我们更清楚地了解和观察城市的发展、变化过程，同时也使得规划过程变得透明可控；
- 大数据技术强化了对规划过程的重视和科学化，尤其是对规划调研、空间分析、公共参与与空间协调规划、空间预测和可视化等过程的科学把握，有助于推动规划过程的科学化



图：大数据驱动的规划决策评估架构

52



目录

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

11

1 大数据时代下的交通规划

2

数据“菜谱”

3

数据分析的“武器库”

4

数据在业务中的应用案例

5

再认识与展望



数据分类

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

12

● 地理信息数据

- ① 交通网络数据
- ② 土地利用和建筑物数据
- ③ 地形图和影像数据

● 静态调查数据

- ① 居民出行调查数据
- ② 跨界调查数据

● 动态大数据

- ① 车辆 GPS 数据
- ② 车牌识别数据
- ③ 公交刷卡数据
- ④ 手机定位数据
- ⑤ 互联网定位数据

● 互联网开放数据

- ① 互联网地图
- ② 交通出行



数据分类

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

12

● 地理信息数据

- ① 交通网络数据
- ② 土地利用和建筑物数据
- ③ 地形图和影像数据

● 静态调查数据

- ① 居民出行调查数据
- ② 跨界调查数据

● 动态大数据

- ① 车辆 GPS 数据
- ② 车牌识别数据
- ③ 公交刷卡数据
- ④ 手机定位数据
- ⑤ 互联网定位数据

● 互联网开放数据

- ① 互联网地图
- ② 交通出行

52



数据分类

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

12

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

● 地理信息数据

- ① 交通网络数据
- ② 土地利用和建筑物数据
- ③ 地形图和影像数据

● 静态调查数据

- ① 居民出行调查数据
- ② 跨界调查数据

● 动态大数据

- ① 车辆 GPS 数据
- ② 车牌识别数据
- ③ 公交刷卡数据
- ④ 手机定位数据
- ⑤ 互联网定位数据

● 互联网开放数据

- ① 互联网地图
- ② 交通出行



数据分类

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

12

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

● 地理信息数据

- ① 交通网络数据
- ② 土地利用和建筑物数据
- ③ 地形图和影像数据

● 静态调查数据

- ① 居民出行调查数据
- ② 跨界调查数据

● 动态大数据

- ① 车辆 GPS 数据
- ② 车牌识别数据
- ③ 公交刷卡数据
- ④ 手机定位数据
- ⑤ 互联网定位数据

● 互联网开放数据

- ① 互联网地图
- ② 交通出行

52

交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

13

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

- 非结构化网络，用于规划编制成果效果图

- 基于节点-弧段模型的结构化网络，用于定量分析和自动化制图



图：规划项目中常用的非结构化网络，无法直接用于定量分析



交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

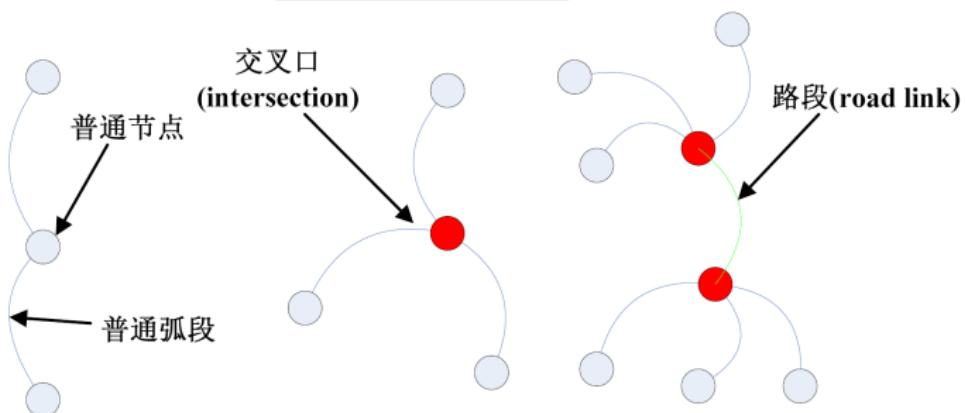
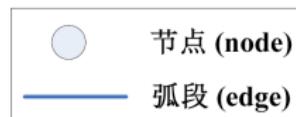
数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

13

- 非结构化网络，用于规划编制成果效果图
- 基于节点-弧段模型的结构化网络，用于定量分析和自动化制图



图：交通网络在计算机中的结构化存储形式

52

交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

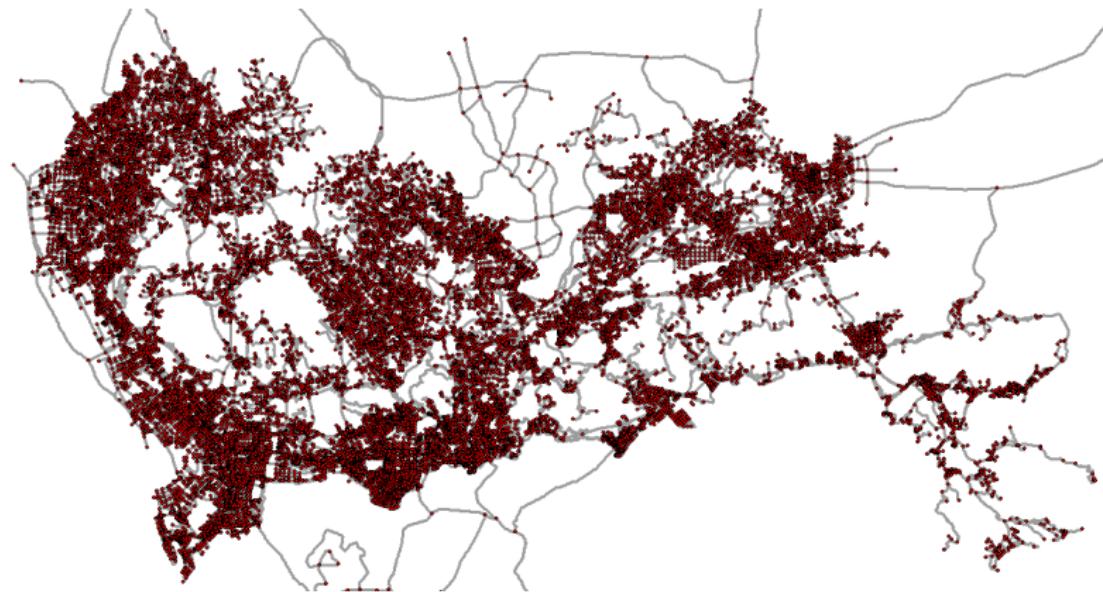
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

14

- 道路网络
- 地铁轨道网络
- 常规公交网络



图：道路网络



交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

14

- 道路网络
- 地铁轨道网络
- 常规公交网络

属性名称	含义	类型
NAME	道路名称	字符型
CDS	车道数	整数型
LEN	道路长度	浮点型
LDKD	道路宽度	浮点型
FJDCKD	非机动车道宽度	浮点型
JDCDKD	机动车道宽度	浮点型
RXDKD	人行道宽度	浮点型
HXKD	红线宽度	浮点型
DLDJ	道路等级	整数型
JXTXJ	可通行交通方式	字符型
GJZYD	是否具备公交专用道	布尔型

表：道路网络包含的主要属性信息



交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

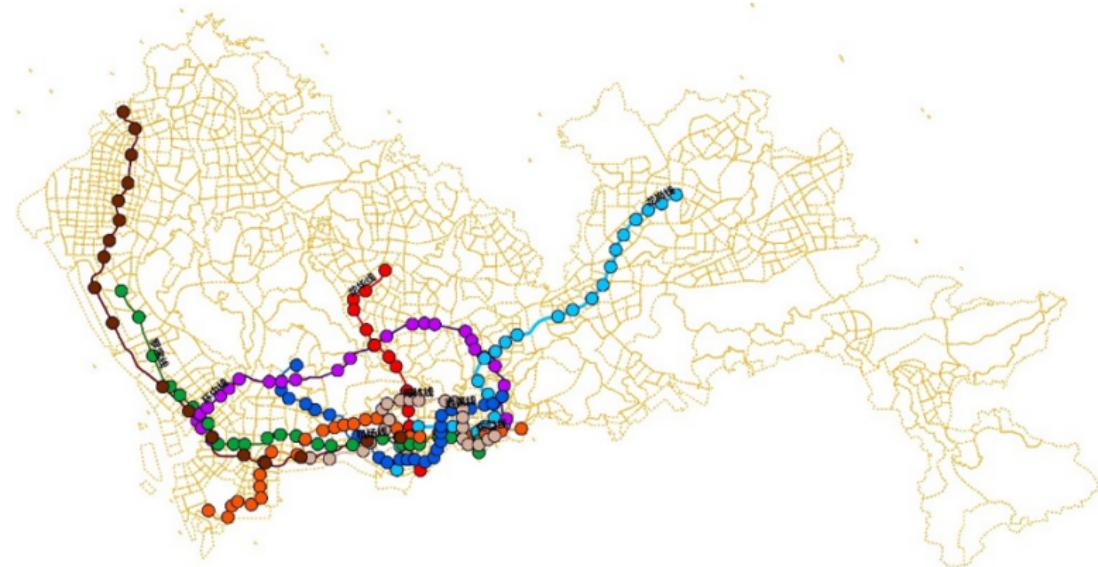
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

14



图：地铁轨道网络

52



交通网络数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

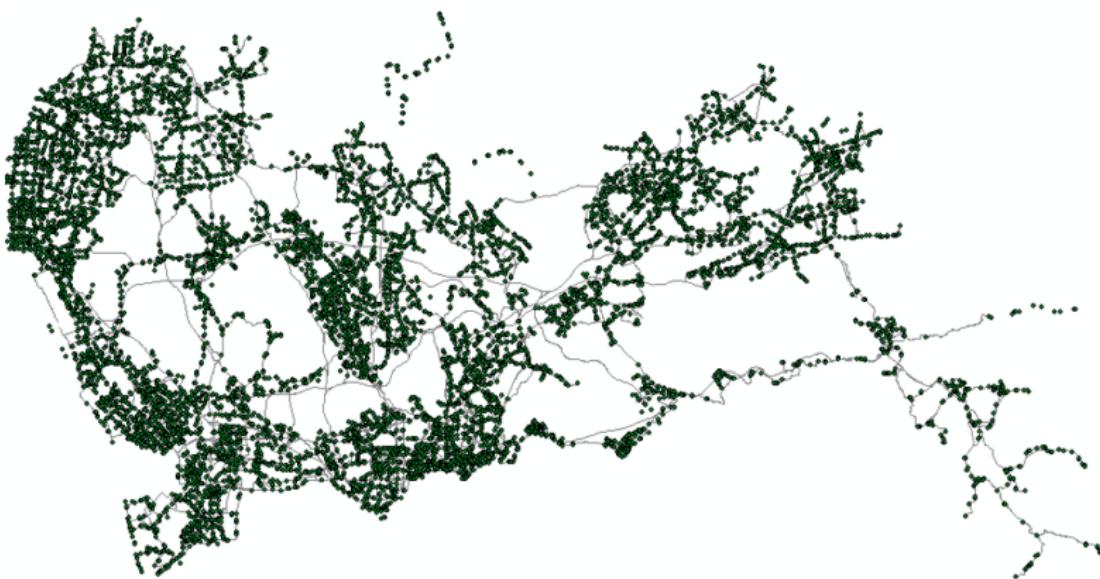
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

14



图：常规公交网络

土地利用和建筑物数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

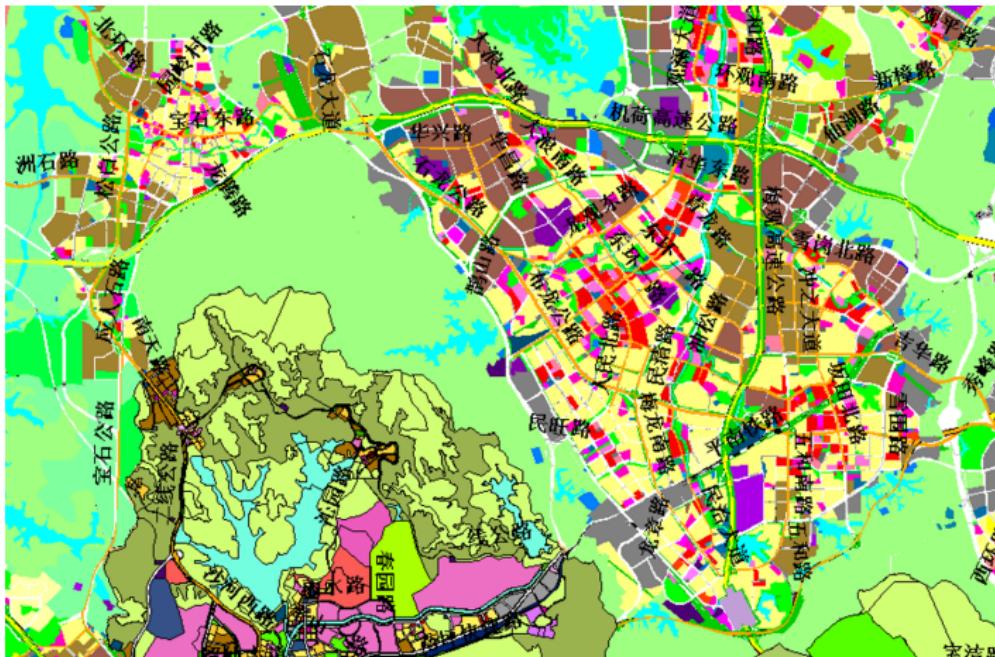
数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

15

- 规土委核心数据，一张图系统提供



图：土地利用数据

52

土地利用和建筑物数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

15

- 规土委核心数据，一张图系统提供



图：建筑物数据

地形图和影像数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

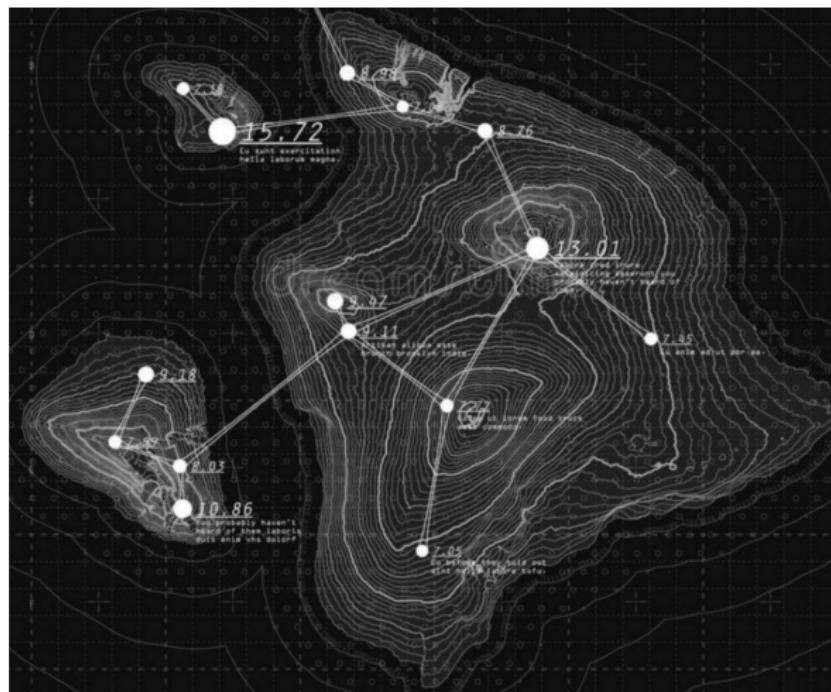
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

16



图：地形图数据



地形图和影像数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

16

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

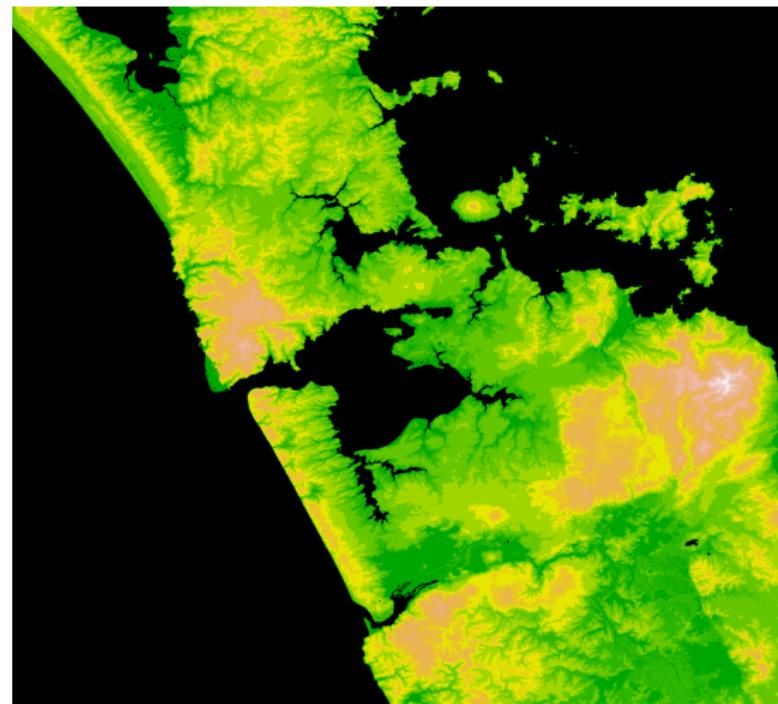
互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望



图：遥感影像数据

52



居民出行调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

17

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 2005、2010、2016 三次居民出行调查数据
- 最终数据成果是户表、人表和出行表共三张表

52

居民出行调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

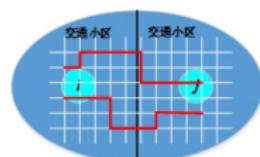
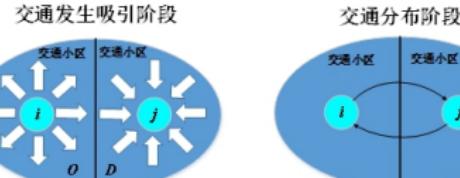
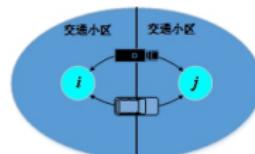
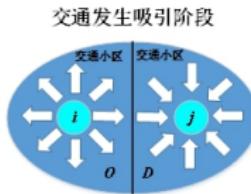
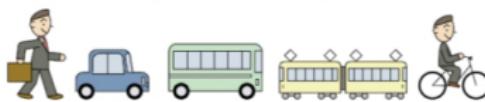
数据在业务中的
应用案例

再认识与展望

17



What kind of transportation is used?



交通方式划分阶段

交通分配阶段

图：以家庭为单位，调查家庭成员出行次数、出行目的、交通方式和目的地



居民出行调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

17

- 2005、2010、2016 三次居民出行调查数据
- 最终数据成果是户表、人表和出行表共三张表

主要字段	说明
户 ID	唯一值
回答时间	填写问卷的时间
建筑物位置	被访问者居住地， 经纬度坐标
户类型	家庭户：1；集体户：2
居住人数	分为 ≥ 4 岁人数和 < 4 岁人数
家庭年收入	≤ 4 万：1；10-20 万：2；20-30 万：3；30-50 万：4； ≥ 50 万：5
住房来源	租赁廉租房、租赁城中村、租赁其他住房、自建房、购买商品房、购买福利房或保障房、集体宿舍
拥车情况	是否拥有小汽车？家庭拥有几辆小汽车？

表：户表



居民出行调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

17

- 2005、2010、2016 三次居民出行调查数据
- 最终数据成果是户表、人表和出行表共三张表

主要字段	说明
人 ID	与户 ID 对应
年龄	
性别	
户口登记情况	本市户籍：1；非本市户籍：2。其中，非本市户籍中是否居住 6 个月以上
文化程度	分为 9 个选项
职业	分为 9 个选项
所属行业	参考经济普查问卷，分为 18 个选项
工作地或学校地址	

表：人表



居民出行调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

17

- 2005、2010、2016 三次居民出行调查数据
- 最终数据成果是户表、人表和出行表共三张表

主要字段	说明
出行 ID	与人 ID 对应
出行和换乘方式	公交、小汽车、地铁等共 12 类
出行目的	上班、上学、公务等 10 类
出发时间	
出发地点	详细地址及经纬度
到达时间	
到达地点	详细地址及经纬度
换乘站点	
步行时间、候车时间、车内时间	公共交通出行

表：出行表



跨界调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

18

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 从 2013 年开始，每两年开展一次跨界客流调查
- 调查范围包括深圳和东莞、惠州的边界、深圳出境口岸和重要对外交通枢纽

52

跨界调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

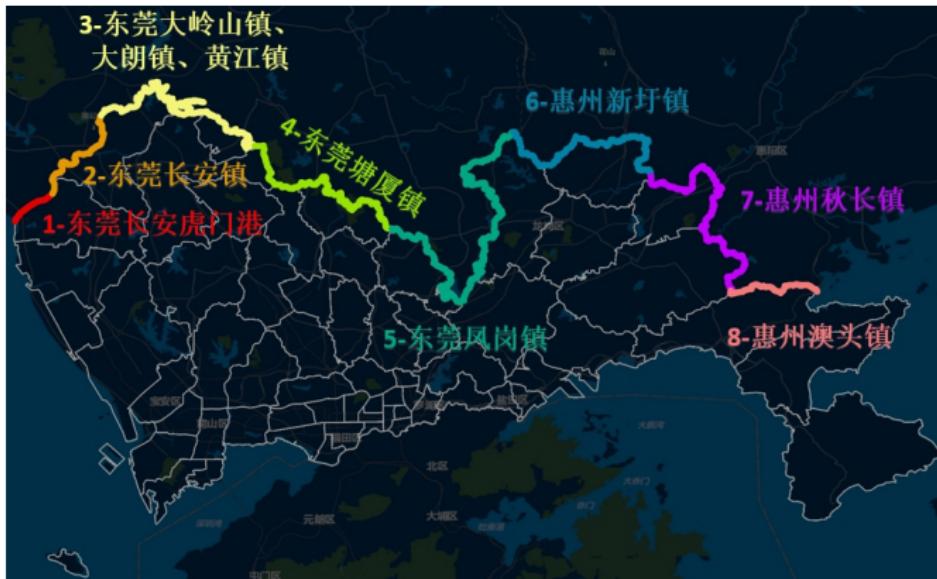
数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

18

- 从 2013 年开始，每两年开展一次跨界客流调查
- 调查范围包括深圳和东莞、惠州的边界、深圳出境口岸和重要对外交通枢纽



图：深莞惠境界线

跨界调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

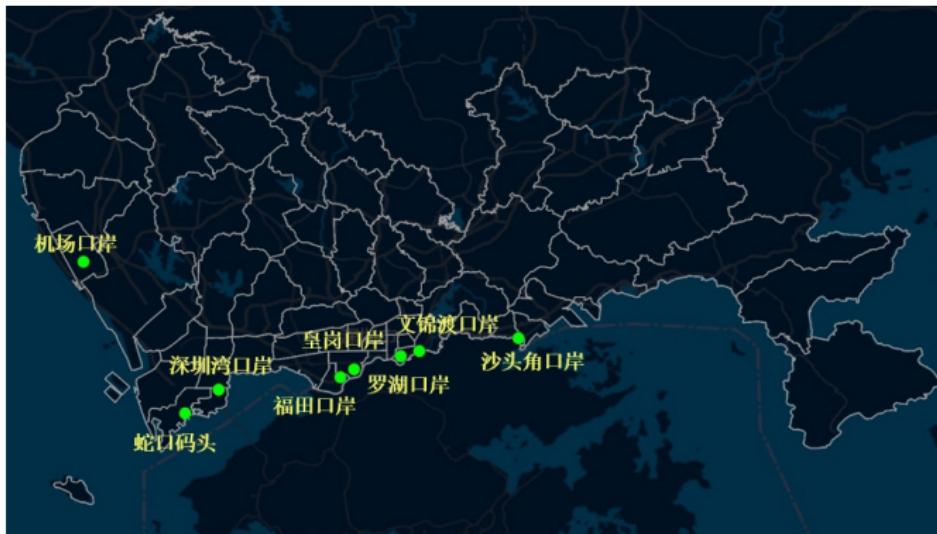
数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

18

- 从 2013 年开始，每两年开展一次跨界客流调查
- 调查范围包括深圳和东莞、惠州的边界、深圳出境口岸和重要对外交通枢纽



图：深港口岸

跨界调查数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

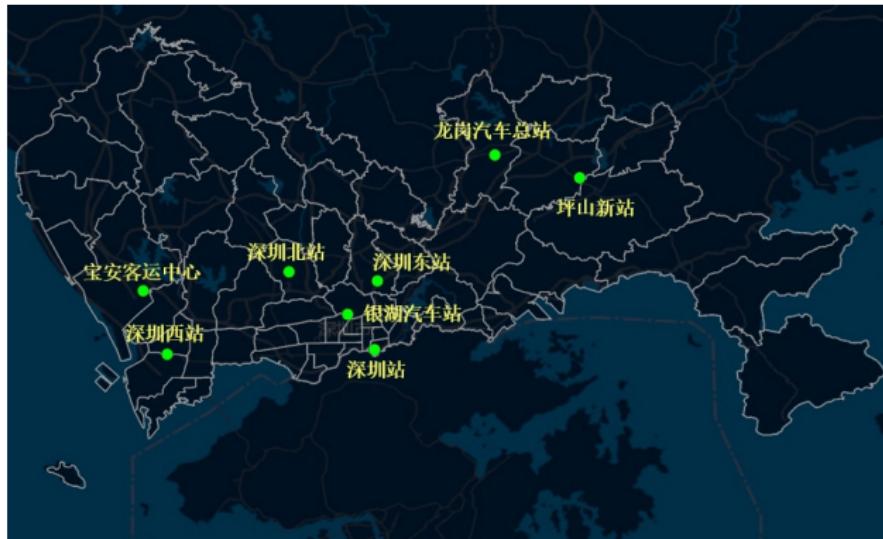
交通出行

数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

18



图：重要对外交通枢纽

52



车辆 GPS 数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

19

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

- 利用 GPS 卫星定位车辆，记录车辆位置、时间、方向、速度和状态等信息
- 覆盖全部出租车、公交车、特种车以及部分货车，约 10 万辆
- 从 2013 年开始收集，10-40 秒回传一次数据，日均数据量约 10GB 左右，超过 1 亿条

```
,117.3711166381836,38.92399978637695,2016-03-07 00:00:23,96,78,225,0
,117.36481475830078,38.92188262939453,2016-03-07 00:00:50,96,81,225,0
,117.35831451416016,38.920166015625,2016-03-07 00:01:17,96,78,225,0
,117.3514633178711,38.91878128051758,2016-03-07 00:01:45,96,79,225,0
,117.34476470947266,38.917449951171875,2016-03-07 00:02:11,96,82,225,0
,117.33816528320312,38.916133880615234,2016-03-07 00:02:37,96,81,225,0
,117.33136749267578,38.9147834777832,2016-03-07 00:03:04,96,81,225,0
,117.32465362548828,38.913448333740234,2016-03-07 00:03:30,96,82,225,0
,117.31790161132812,38.91193389892578,2016-03-07 00:03:58,96,77,225,0
,117.31183624267578,38.909366607666016,2016-03-07 00:04:26,96,73,225,0
,117.30690002441406,38.90566635131836,2016-03-07 00:04:53,96,82,180,0
,117.30332946777344,38.9011344909668,2016-03-07 00:05:19,96,82,180,0
,117.30133056640625,38.895931243896484,2016-03-07 00:05:46,96,77,180,0
```

图：GPS 原始数据文件，每辆车存储一个文件



车辆 GPS 数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

19

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 利用 GPS 卫星定位车辆，记录车辆位置、时间、方向、速度和状态等信息
- 覆盖全部出租车、公交车、特种车以及部分货车，约 10 万辆
- 从 2013 年开始收集，10-40 秒回传一次数据，日均数据量约 10GB 左右，超过 1 亿条



图：空间中的 GPS 数据

车辆 GPS 数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

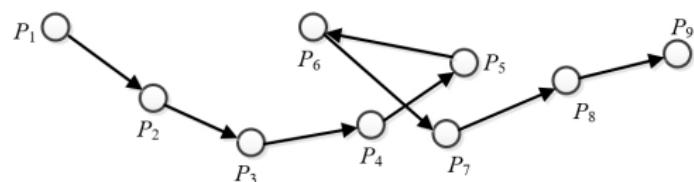
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 利用 GPS 卫星定位车辆，记录车辆位置、时间、方向、速度和状态等信息
- 覆盖全部出租车、公交车、特种车以及部分货车，约 10 万辆
- 从 2013 年开始收集，10-40 秒回传一次数据，日均数据量约 10GB 左右，超过 1 亿条

	经度	纬度	时间
$P_1:$	Lat_1	Lon_1	T_1
$P_2:$	Lat_2	Lon_2	T_2
.....			
$P_n:$	Lat_n	Lon_n	T_n



图：GPS 轨迹数据示意图



车牌识别数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

20



图：道路上的车辆拍摄设备

52

车牌识别数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

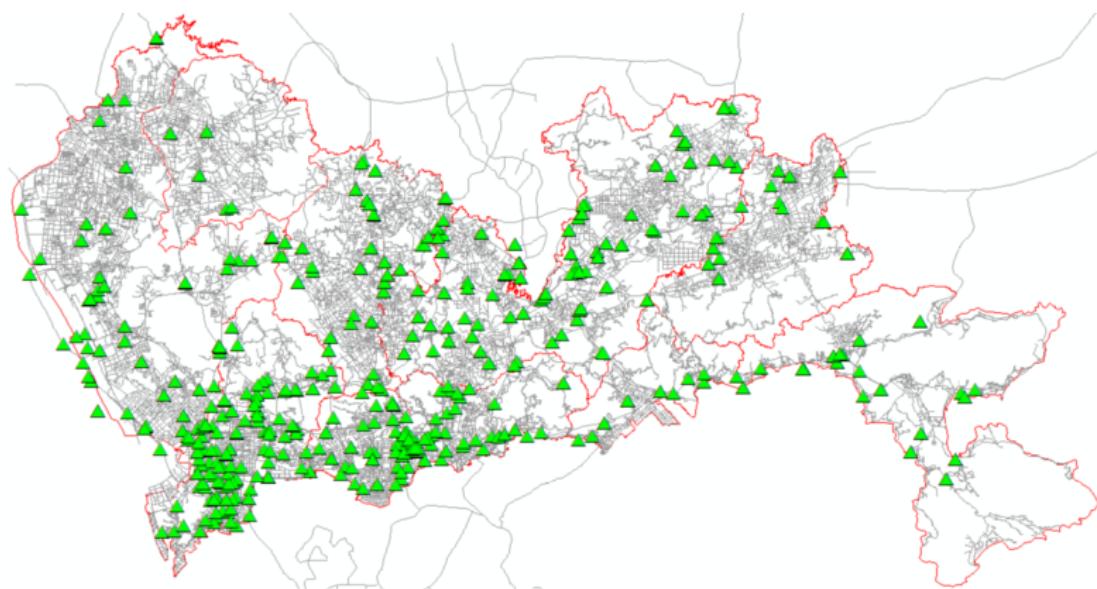
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

20



图：车牌识别点位在全市的分布



车牌识别数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

20

- 通过车牌识别算法，从布设在道路上的拍摄视频中提取车牌
- 全市目前有 300 多个检测点位
- 从 2014 年开始收集，日均数据量是约 2.5GB，超过 1200 万条

#D e08b15d59e4f932b0d4cfb2ae5ac5b54_2.0_2017-02-06 00:00:09 2017-02-06 00:00:19.021102210.3....-1.00_1.2017020600009360211022131_0_0_0.10....1.0.0.0.1.0.6....
#B 19a5d2cae2160da9a0e67422a5ea00_0.0_2017-02-06 00:00:18 2017-02-06 00:00:20.2017020600009360211022131_0_0_0.10....1.0.1.95_0.0.6.Uf....
#B 8a5f07a6b3b6e0ec6e6712f788001d_0.0_2017-02-06 00:00:19 2017-02-06 00:00:21.10300207_0.2....74.00_2.2017020600009360211022131_0_0_0.10....1.0.1.97_0.0.6.Uf....
#B 904dcf372754292aa043a81bd1bf3_0.0_2017-02-05 23:59:47 2017-02-06 00:00:21.20204001_0.2....其它_105.00_2.2017020523594760200400121_0_0_0.10....1.0.17020523594760200400122_0_0_0.10....
#B 237a93d34bf6f6605932238f7312_2.0_2017-02-06 00:00:22.21102121_0.1....1.00_1.2017020600001910211021211_0_0_0_10....1.0.0.0.1.0.6....
#B fb8b72c2a59e6ea273d07e56df4d1c_0.0_2017-02-05 23:59:19 2017-02-06 00:00:22.20507301_0.3....1.00_2.20170205235919042050730132_10_0.1.0.63_0.0.5.Uf....
#B 5bdcab1f6320836d53b7952f1622f_0.0_2017-02-05 23:59:43 2017-02-06 00:00:23.20507301_0.3....1.00_2.20170205235943922050730131_10_0.1.0.60_0.0.5.Uf....
#S 42bdff628c5787213262a6665355ce8_2.0_2017-02-06 00:00:03 2017-02-06 00:00:23.20507301_0.3....1.00_2.20170206000003842050730131_10_0.1.0.81_0.0.6.Uf....
#B 789c0b54bf624b1f6396ea6e5fb00954_0.0_2017-02-06 00:00:02.2017-02-06 00:00:23.20501803_0.1....0.00_2.2017020600001002050180311_10_0.1.0.80_0.0.6.Uf....
#S 7fa2377ae7c7c0_19245677429bc32b_0.0_2017-02-06 00:00:22.2017-02-06 00:00:25.10102040_0.5....49.00_2.20170206000022621010020451_10_0.170206000023901010020531_10_0.1.0.83_0.0.6.Uf....
#B dd0205284b159807cd396a63389271_0.0_2017-02-06 00:00:23.2017-02-06 00:00:26.10100205_0.3....71.00_2.20170206000023901010020531_10_0.170206000023981010020532_10_0.1.0.83_0.0.6.Uf....
#B 92c04c6c7b701fba6e262b739457b6c_0.0_2017-02-06 00:00:24.2017-02-06 00:00:27.20104404_0.4....其它_82.00_2.20170206000024452010440441_0_0_0_10....1.0.170206000024582010440442_0_0_0_10....
#B 96dcb8ef32b4befad3baef6f2560bf76_0.0_2017-02-06 00:00:25.2017-02-06 00:00:27.10100204_0.2....70.00_2.20170206000025701010020421_10_0.1.94_0.0.6.Uf....
#B ea394c0a3532ce3a5a45bd6ed900_0.0_2017-02-06 00:00:25.2017-02-06 00:00:28.10100406_0.1....163.00_2.20170206000025611010040611_10_0.170206000025681010040612_10_0.1.95_0.0.6.Uf....
#B ea81e299d3fcac90e8b8c1a5ae4f73_0.0_2017-02-06 00:00:28.2017-02-06 00:00:29.10100211_0.1....52.00_2.2017020600002801010021111_10_0.2017020600002801010021112_10_0.1.0.97_0.0.6.Uf....
#B 8b00e92dc0ca97e269d11f56fc2b27_0.0_2017-02-06 00:00:27.2017-02-06 00:00:29.10300201_0.3....42.00_2.2017020600002771030020131_10_0.170206000027851030020132_10_0.1.0.63_0.0.6.Uf....
#B fe9068a4bcfd803dd370bd1d6e1b_0.0_2017-02-05 23:55:05.2017-02-06 00:00:31.20110038_0.1....其它_0.00_1.20170205235505962011003811_0_0_0_10....1.0.0.0.1.0....
#B 30f2683674a8584bc3fe5402cea5cba_0.0_2017-02-06 00:00:28.2017-02-06 00:00:31.10100406_0.8....52.00_2.20170206000028601010040681_10_0.170206000029001010040682_10_0.1.96_0.0.6.Uf....
#B 39b2db62b0bd7a/dc66036350a6bad2_0.0_2017-02-06 00:00:30.2017-02-06 00:00:32.10100206_0.4....105.00_2.20170206000030121010020641_10_0.170206000030201010020642_10_0.1.72_0.0.6.Uf....
#B d5d242e19f72410eb72b038c20e1105_0.0_2017-02-06 00:00:24.2017-02-06 00:00:33.02110189_0.6....-1.00_2.01702060000243402110189_1_0_0_0_10....1.0.0.0.1.0....
#B e617e5c35748b5708d3f8819e00b4_0.0_2017-02-06 00:02:06.2017-02-06 00:00:34.02110256_0.2....-1.00_1.20170206000020647021102562_1_0_0_0_10....1.0.0.0.1.0....
#B 6bb6c6594f5cb1a570698e9a7d72619_13.0_2017-02-06 00:15.25.2017-02-06 00:00:35.02110027_0.1....其它_0.00_1.201702060001525850211002711_0_0_0_10....1.0.0.0.1.0....

图：车牌识别原始数据

公交刷卡数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

21



图：深圳通刷卡设备



公交刷卡数据

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

21

- 全部深圳通刷卡数据，包括地铁和常规公交
- 包含刷卡时间、终端编号、卡号、进出标识等信息
- 从 2013 年开始收集，日均数据量 3.5GB，超过 1500 万条

CARDID	TRADETYPE	TRADEDATE	TERMINALID
329740393	22	2018/7/2	261032119
360639696	22	2018/7/2	261032104
684992446	22	2018/7/2	261032103
843962946	22	2018/7/2	261012143
690331483	22	2018/7/2 0:00:01	261009141
281229535	22	2018/7/2 0:00:02	263035108
293554817	22	2018/7/2 0:00:02	261030107
685548241	22	2018/7/2 0:00:02	268005120
880025720	22	2018/7/2 0:00:02	263035104
361826874	22	2018/7/2 0:00:03	261012143
841593855	22	2018/7/2 0:00:03	261027107
665631757	22	2018/7/2 0:00:04	261032122
852840228	22	2018/7/2 0:00:04	261031107
844677279	22	2018/7/2 0:00:05	261032102
860132723	22	2018/7/2 0:00:05	260012113

图：深圳通刷卡原始数据

52

手机定位数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

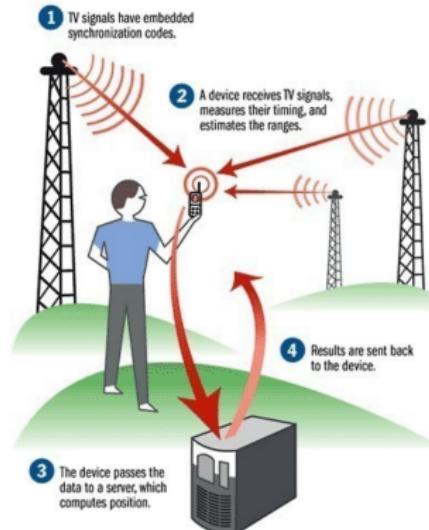
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 利用手机与基站的通信定位手机位置、时间信息；另外，运营商还掌握机主实名信息
- 数据量受采样频率影响，深圳市日均通常可以达到 TB 级别
- 目前只有电信和联通少量处理后的数据



图：手机定位原理



手机定位数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 利用手机与基站的通信定位手机位置、时间信息；另外，运营商还掌握机主实名信息
- 数据量受采样频率影响，深圳市日均通常可以达到 TB 级别
- 目前只有电信和联通少量处理后的数据

互联网定位数据

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

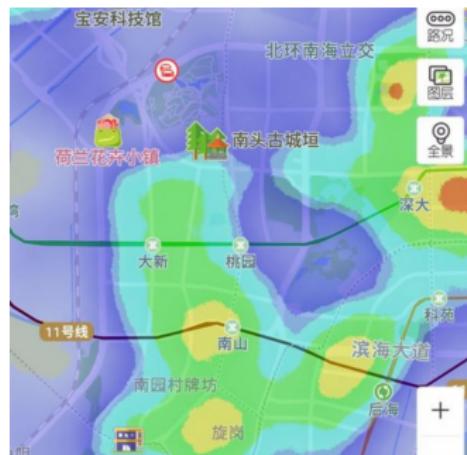
交通出行

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

23



百度利用定位数据制作热力图



腾讯实时定位数据分布

52



互联网地图

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

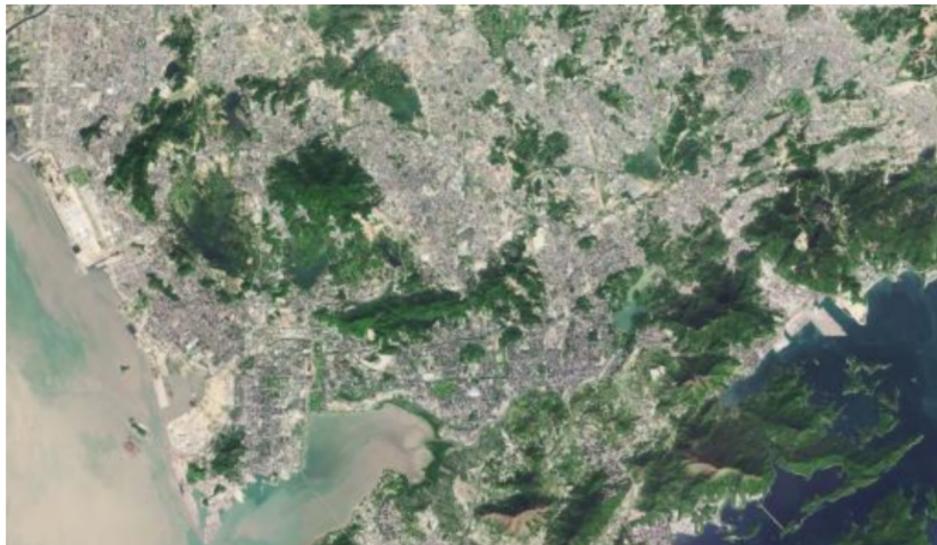
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

24

- 卫星遥感影像地图
- 制图综合后的瓦片地图
- 兴趣点 (POI) 数据
- 矢量 GIS 数据



图：经过处理后的卫星遥感影像地图



互联网地图

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

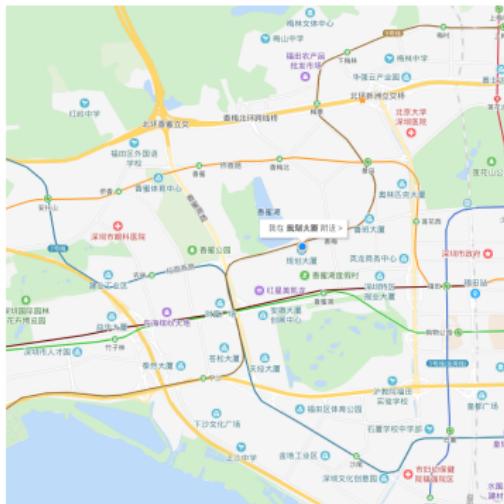
数据分析的“武
器库”

数据在业务中的
应用案例

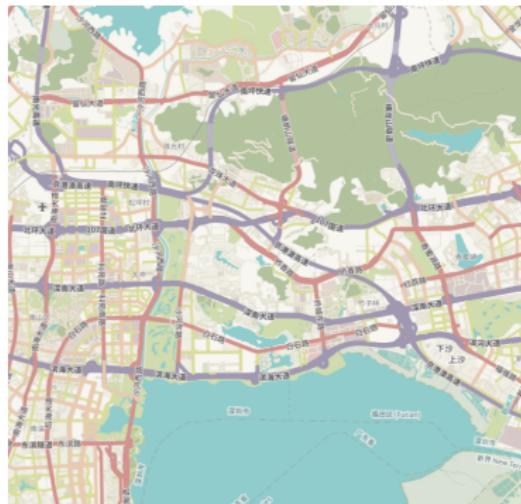
再认识与展望

- 卫星遥感影像地图
- 制图综合后的瓦片地图

- 兴趣点 (POI) 数据
- 矢量 GIS 数据



(a) 百度地图



(b) openstreetmap 地图

图：通过制图综合技术对矢量数据进行渲染优化，然后制作而成的瓦片地图



互联网地图

交通规划行业中
的数据应用现状
及思考

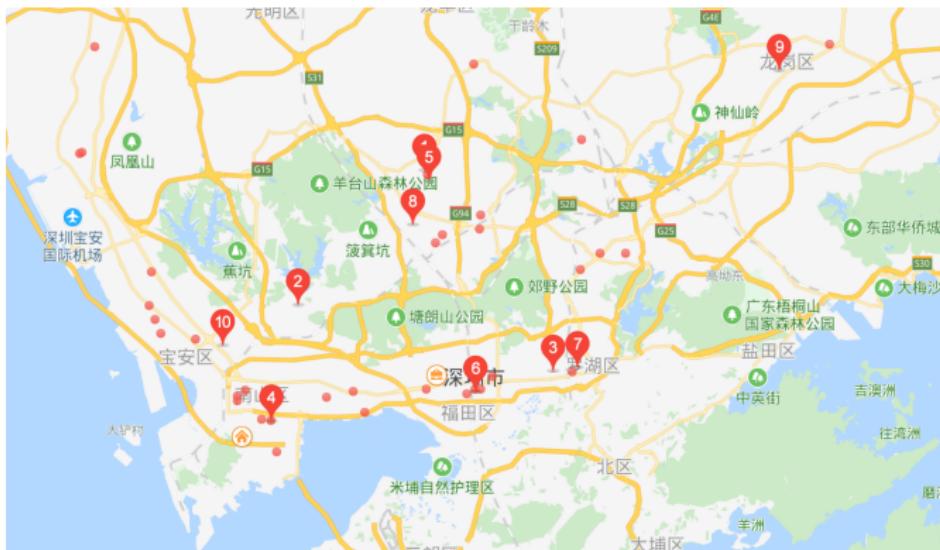
大数据时代下的
交通规划

数据“菜谱”
交通网络数据
土地利用和建筑物
数据
地形图和影像数据
居民出行调查数据
跨界调查数据
车辆 GPS 数据
车牌识别数据
公交刷卡数据
手机定位数据
互联网定位数据
互联网地图
交通出行

数据分析的“武器库”
数据在业务中的
应用案例

再认识与展望

- 卫星遥感影像地图
- 制图综合后的瓦片地图
- 兴趣点 (POI) 数据
- 矢量 GIS 数据



图：兴趣点数据



互联网地图

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

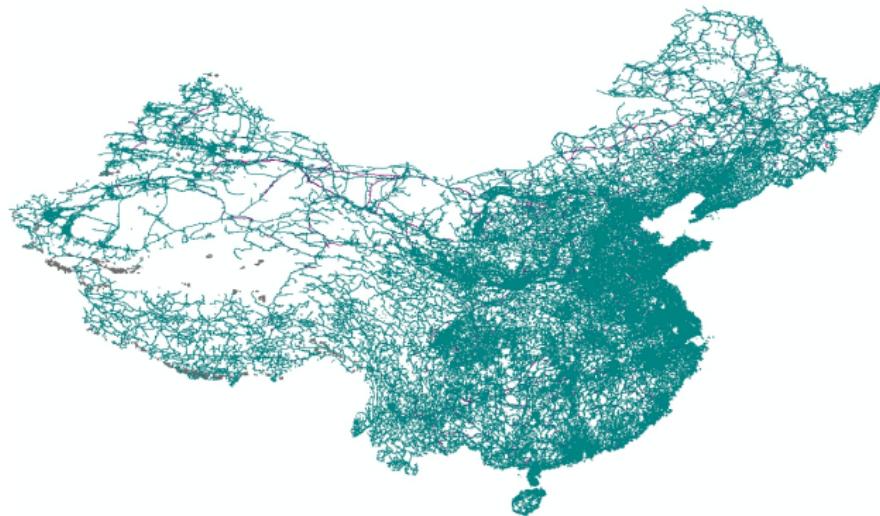
数据“菜谱”
交通网络数据
土地利用和建筑物
数据
地形图和影像数据
居民出行调查数据
跨界调查数据
车辆 GPS 数据
车牌识别数据
公交刷卡数据
手机定位数据
互联网定位数据
互联网地图
交通出行

数据分析的“武
器库”
数据在业务中的
应用案例

再认识与展望

24

- 卫星遥感影像地图
- 制图综合后的瓦片地图
- 兴趣点 (POI) 数据
- 矢量 GIS 数据



图：OSM 网站获取的矢量 GIS 数据，由来自世界各地的自愿者生产

52

交通出行

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

25

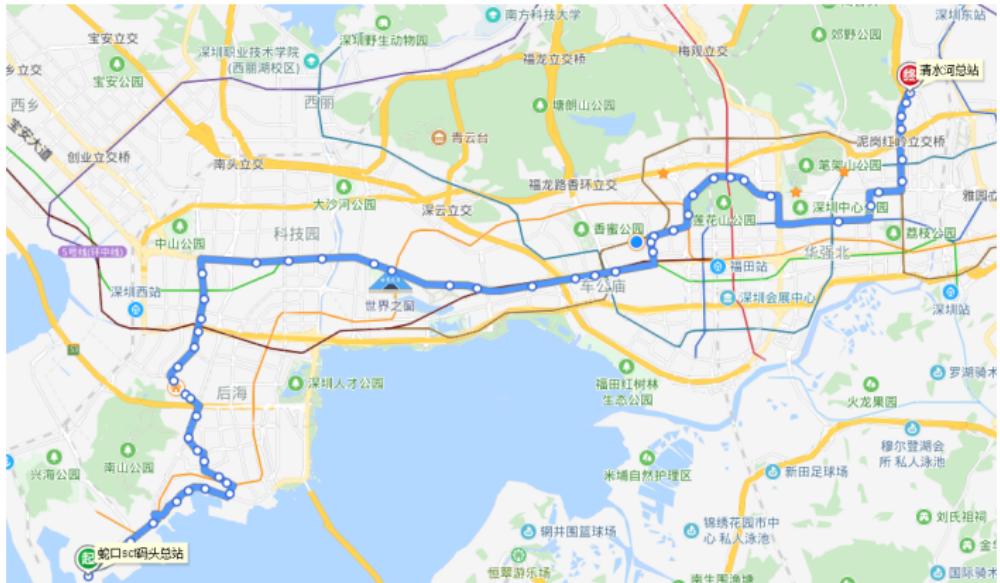
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

● 公共交通线路和站点

● 交通出行 API



图：互联网上同步更新的公共交通线路和站点

52



交通出行

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

交通网络数据

土地利用和建筑物
数据

地形图和影像数据

居民出行调查数据

跨界调查数据

车辆 GPS 数据

车牌识别数据

公交刷卡数据

手机定位数据

互联网定位数据

互联网地图

交通出行

25

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 公共交通线路和站点
- 交通出行 API

The screenshot shows the Baidu Map API documentation page. At the top, there's a navigation bar with links for Home, Functions & Services (NEW!), Solutions, Developer Documentation, Feedback & Help, Service Upgrade (NEW!), and Control Panel. Below the navigation is a banner for the "WEB Service API". The main content area features a map with several location markers and a red callout bubble. To the left, a sidebar lists categories: Web Service API (selected), Overview, Get API Key, Location Search, Location Input Hint, Forward/Reverse Geocoding, and Route Planning. Below the sidebar are four service examples: Location Search Service, Location Input Hint Service, Address Coding Service, and Route Planning Service, each with a small map diagram.

图：百度地图提供 API 接口供用户调取其交通出行数据

52



目录

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”
26

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

① 大数据时代下的交通规划

② 数据“菜谱”

③ 数据分析的“武器库”

④ 数据在业务中的应用案例

⑤ 再认识与展望

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”27

数据采集

数据预处理

业务分析

数据建模

数据仓库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望



谢逊取过手边的屠龙宝刀，拔刀出鞘，擦的一声，在大树的树干上斜砍一刀，只听得砰的一响，大树的上半段向外跌落。

工欲善其事，必先利其器
—《论语·卫灵公》



数据分析的“七种武器”

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”28

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- ① 数据采集
- ② 数据处理
- ③ 业务分析
- ④ 数据建模
- ⑤ 数据库
- ⑥ 分析算法
- ⑦ 可视化



数据采集

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

29

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- 利用服务器端提供的 API 接口调取数据
- 编写爬虫程序，从网页上抓取数据

52



数据采集

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

29

- 利用服务器端提供的 API 接口调取数据
- 编写爬虫程序，从网页上抓取数据

字段名称	字段含义	说明
alter	交通出行方式	推荐公交、巴士优先、骑车、小汽车、出租车等
O_lon & O_lat	出发地坐标	百度公司加密后的经纬度坐标
D_lon & D_lat	到达地坐标	百度公司加密后的经纬度坐标
distance	出行距离	包括步行在内的实际路网距离
duration	出行时间	包括步行在内
price	票价	
walkDistance	步行距离	
walkTime	步行时间	
transferNumber	换乘次数	

表：百度地图 API 返回的公交路径规划数据结果

52

数据采集

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武器库”

数据采集

数据预处理

业务分析

数据建模

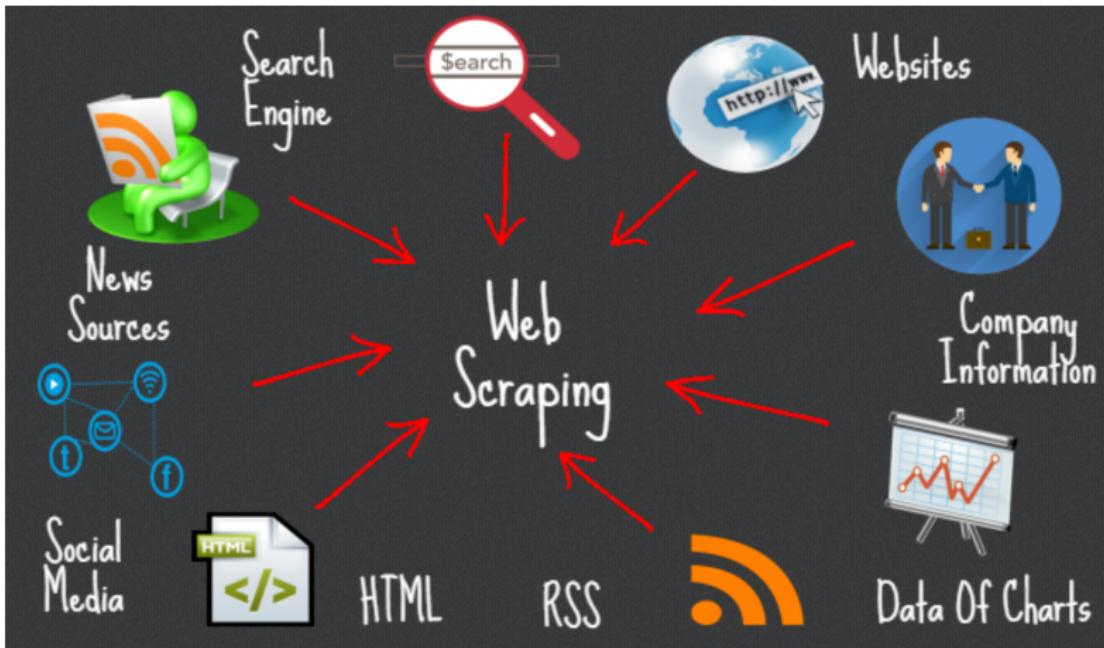
分析算法

可视化

数据在业务中的
应用案例

再认识与展望

29



图：利用爬虫程序，可以从丰富的互联网资源中自动化采集数据

52



数据预处理

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

30

52

- 现实中的原始数据都是不完整、不一致的脏数据，需要编写程序对数据进行**清洗、集成、变换和归约**等自动化处理
- 示例：GPS 数据预处理



数据预处理

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

30

- 现实中的原始数据都是不完整、不一致的脏数据，需要编写程序对数据进行 **清洗、集成、变换和归约** 等自动化处理
- 示例：GPS 数据预处理

算法：GPS 原始数据清洗

```
Data: 原始车牌 GPS 文件:GPS_File, 每行 line=(ID,Postion,Time,Direction,Speed,State)
Result: 清洗后的 GPS 文件:GPS_File_P
输入: 时间容差  $\Delta t$ , 距离容差  $\Delta d$ , 原始 GPS 数据列表 G
      /* 对 G 按照时间排序
      SortByTime(G);
      */
      2 bFirst  $\leftarrow$  true;
      3 for i = 1  $\leftarrow$  1 to N do
          /* 删除位置不在市域范围 Rect 的错误点
          if G[i].position  $\notin$  Rect then continue;
          if bFirst = true then
              GPS_File_P  $\leftarrow$  OutputResult(G[i]);
              bFirst  $\leftarrow$  true; tt  $\leftarrow$  G[i].time; tp  $\leftarrow$  G[i].position;
              /* 输出结果 */
          else
              /* 根据前后点时间容差和距离容差删除错误点
              if (5  $\leqslant$  (G[i].time - tt)  $\leqslant$   $\Delta t$ )  $\cap$  Distance(G[i].position, tp)  $\leqslant$   $\Delta d$  then
                  GPS_File_P  $\leftarrow$  OutputResult(G[i]);
                  tt  $\leftarrow$  G[i].time; tp  $\leftarrow$  G[i].position;
                  /* 输出结果 */
              end
          end
      end
      */

      30
```

52

数据预处理

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

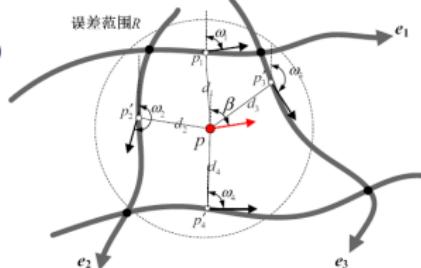
分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- 现实中的原始数据都是不完整、不一致的脏数据，需要编写程序对数据进行 **清洗、集成、变换和归约** 等自动化处理
- 示例：GPS 数据预处理



算法：GPS 与道路地图匹配

输入：路网模型 G 和 GPS 轨迹 $T: p_1 \rightarrow p_2 \rightarrow \dots \rightarrow p_n$

输出：处理后的 GPS 点序列 $P: c_1^{j1} \rightarrow c_2^{j2} \rightarrow \dots \rightarrow c_n^{jn}$

```

/* 清空列表 tList
1  tList ← empty;
2  for i = 1 to n do
3      /* 错误范围 R 内的候选路段集
4      s = GetCandidates(p_i, G, R);
5      tList.add(s);
6  end
7  G'_T = ConstructGraph(tList);           /* 构建图 G'_T */
8  return FindMatchedSequence(G'_T);
    */

```



业务分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

31



图：轨道刷卡数据分析的业务流程

52



数据建模

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- 根据业务需求对数据进行抽象，形成计算机能够理解的逻辑关系和物理结构
- 示例：车辆轨迹模型

32

52

数据建模

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

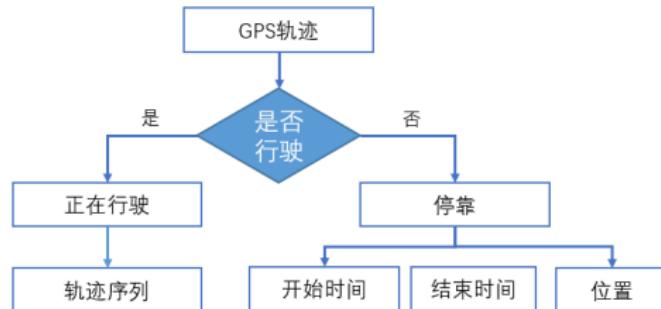
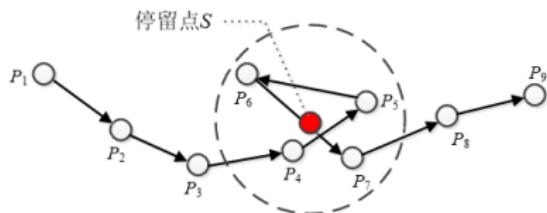
数据在业务中的
应用案例

再认识与展望

32

- 根据业务需求对数据进行抽象，形成计算机能够理解的逻辑关系和物理结构
- 示例：车辆轨迹模型

	经度	纬度	时间
P_1 :	Lat_1	Lon_1	T_1
P_2 :	Lat_2	Lon_2	T_2
P_n :	Lat_n	Lon_n	T_n



图：车辆轨迹模型



数据库

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

33



图：数据库技术的核心是对数据进行高效组织存储，以及多终端并发查询访问

52



数据库

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

33

- 通过高效的组织和存储，实现数据“增删改查”操作
- 根据数据量级、数据内容、应用场景选择最适合的数据库技术
- 数据库是大数据最核心的技术之一，目前主流采用分布式存储方案来解决



图：根据实际需求选择最合适的数据库

52



数据库

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

33

- 通过高效的组织和存储，实现数据“增删改查”操作
- 根据数据量级、数据内容、应用场景选择最适合的数据库技术
- 数据库是大数据最核心的技术之一，目前主流采用**分布式存储方案**来解决



图：目前主流的大数据数据库产品

52



分析算法

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- 算法不是简单的统计，而要能挖掘数据的规律和关系
- 交通规划中最常用的分析算法是空间分析算法
- 示例：核密度分析



分析算法

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

- 算法不是简单的统计，而要能挖掘数据的规律和关系
- 交通规划中最常用的分析算法是**空间分析算法**
- 示例：核密度分析

34

52

分析算法

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

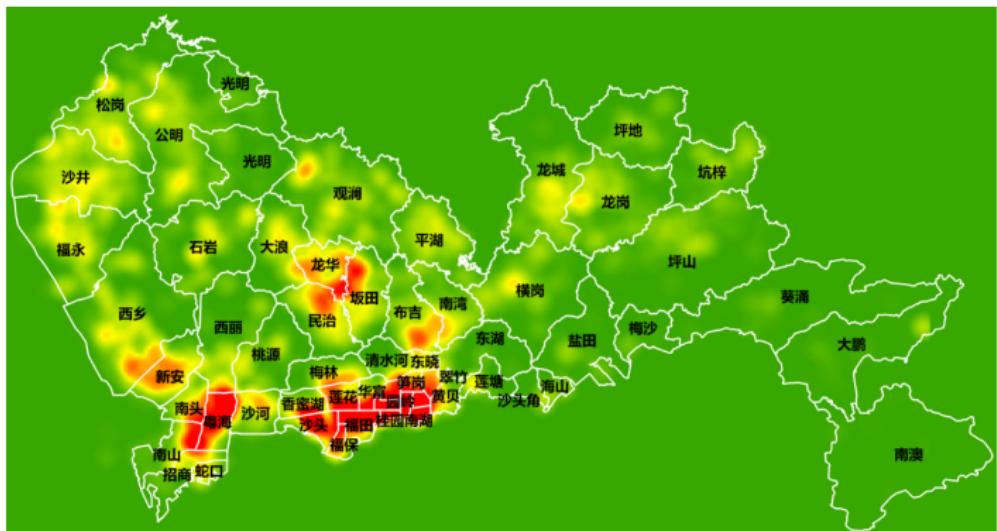
可视化

数据在业务中的
应用案例

再认识与展望

34

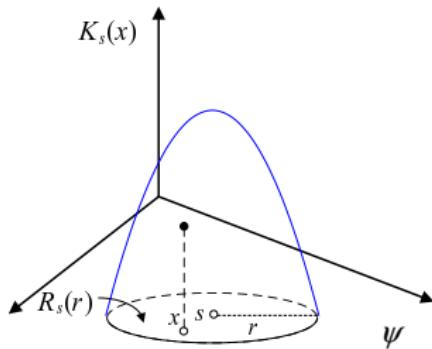
- 算法不是简单的统计，而要能挖掘数据的规律和关系
- 交通规划中最常用的分析算法是**空间分析算法**
- 示例：核密度分析



图：核密度分析

52

34



带权重的核密度估计^{1,2}

$$\hat{\lambda}(s) = \frac{1}{n\pi r^2} \sum_{i=1}^n w_i K_s \left(\frac{d_{sx}}{r} \right)$$

其中, $d_{sx} = \|s - x_i\|$ 表示位置 s 和第 i 个样本 x_i 之间的欧式距离; r 被称为带宽, 表示区域 $R_s(r)$ 的面积大小。

- KDE 的核心思想就是通过计算单位面积内样本的核密度平滑处理来估计位置的属性值
- KDE 结果表示区域 $R_s(r)$ 内不同样本对估计 s 位置属性值的权重
- 核函数是一个与距离有关的函数, 而且其赋予区域 $R_s(r)$ 内每个样本点的值是不同的, 离中心位置 s 越远的样本点, 其值应该越小, 而离中心位置 s 越近的样本点, 其值应该越大
- 对 KDE 结果产生影响的两个因素包括核函数的形式以及带宽的选择

¹ Silverman B W. 1986. Density Estimation for Statistics and Data Analysis[M]. London: Chapman Hall.

² Wand M P, Jones M C. 1995. Kernel Smoothing[M]. Boca Raton, Florida: Chapman & Hall/CRC.

52

分析算法

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集
数据预处理
业务分析
数据建模

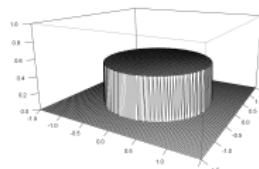
数据库
分析算法
可视化

数据在业务中
的应用案例

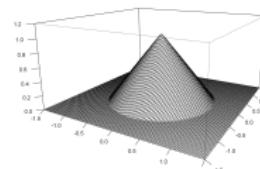
再认识与展望

34

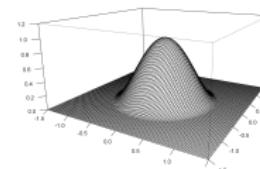
- 算法不是简单的统计，而要能挖掘数据的规律和关系
- 交通规划中最常用的分析算法是**空间分析算法**
- 示例：核密度分析



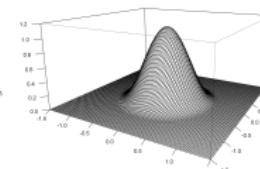
(a) Uniform



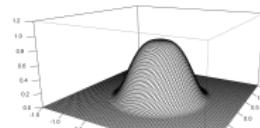
(b) Triangular



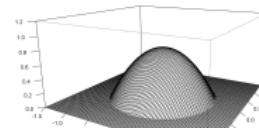
(c) Quartic



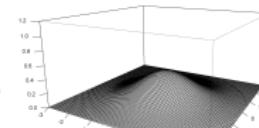
(d) Triweight



(e) Tricube



(f) Epanechnikov



(g) Gaussian

图：常见的核函数

52

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

数据在业务中的
应用案例

再认识与展望

35



图：Edward Tufte(1942-), 美国统计学家，数据可视化理论的先驱者和领军人物，人称“数据达芬奇”

52

The commonality between science and art is in trying to see profoundly - to develop strategies of seeing and showing.

—Edward Tufte



可视化

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

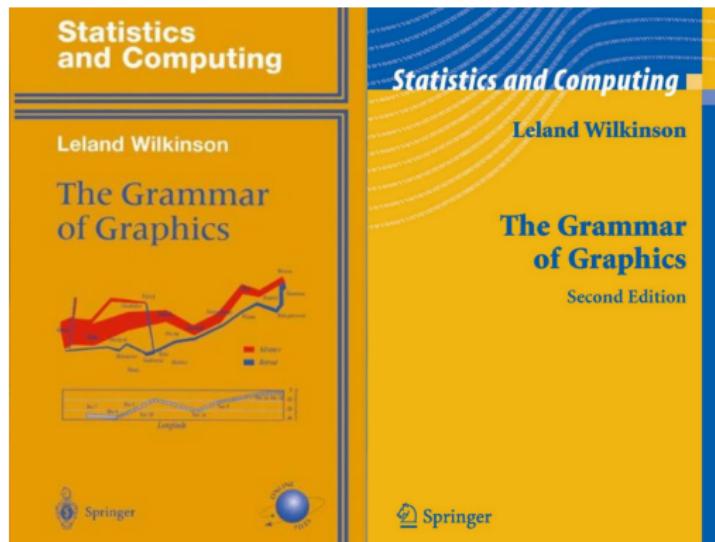
35

数据在业务中的
应用案例

再认识与展望

52

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：变形地图 (cartogram)



图：可视化领域经典著作《The Grammar of Graphics》的第一版 (1999) 和第二版 (2005)

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

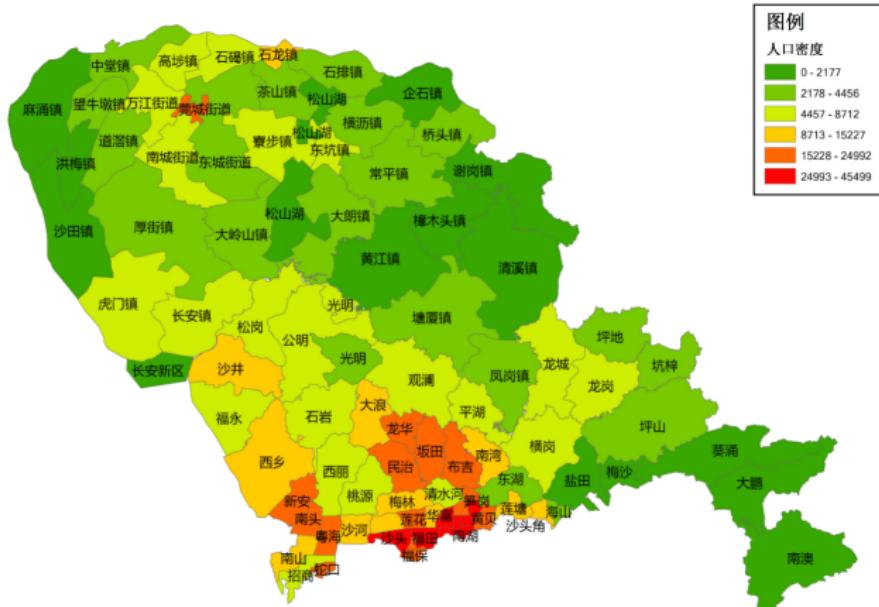
可视化

35

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：变形地图 (cartogram)



图：深莞街道人口密度

52

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

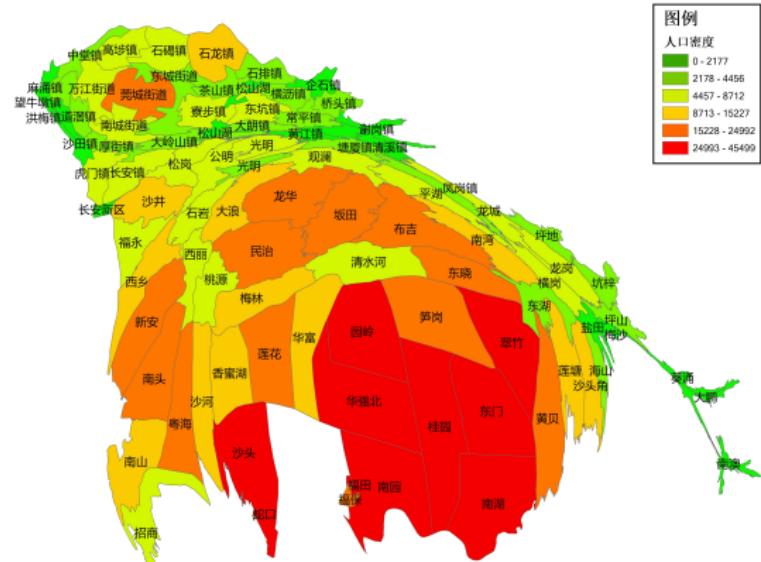
可视化

35

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：变形地图 (cartogram)



图：深莞街道人口密度变形地图 (cartogram)

52

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

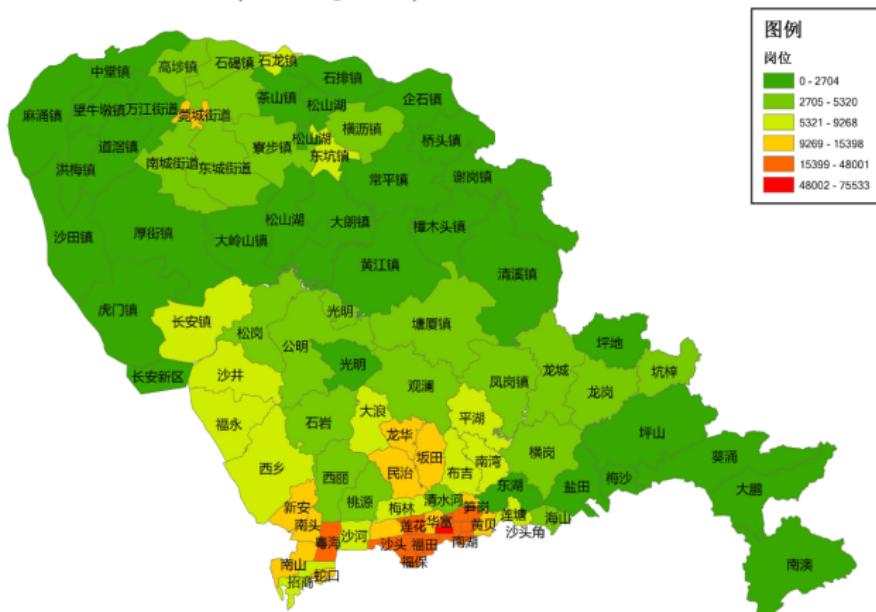
可视化

35

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：变形地图 (cartogram)



图：深莞街道岗位密度

52

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

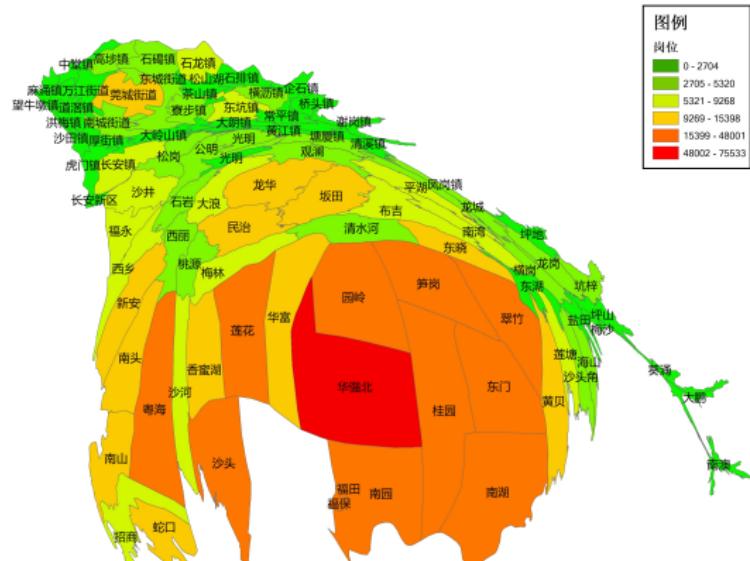
可视化

35

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：变形地图 (cartogram)



图：深莞街道岗位密度变形地图 (cartogram)

52



可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

36

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：交通流量图



可视化

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

可视化

36

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：交通流量图

52

可视化

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据采集

数据预处理

业务分析

数据建模

数据库

分析算法

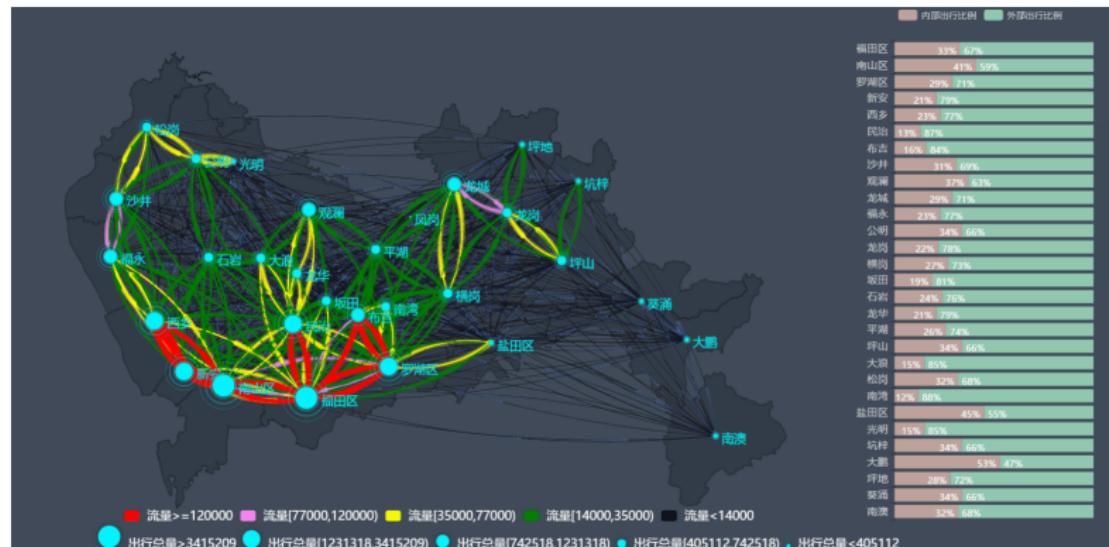
可视化

36

数据在业务中的
应用案例

再认识与展望

- 计算机技术与艺术的结合
- 将绘图与数据分离，数据相关绘图与数据无关绘图分离
- 示例：交通流量图



图：在一张图中同时展现总流量、OD量和内外部出行比例，数据在外部单独存储，与样式无关

52



目录

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

37

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

① 大数据时代下的交通规划

② 数据“菜谱”

③ 数据分析的“武器库”

④ 数据在业务中的应用案例

⑤ 再认识与展望

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

38

● 出租车

- ① 客流量分析（乘次、产生量和吸引量统计）
- ② 出行分析（OD 矩阵、OD 行程时间和出行距离）
- ③ 运营分析（空驶比例、路段行程车速）

● 定点车流

- ① 流量统计
- ② 流量时变分析
- ③ 流向分析（各关口每日出入关的总流量）

● 常规公交

- ① 客流量分析（线路、站点、换乘）
- ② 出行分析（站点 OD、出行 OD）
- ③ 运营分析（车速、发车频率、候车时间）
- ④ 可达性分析
- ⑤ 站点覆盖范围分析

● 轨道

- ① 客流量统计
- ② 出行统计（站点 OD、出行 OD、OD 行程时间、出行距离）
- ③ 运营统计（线路拥挤度、站点拥挤度）

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

38

● 出租车

- ① 客流量分析（乘次、产生量和吸引量统计）
- ② 出行分析（OD 矩阵、OD 行程时间和出行距离）
- ③ 运营分析（空驶比例、路段行程车速）

● 定点车流

- ① 流量统计
- ② 流量时变分析
- ③ 流向分析（各关口每日出入关的总流量）

● 常规公交

- ① 客流量分析（线路、站点、换乘）
- ② 出行分析（站点 OD、出行 OD）
- ③ 运营分析（车速、发车频率、候车时间）
- ④ 可达性分析
- ⑤ 站点覆盖范围分析

● 轨道

- ① 客流量统计
- ② 出行统计（站点 OD、出行 OD、OD 行程时间、出行距离）
- ③ 运营统计（线路拥挤度、站点拥挤度）

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

38

● 出租车

- ① 客流量分析（乘次、产生量和吸引量统计）
- ② 出行分析（OD 矩阵、OD 行程时间和出行距离）
- ③ 运营分析（空驶比例、路段行程车速）

● 定点车流

- ① 流量统计
- ② 流量时变分析
- ③ 流向分析（各关口每日出入关的总流量）

● 常规公交

- ① 客流量分析（线路、站点、换乘）
- ② 出行分析（站点 OD、出行 OD）
- ③ 运营分析（车速、发车频率、候车时间）
- ④ 可达性分析
- ⑤ 站点覆盖范围分析

● 轨道

- ① 客流量统计
- ② 出行统计（站点 OD、出行 OD、OD 行程时间、出行距离）
- ③ 运营统计（线路拥挤度、站点拥挤度）

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

38

● 出租车

- ① 客流量分析（乘次、产生量和吸引量统计）
- ② 出行分析（OD 矩阵、OD 行程时间和出行距离）
- ③ 运营分析（空驶比例、路段行程车速）

● 定点车流

- ① 流量统计
- ② 流量时变分析
- ③ 流向分析（各关口每日出入关的总流量）

● 常规公交

- ① 客流量分析（线路、站点、换乘）
- ② 出行分析（站点 OD、出行 OD）
- ③ 运营分析（车速、发车频率、候车时间）
- ④ 可达性分析
- ⑤ 站点覆盖范围分析

● 轨道

- ① 客流量统计
- ② 出行统计（站点 OD、出行 OD、OD 行程时间、出行距离）
- ③ 运营统计（线路拥挤度、站点拥挤度）

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

● 居民出行调查

- ① 人口、就业、年龄、收入等分布
- ② 交通方式和交通目的的比例
- ③ 分交通方式和交通目的的出行量
- ④ 通勤交通分析

● 手机

- ① 就业和居住人口分布
- ② 不分交通方式和目的的出行量
- ③ 通勤交通分析

39

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

39

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

● 居民出行调查

- ① 人口、就业、年龄、收入等分布
- ② 交通方式和交通目的的比例
- ③ 分交通方式和交通目的的出行量
- ④ 通勤交通分析

● 手机

- ① 就业和居住人口分布
- ② 不分交通方式和目的的出行量
- ③ 通勤交通分析

52



常规交通指标分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

39

● 居民出行调查

- ① 人口、就业、年龄、收入等分布
- ② 交通方式和交通目的的比例
- ③ 分交通方式和交通目的的出行量
- ④ 通勤交通分析

● 手机

- ① 就业和居住人口分布
- ② 不分交通方式和目的的出行量
- ③ 通勤交通分析

多种数据融合使用

当遇到同一指标存在多种数据源都可以计算的情况，针对应用场景选择最合适的数据进行计算

52



都市圈范围分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

40

- 数据:2010 和 2016 年居民出行调查

- 分析指标

通勤率定义

通勤率 = 区域 R 至中心区域通勤人数 / 区域 R 常住人口

- 解决问题: 深圳市城市边界的空间增长与发展规律

52



都市圈范围分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 数据:2010 和 2016 年居民出行调查
- 分析指标

通勤率定义

$$\text{通勤率} = \text{区域 R 至中心区域通勤人数} / \text{区域 R 常住人口}$$

- 解决问题: 深圳市城市边界的空间增长与发展规律

40

52



都市圈范围分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 数据:2010 和 2016 年居民出行调查
- 分析指标

通勤率定义

$$\text{通勤率} = \text{区域 R 至中心区域通勤人数} / \text{区域 R 常住人口}$$

- 解决问题: 深圳市城市边界的空间增长与发展规律

40

52



都市圈范围分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

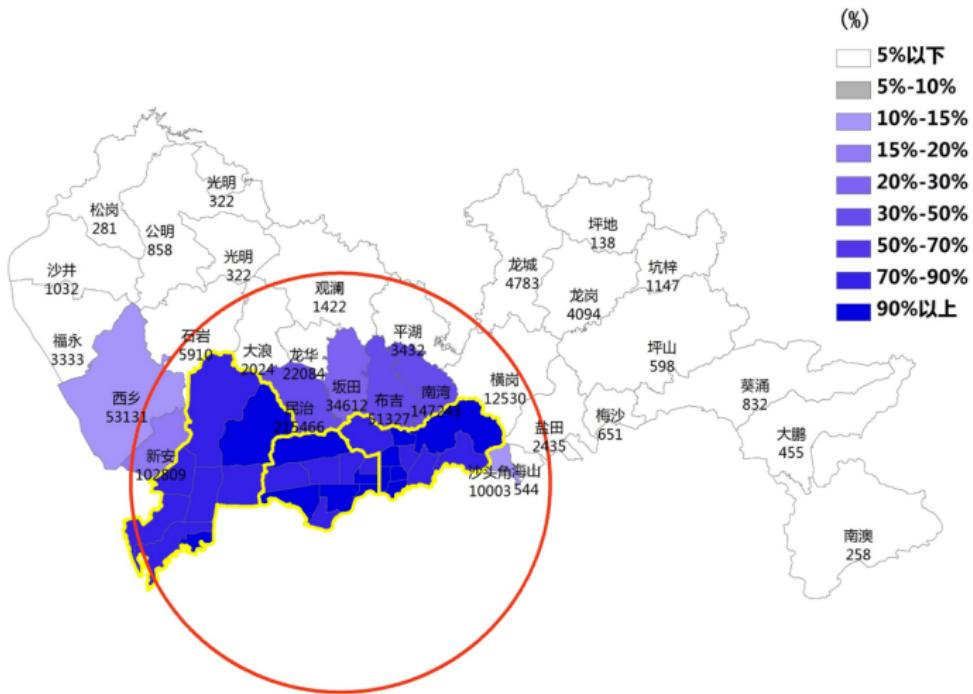
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

41



图：2010 年中心区影响范围（圆心为中心区中心点，圆半径 20 公里）

52

- (%)
- 5%以下
 - 5%-10%
 - 10%-15%
 - 15%-20%
 - 20%-30%
 - 30%-50%
 - 50%-70%
 - 70%-90%
 - 90%以上

都市圈范围分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

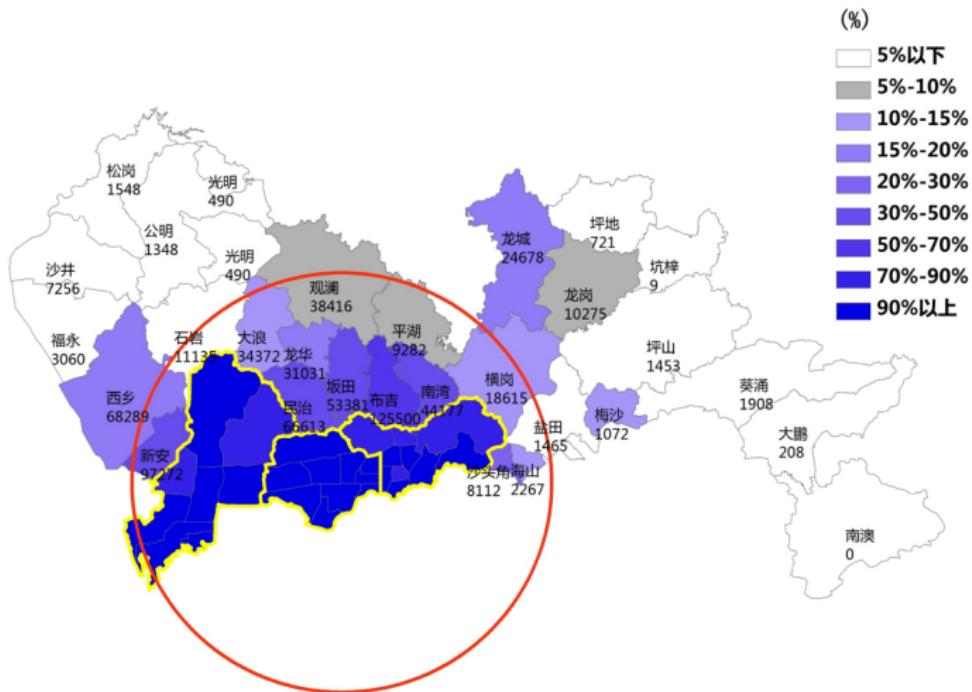
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

41



图：2016 年中心区影响范围（圆心为中心区中心点，圆半径 20 公里）

52



区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 数据：2016 年手机数据处理后的交通出行量
- 解决问题：从交通出行角度探讨深莞边界区域的联系，以及交
通组团的划分

42

52

区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

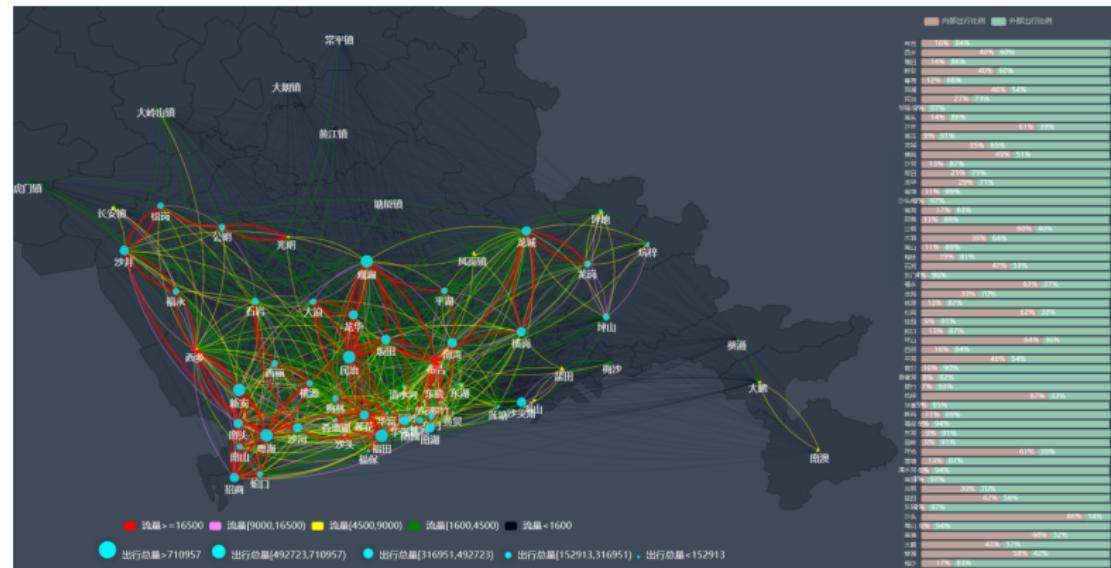
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

42



图：深莞交通出行

52

区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

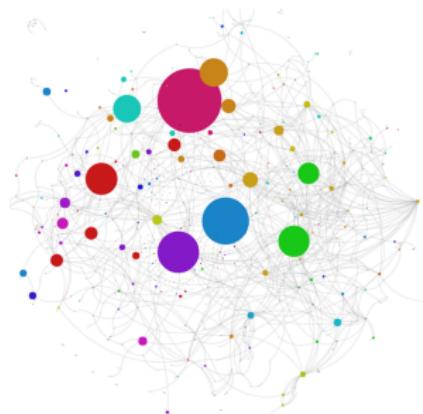
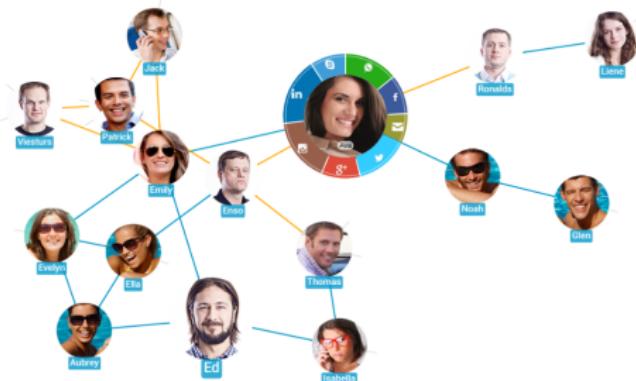
数据在业务中的
应用案例

常规交通指标分析
都市圈范围分析

区域联系强度分析
货运交通分析

再认识与展望

43



图：社交网络与网络分析算法

52



区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 网络分析中的社区发现 (Community Detection) 算法
- 分析指标: 模块度 (Modularity)

社区划分的目标是使得划分后的社区内部的连接较为紧密，而在社区之间的连接较为稀疏，网络分析算法中通过**模块度**指标可以刻画这样的划分的优劣，模块度越大，则社区划分的效果越好

43

52



区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

43

- 网络分析中的社区发现 (Community Detection) 算法
- 分析指标：模块度 (Modularity)

模块度定义¹

网络中连接社区结构内部顶点的边占的比例与另一个随机网络中连接社区结构内部顶点的边所占比例的期望值的差值.

$$Q = \frac{1}{2m} \sum_{i,j} \left[\omega_{i,j} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) = \sum_s \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 \right]$$

其中， $m = \frac{1}{2} \sum_{i,j} \omega_{i,j}$ 表示网络中所有边的权重之和， $\omega_{i,j}$ 表示节点 i 和节点 j 之间的权重； $k_i = \sum_j \omega_{i,j}$ 表示与顶点 i 连接的所有边的权重； c 表示的是顶点被分配到的社区， $\delta(s_i, s_j)$ 用于判断节点 i 与节点 j 是否被划分在同一个社区中，若是则返回 1，否则返回 0

\sum_{in} 表示社区 c 内部的权重之和， \sum_{tot} 表示的是与社区 c 内部所有点的连接边的权重之和（包括社区内部的边和社区外部的边）

¹ M.E.J. Newman and M.Girvan. 2004. Finding and evaluating community structure in networks[J]. Phys. Rev. E 69, 026113.



区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 网络分析中的社区发现 (Community Detection) 算法
- 分析指标：模块度 (Modularity)

算法：FastUnfolding 算法¹

输入：图 $G = \{V, E, \omega\}$, 节点 V 的数目是 m , 边 E 的数目是 n , ω 是节点之间的权重

输出：分类后的社区集合 $C = \{c_1, c_2, \dots, c_k\}$

```

1  for i ∈ V do C.add(ci);                                /* 把每个节点初始化到不同的社区 */
2  ;
3  best_gain ← 0;
4  token ← 0;
5  while true do
6      for i ∈ V do
7          C.remove(ci);
8          /* 计算将节点 i 合并到邻接节点 j 所属社区 cj 后的模块度增益 ΔQ */
9          for j ∈ neighbor of i do
10             c = C[j];
11             if ΔQ > best_gain then
12                 best_gain ← ΔQ;
13                 best_community ← c;
14             end
15             C.add(best_community);
16         end
17         if ci ≠ best_community then token = 1;
18     end
19   if token = 0 then break;

```

¹ V.D. Blondel, J.L. Guillaume, R. Lambiotte, E. Lefebvre. 2008. Fast unfolding of communities in large networks[J]. J. Stat. Mech: Theory Exp.

区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

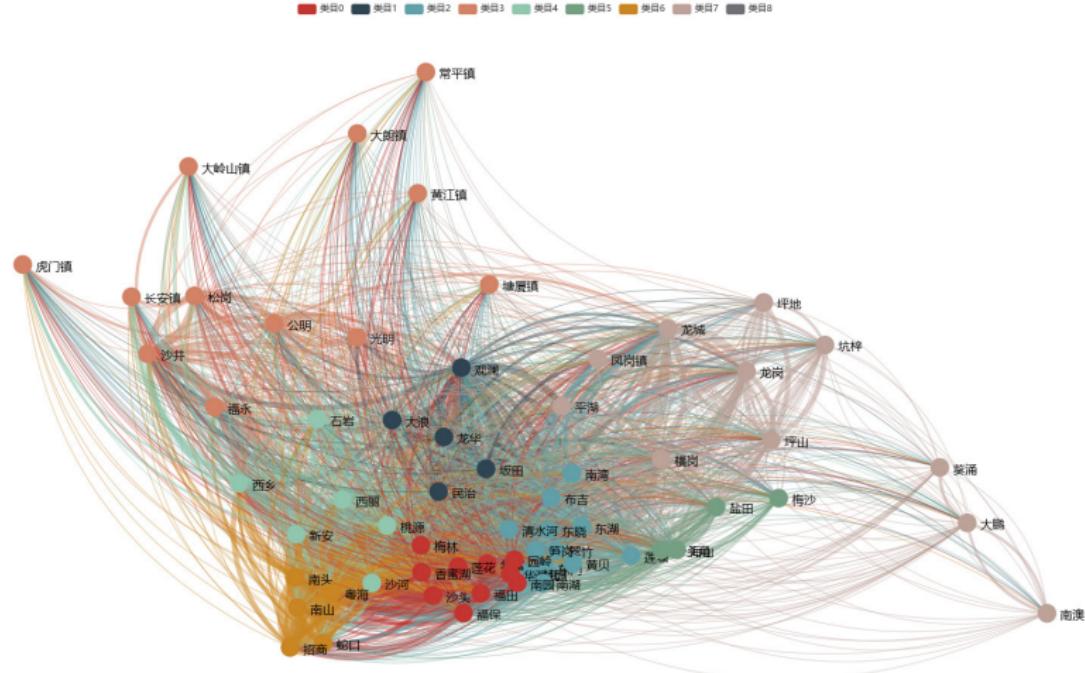
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

44



图：基于交通出行联系的社区发现算法实现

52



区域联系强度分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

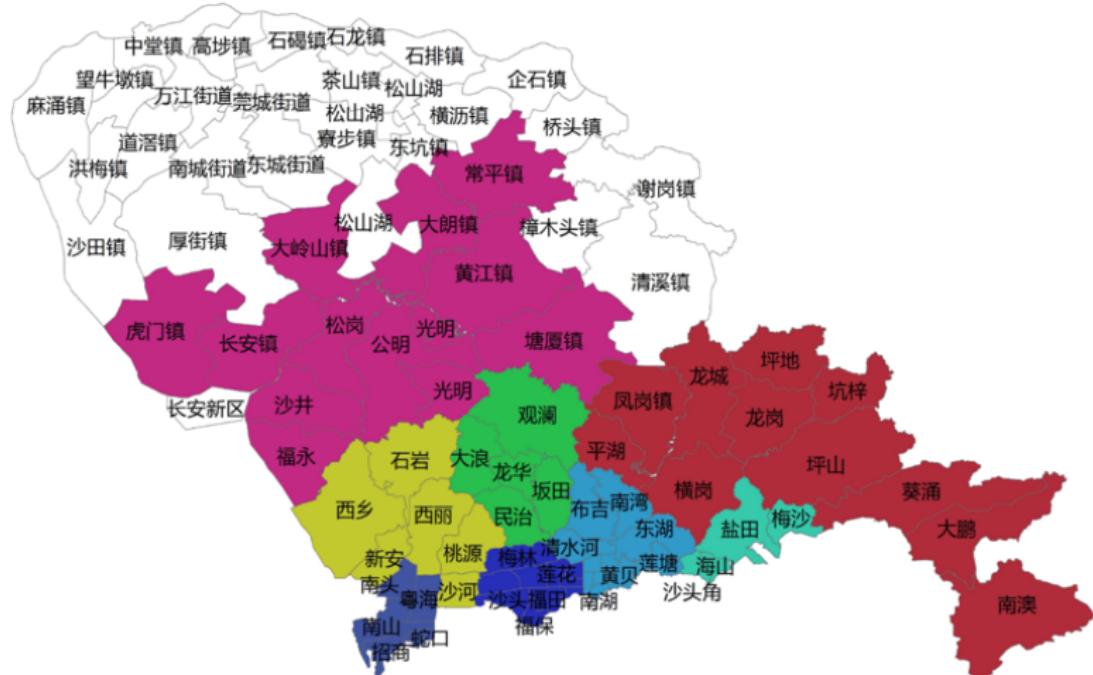
都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

44



图：深圳与东莞的新组团范围（不同颜色表示模块度最大的社区）

52



货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系

45

52

货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

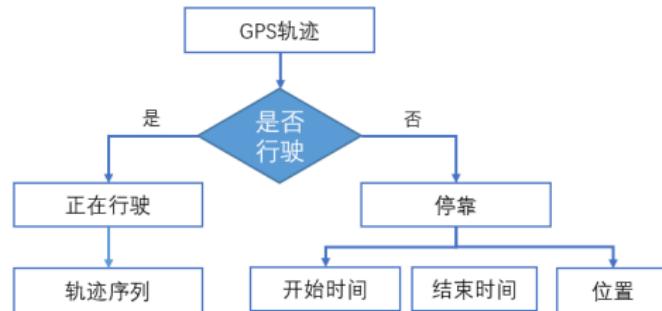
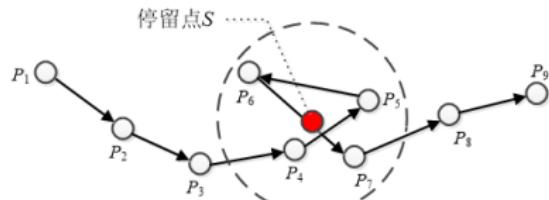
货运交通分析

再认识与展望

45

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系

	经度	纬度	时间
P_1 :	Lat_1	Lon_1	T_1
P_2 :	Lat_2	Lon_2	T_2
.....			
P_n :	Lat_n	Lon_n	T_n



图：车辆轨迹模型

52



货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系

算法：停留位置识别

输入：GPS 点序列 P , 距离阈值 $distThreh$, 时间阈值 $timeThreh$

输出：停留点集合 $SP = \{S\}$

```
1 i = 0; pointNum = |P|;
2 while i < pointNum do
3     j ← i + 1; Label = 0;
4     while j < pointNum do
5         dist = Distance( $p_i, p_j$ );
6         if dist > distThreh then
7              $\Delta T \leftarrow p_j.T - p_i.T$ ;
8             if  $\Delta T > timeThreh$  then
9                 S.coord ← ComputeMeanCoord( $p_k | i \leq k \leq j$ ); /* 计算静止点集的中心坐标 */
10                S.arrvT ←  $p_i.T$ ; S.levT ←  $p_j.T$ ;
11                SP.insert(S);
12                i ← j; Label ← 1;
13            end
14            break;
15        end
16        j = j + 1;
17    end
18    if Label ≠ 1 then i = i + 1;
19 end
20 return SP;
```



货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

45

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系



图：深圳市货车在全国的货场分布

52

货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

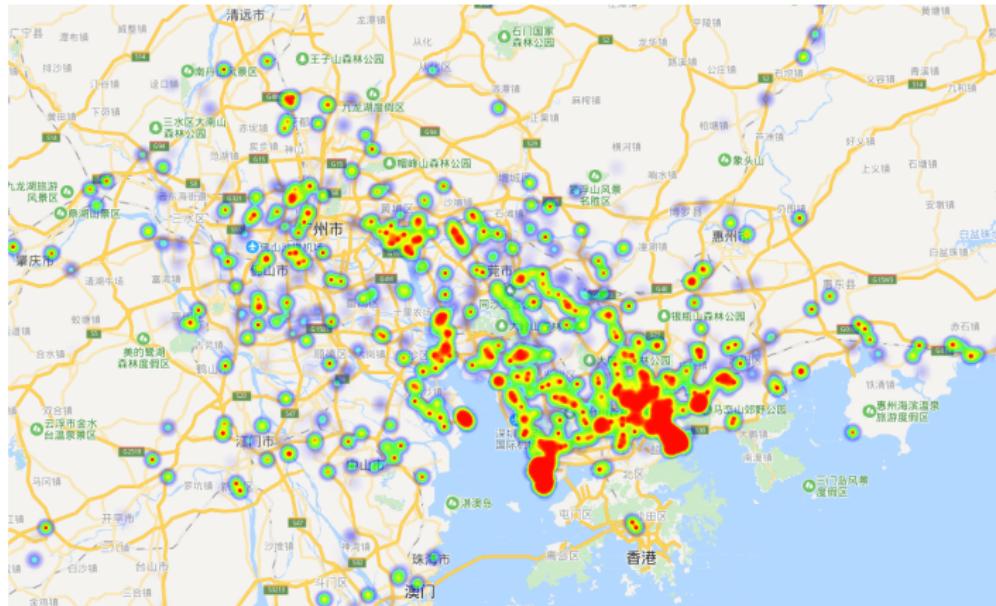
区域联系强度分析

货运交通分析

再认识与展望

45

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系



图：深圳市货车在珠三角的货场分布

52

货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

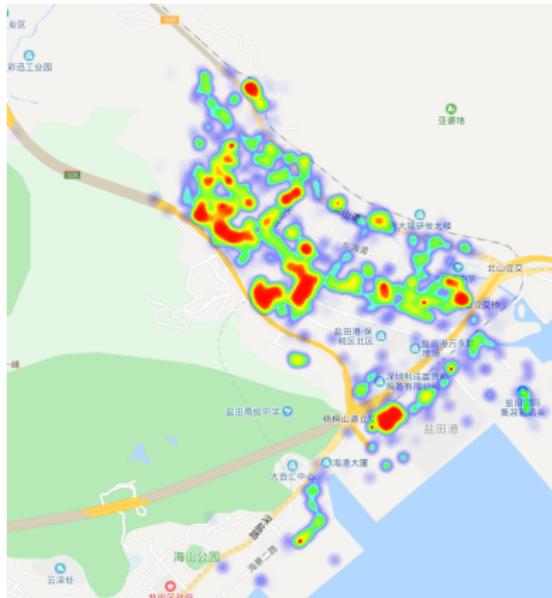
区域联系强度分析

货运交通分析

45

再认识与展望

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系



图：深圳市货车在盐田港区的货场分布

52

货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

再认识与展望

45



图：车辆轨迹地图

52

货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

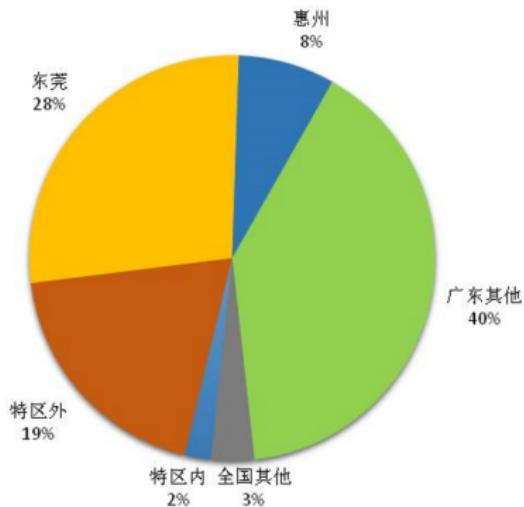
货运交通分析

45

再认识与展望

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系

全年各地货源比例



图：除了空间表达形式之外，还可以分析各类统计指标

52



货运交通分析

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

常规交通指标分析

都市圈范围分析

区域联系强度分析

货运交通分析

45

再认识与展望

- 数据：货车 GPS 数据
- 解决问题：深圳市对内和对外货运交通的联系

物流 节点区域	运输 方向	大铲湾 港区	西部 港区	东部 港区	平湖 物流园	机场 物流区	龙华 物流园	横岗 物流区	大浪 物流区	西丽 物流区
大铲湾港区	起点		17.02%	12.77%	0.00%	53.19%	0.00%	4.26%	2.13%	10.64%
	终点		24.07%	20.37%	5.56%	31.48%	0.00%	7.41%	1.85%	9.26%
西部港区	起点	3.94%		38.79%	11.21%	5.76%	7.58%	6.36%	18.18%	8.18%
	终点	2.99%		47.01%	9.70%	6.72%	5.60%	7.46%	13.43%	7.09%
东部港区	起点	3.46%	39.62%		11.32%	5.66%	7.86%	26.73%	3.77%	1.57%
	终点	1.85%	39.38%		8.62%	3.69%	4.92%	38.77%	2.15%	0.62%
平湖物流园	起点	3.61%	31.33%	33.73%		4.82%	12.05%	8.43%	1.20%	4.82%
	终点	0.00%	36.63%	35.64%		1.98%	11.88%	8.91%	2.97%	1.98%
机场物流区	起点	27.42%	29.03%	19.35%	3.23%		6.45%	6.45%	3.23%	4.84%
	终点	30.12%	22.89%	21.69%	4.82%		9.64%	7.23%	2.41%	1.20%
龙华物流园	起点	0.00%	25.00%	26.67%	20.00%	13.33%		1.67%	13.33%	0.00%
	终点	0.00%	30.86%	30.86%	12.35%	4.94%		13.58%	7.41%	0.00%
横岗物流区	起点	2.25%	11.24%	70.79%	5.06%	3.37%	6.18%		1.12%	0.00%
	终点	1.65%	17.36%	70.25%	5.79%	3.31%	0.83%		0.83%	0.00%
大浪物流区	起点	1.69%	61.02%	11.86%	5.08%	3.39%	10.17%	1.69%		5.08%
	终点	1.14%	68.18%	13.64%	1.14%	2.27%	9.09%	2.27%		2.27%
西丽物流区	起点	16.13%	61.29%	6.45%	6.45%	3.23%	0.00%	0.00%	6.45%	
	终点	10.64%	57.45%	10.64%	8.51%	6.38%	0.00%	0.00%	6.38%	

图：除了空间表达形式之外，还可以分析各类统计指标

52



目录

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

46

① 大数据时代下的交通规划

② 数据“菜谱”

③ 数据分析的“武器库”

④ 数据在业务中的应用案例

⑤ 再认识与展望

52



加大现有数据的利用率和产品化思维

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

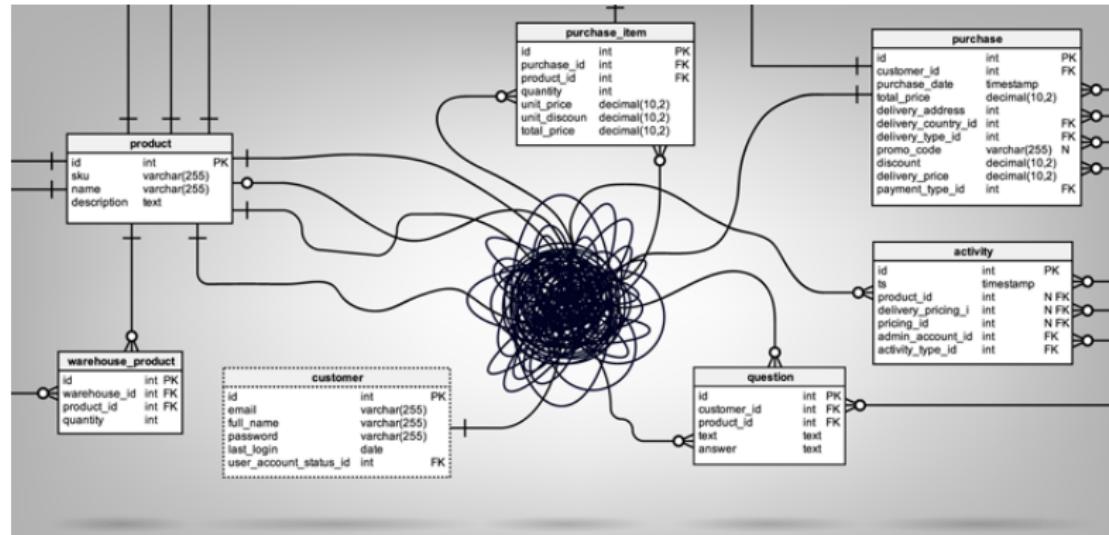
数据在业务中的
应用案例

再认识与展望

47

● 面向业务的数据整合

● 查询功能设计与开发



图：数据整合要梳理数据结构特征，设计面向业务的数据库对象

52



加大现有数据的利用率和产品化思维

交通规划行业中
的数据应用现状
及思考

- 面向业务的数据整合
- 查询功能设计与开发

大数据时代的
交通规划

数据“菜谱”

数据分析的“武器库”

数据在业务中的
应用案例

再认识与展望

47

文件	编辑	插入	页脚	公式	数据	审核	视图	ExoCharts	LOAD TEST	Power Pivot	TEAMS	反馈	常规	常规	常规	常规	常规	常规	常规	常规	常规	常规	常规	常规		
transfer.xls 20170525.xlsx - Excel												开始	结束	反馈	操作	筛选	格式	排序	筛选	操作	筛选	格式	排序	筛选		
	新选项卡	快速访问	文件	插入	页脚	公式	数据	审核	视图	ExoCharts	LOAD TEST	Power Pivot	TEAMS	反馈	常规											
1	步行时耗	乘车时耗	车平均耗	frun2x	toRun2x	中段时耗	24小时	时耗	里程	方式	是否	AM	FM	AM	FM	AM	FM	AM	FM	AM	FM	AM	FM	AM		
2																										
3	2005	2081	62.50786 8:05				8 AM			8 AM																
4	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
5	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
6	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
7	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
8	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
9	2005	2081	62.50786 8:15	8 AM			8 AM			8 AM																
10	2005	2081	62.50786 15:10	8 AM			8 AM			8 AM																
11	2005	2081	62.50786 18:15	19 FM			7 AM			7 AM																
12	2005	2081	62.50786 15:10	8 AM			8 AM			8 AM																
13	2005	2081	62.50786 15:10	19 FM			7 AM			7 AM																
14	2005	2081	62.50786 8:30	8 AM			8 AM			8 AM																
15	2005	2081	62.50786 18:30	19 FM			7 AM			7 AM																
16	2005	2081	62.50786 8:30	8 AM			8 AM			8 AM																
17	2005	2081	62.50786 18:30	19 FM			7 AM			7 AM																
18	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
19	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
20	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
21	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
22	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
23	2005	2081	62.50786 18:05	19 FM			7 AM			7 AM																
24	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
25	2005	2081	62.50786 9:35	9 AM			1步行			9 AM																
26	2005	2081	62.50786 11:15	11 OP			8 AM			8 AM																
27	-1	2005	62.50786 19:15	19 OP			8 AM			8 AM																
28	-1	2005	62.50786 19:15	19 OP			8 AM			8 AM																
29	2005	-1	2005	62.50786 19:15	19 OP		8 AM			8 AM																
30	-1	2005	62.50786 19:15	19 OP			8 AM			8 AM																
31	2005	-1	2005	62.50786 19:15	19 OP		8 AM			8 AM																
32	2005	-1	2005	62.50786 19:15	19 OP		8 AM			8 AM																
33	2005	2081	62.50786 10:07	10 OP			7 AM			7 AM																
34	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
35	2005	2081	62.50786 10:07	10 OP			7 AM			7 AM																
36	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
37	2005	2081	62.50786 10:07	10 OP			7 AM			7 AM																
38	2005	2081	62.50786 8:05	8 AM			8 AM			8 AM																
39	2005	2081	62.50786 10:07	10 OP			7 AM			7 AM																
40	2005	5031	105.558 9:15	9 OP			8 AM			8 AM																
41	5031	5106	105.558 10:15	10 PM			7 AM			7 AM																
42																										

图：目前还是基于 excel 进行手动的数据查询分析，没有形成产品化的工具，在日常工作中很难推广和复用



加大现有数据的利用率和产品化思维

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

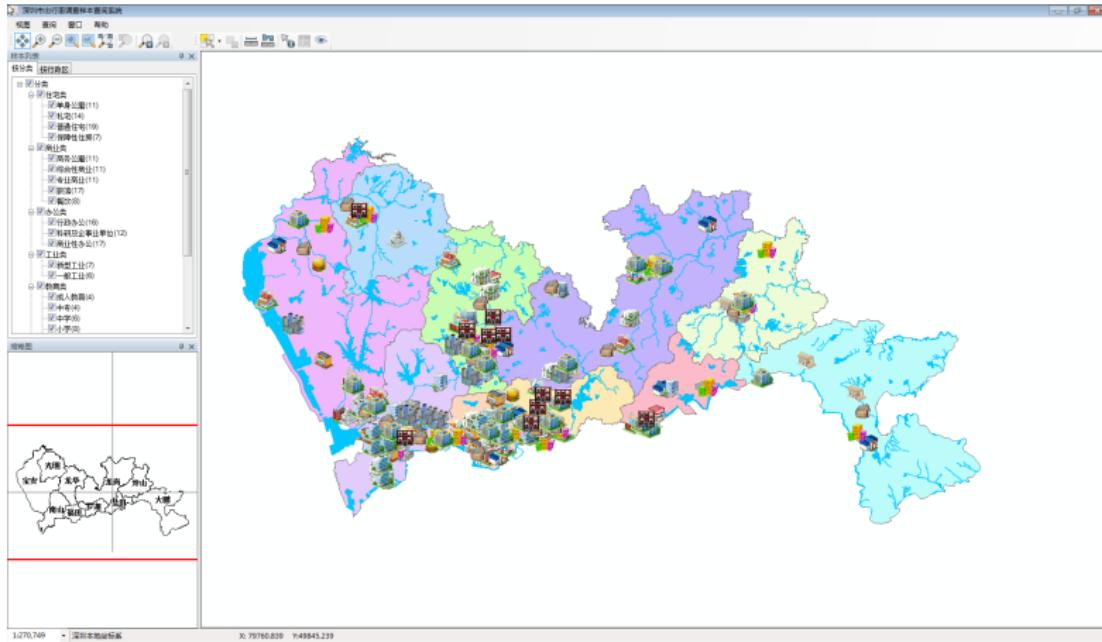
数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

47



图：基于 GIS 开发独立的面向业务人员的数据查询系统

52

加大现有数据的利用率和产品化思维

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

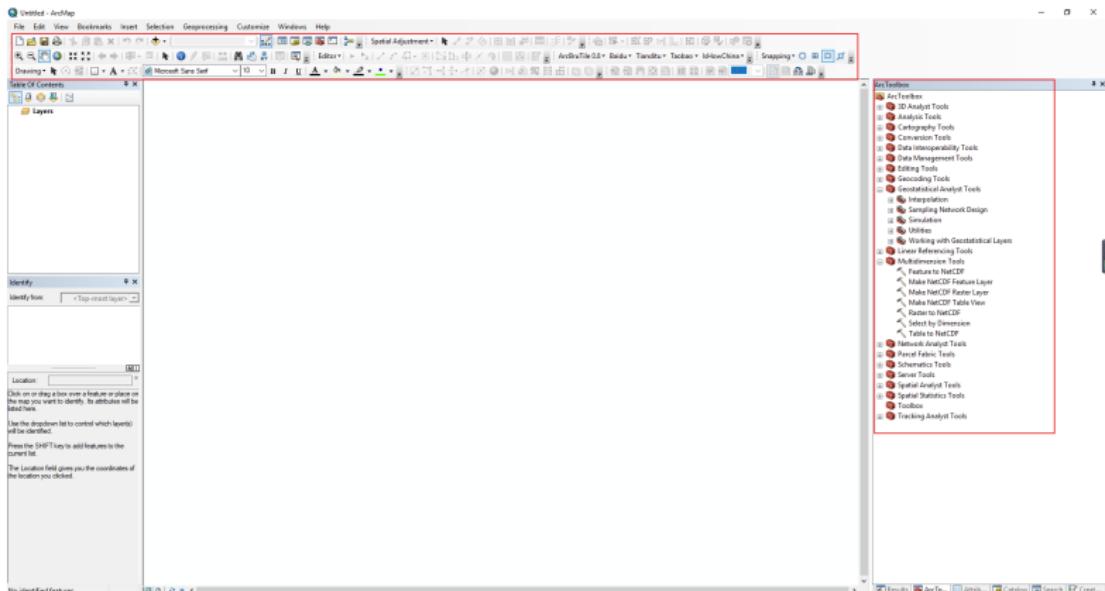
数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

47



图：在熟悉的工作平台中以插件的方式实现数据查询

52

提升数据分析的深度

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

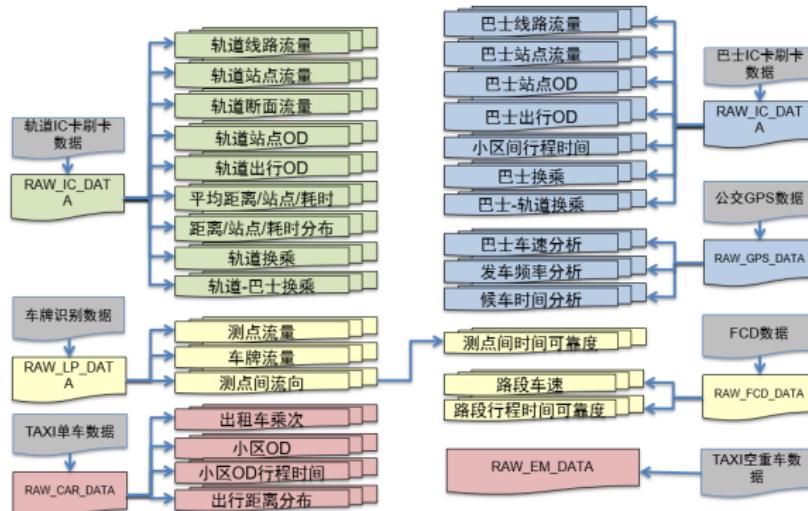
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

48

- 现有数据应用主要在传统交通模型和现状的简单指标统计，没有形成分析方法体系，**分析程度非常浅**
- 交通模型体系在大数据环境下的理论创新，以及空间分析、机器学习等技术手段在交通规划行业的广泛应用



图：仿真二期系统中同济设计的交通专项统计指标

52

提升数据分析的深度

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

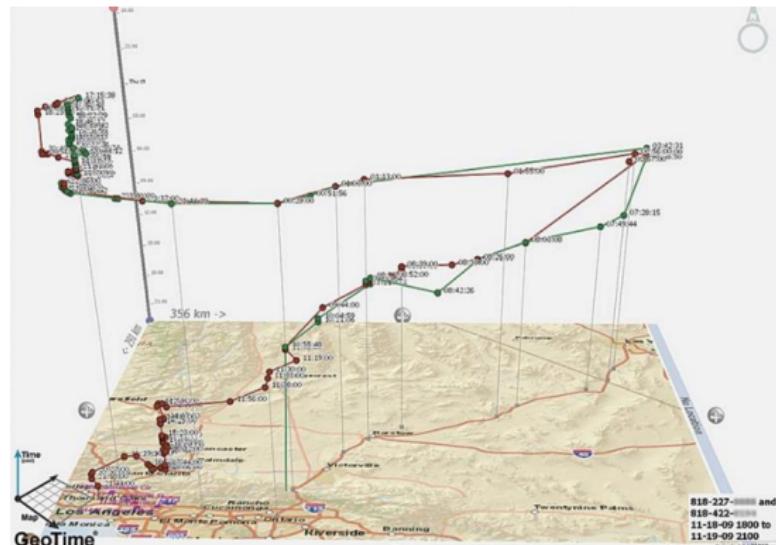
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

48

- 现有数据应用主要在传统交通模型和现状的简单指标统计，没有形成分析方法体系，**分析程度非常浅**
- 交通模型体系在大数据环境下的理论创新，以及空间分析、机器学习等技术手段在交通规划行业的广泛应用



图：交通模型理论的创新

52

提升数据分析的深度

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

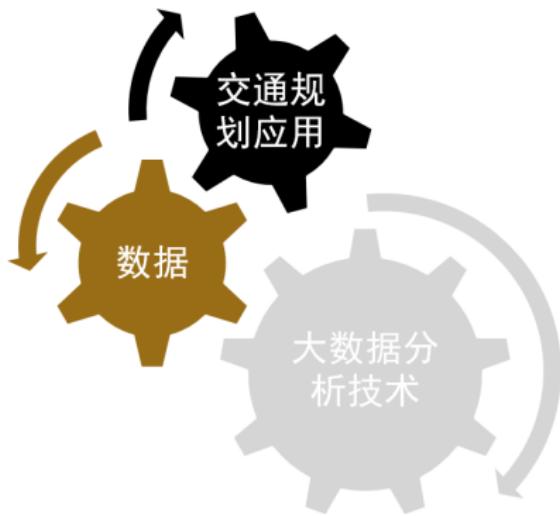
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

48

- 现有数据应用主要在传统交通模型和现状的简单指标统计，没有形成分析方法体系，**分析程度非常浅**
- 交通模型体系在大数据环境下的理论创新，以及空间分析、机器学习等技术手段在交通规划行业的广泛应用



图：专业的大数据分析手段在行业内的推广普及

52



提升数据分析的深度

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 现有数据应用主要在传统交通模型和现状的简单指标统计，没有形成分析方法体系，**分析程度非常低下**
- 交通模型体系在大数据环境下的理论创新，以及空间分析、机器学习等技术手段在交通规划行业的广泛应用
- 交通规划业务人员需要提升自身的技术水平，才能够提出具有针对性和可行性的需求

49

52



提升数据分析的深度

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 现有数据应用主要在传统交通模型和现状的简单指标统计，没有形成分析方法体系，**分析程度非常低下**
- 交通模型体系在大数据环境下的理论创新，以及空间分析、机器学习等技术手段在交通规划行业的广泛应用
- 交通规划业务人员需要提升自身的技术水平，才能够提出具有针对性和可行性的需求

49

52



- 数据可以实现规划编制过程中政府、规划师、市民之间的动态互动，实现规划调研、方案制定、规划评估、规划实施等全过程的公众参与
 - 利用政府开放数据，让公众参与规划过程，可以更加精细化地研究人的行为需求与城市空间的内在关系，体现“以人为本”的规划设计



图：武汉市众规平台

“自下而上”的规划决策

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

50

- 数据可以实现规划编制过程中政府、规划师、市民之间的动态互动，实现规划调研、方案制定、规划评估、规划实施等全过程的公众参与
- 利用政府开放数据，让公众参与规划过程，可以更加精细化地研究人的行为需求与城市空间的内在关系，体现“以人为本”的规划设计



图：微软提出的基于大数据平台建设公共参与系统架构

52



国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

● 智慧城市建设已经成为国家战略

- ① 《中共中央关于制定国民经济和社会发展第十三个五年规划的建议》
- ② 国务院关于印发促进大数据发展行动纲要的通知
- ③ 《国家新型城镇化规划（2014—2020年）》



国家智慧城市发展战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

● 智慧城市建设已经成为国家战略

- ① 《中共中央关于制定国民经济和社会发展第十三个五年规划的建议》
- ② 国务院关于印发促进大数据发展行动纲要的通知
- ③ 《国家新型城镇化规划（2014—2020年）》

智慧城市建设目标

更加智慧地推进新型城镇化建设，实现城乡一体化发展；产业升级、社会治理、生态环境修复、民生改善和基础设施功能提升，是智慧城市建设的方向。

51

52



国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

52

- 越来越多的非传统规划企业借助国家智慧城市建設的契机介入规划行业
- 越来越多的数据通过大数据技术可以被连通和使用，但同时也对交通规划行业未来的发展提出了新的挑战

52

国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代的
交通规划

数据“菜谱”

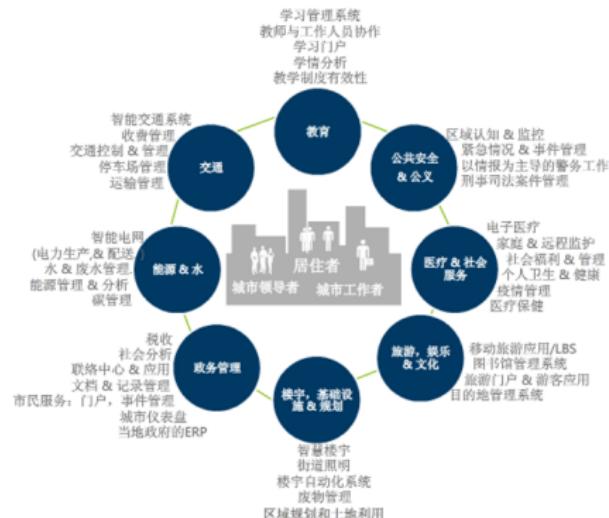
数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

52

- 越来越多的非传统规划企业借助国家智慧城市建設的契机介入规划行业
- 越来越多的数据通过大数据技术可以被连通和使用，但同时也对交通规划行业未来的发展提出了新的挑战



图：微软公司的智慧城市建設解决方案

52



国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

52

- 越来越多的非传统规划企业借助国家智慧城市建設的契机介入规划行业
- 越来越多的数据通过大数据技术可以被连通和使用，但同时也对交通规划行业未来的发展提出了新的挑战

智慧城市整体框架

智慧城市需要打造一个统一平台，设立城市数据中心，构建三张基础网络，通过分层建设，达到平台能力及应用的可成长、可扩充，创造面向未来的智慧城市系统框架。



图：华为公司的智慧城市建設解决方案

52

国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

52

- 越来越多的非传统规划企业借助国家智慧城市建設的契机介入规划行业
- 越来越多的数据通过大数据技术可以被连通和使用，但同时也对交通规划行业未来的发展提出了新的挑战



图：阿里巴巴参与雄安新区规划建设

52



国家智慧城市建設战略的基础

交通规划行业中
的数据应用现状
及思考

大数据时代下的
交通规划

数据“菜谱”

数据分析的“武
器库”

数据在业务中的
应用案例

再认识与展望

- 越来越多的非传统规划企业借助国家智慧城市建設的契机介入规划行业
- 越来越多的数据通过大数据技术可以被连通和使用，但同时也对交通规划行业未来的发展提出了新的挑战

汇报结束
谢谢!

