



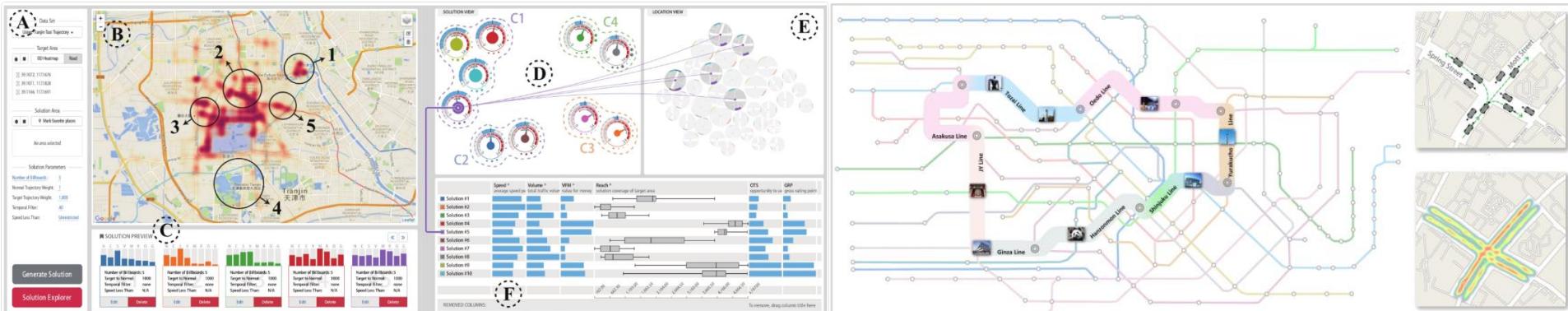
城市大数据的可视分析

巫英才

浙江大学CAD&CG国家重点实验室

ycwu@zju.edu.cn

<http://www.ycwy.org>





大纲

- 城市大数据的可视化
- 城市大数据可视分析应用案例

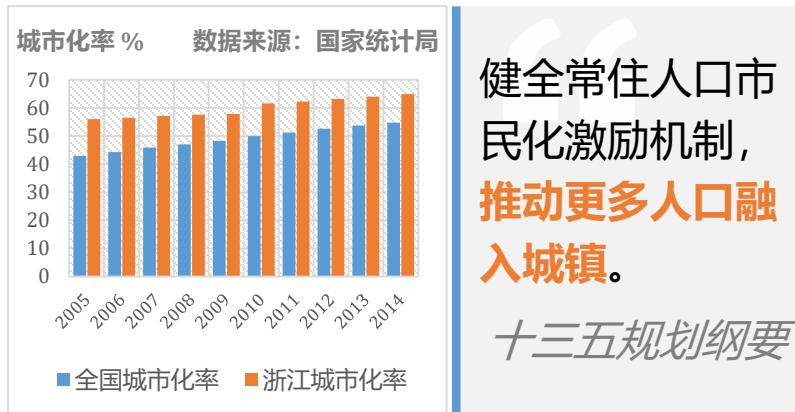


城市大数据的可视化

城市化面临的问题，以及可视化是如何得到应用的

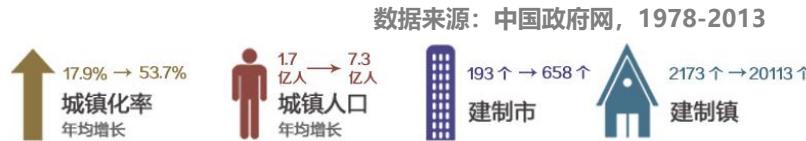
城市问题是国家重大社会问题

城市化是中国重大战略



健全常住人口市民化激励机制，
推动更多人口融入城镇。

十三五规划纲要



更**充裕**的工作机会



更**丰富的**物质生活

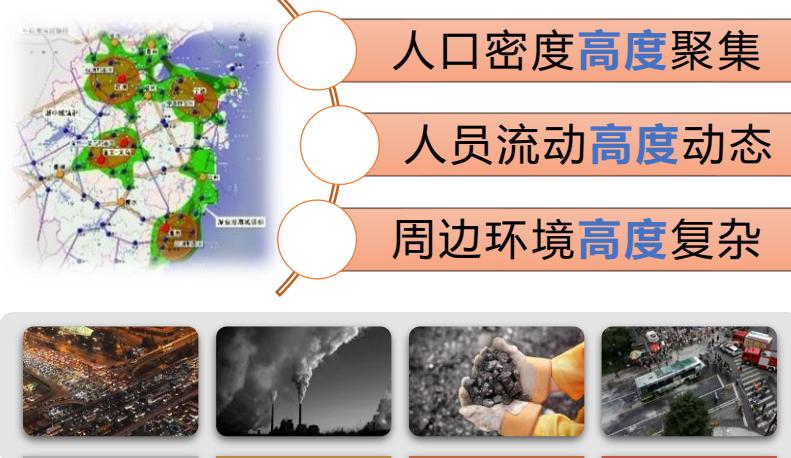


更**便利**的城市交通



更**安心**的公共服务

各种城市问题浮现



近三年来我国因**管理与服务**不当引发的重大坍塌、踩踏等事故造成直接经济损失**5000多亿元**。

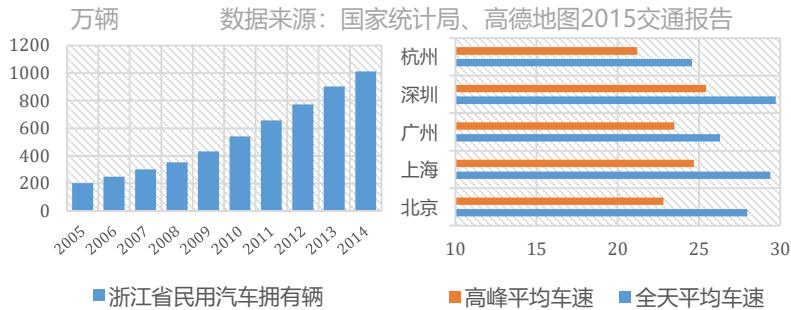
数据来源：国家安全生产监督管理总局

城市问题成为制约城市健康和可持续发展的难题

浙江重大需求

杭州市的常住和暂住人口
总计 **2000 余万**

交通拥堵



公共安全



2022年杭州亚运会

金融、安全生产、社会治安、网络等领域**潜在风险隐患较多**，维护国家安全和保持社会大局稳定**面临新的挑战**。

浙江省十三五规划

保障城市的安全与舒适是一个重大的社会需求

重大社会意义

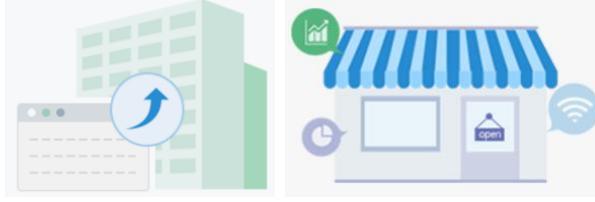
智慧城市：基于城市大数据，利用信息技术精准地分析以及高效地解决城市的问题。

建设城市信息基础



分析相关城市数据

产生明显社会经济效益



大数据地产分析
把握地产价值趋势
大数据店铺分析
洞悉顾客提升服务

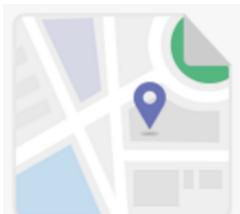
形成新的经济增长点



提升传统优势产业升级



大数据客群分析
洞晓客流分布迁徙



大数据选址分析
空间洞察科学选址

牵引现代服务业的发展

构建智慧城市具有重大的社会意义



借助两化融合，构建智慧城市

智慧城市：基于城市大数据，利用信息技术精准地分析以及高效地解决城市的问题。

浙江20个智慧城市建设项目

- 完善城市信息化基础设施
- 建立健全各省级数据中心
- 每日积累海量城市数据



数据中心的基础设施建设

多源异构

海量高维

动态模糊

全局稀疏

局部冗余

时空断裂

“理想”的解决方案

帮助我分析最近雾霾的成因



用户引导

发生雾霾的原因是.....，我认为应该采取如下措施.....



计算机引导

强人工智能短期内几乎不可能实现



关键问题 (1/3)

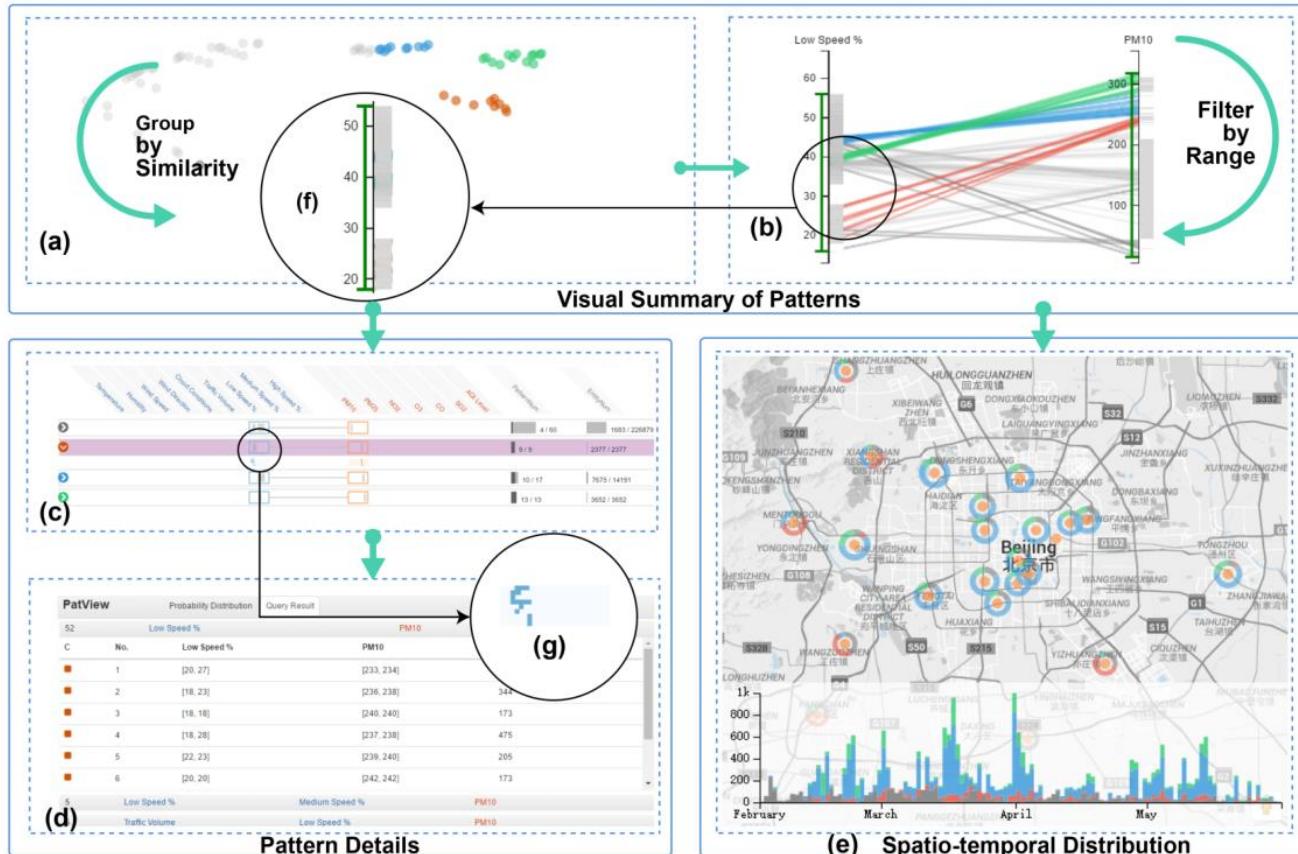
- 如何从大规模复杂的城市数据集中得出结论
- “这么多杂乱无章的数据该从何着手？”

Data Source	Property	Value Range
Air Quality	PM 2.5, PM 10, O ₃ (ug/m ³)	[0, 500]
	NO ₂ (ug/m ³)	[0, 300]
	CO (ug/m ³)	[0, 20]
	SO ₂ (ug/m ³)	[0, 50]
	Level	6 levels
Weather	Temperature °C	[-20, 40]
	Humidity	[0, 100]
	Wind Speed (m/s)	[0, 300]
	Wind Direction	[1, 24]
	Cloud Conditions	14 conditions
Traffic	Total Cars	[9, 200]
	Low Speed (0 ~ 20 km/h) %	[6, 80]
	Medium Speed (20 ~ 50 km/h) %	[17, 74]
	High Speed (above 50 km/h) %	[0, 53]

如何描述复杂的多属性关联

关键问题 (1/3)

- 如何从大规模复杂的城市数据集中得出结论
- “这么多杂乱无章的数据该从何着手？”



关键问题 (2/3)

- 如何利用专家的领域知识发现新的见解
- “这不是我想要的答案。我想要..... 再处理一次。”



这里有太多轨迹了，
我想抽样一下。

这个假设的前提太牵
强了，我想修改成....

这些数据点和分析没
有关系，能否删除？

结果不太对。能否尝
试一下其他方法？

这个模式看起来很可
疑，具体分析一下。

关键问题 (3/3)

- 如何传递并解释从数据中得到的见解
- “给我一份好看的报告哄走我的老板。”



如何用可视化解决城市问题

三个目标

- 多层次的可视化
 - Overview first, zoom and filter, then details-on-demand
- 支持混合引导探索的富交互行为
- 叙述性地总结数据中挖掘得到的见解

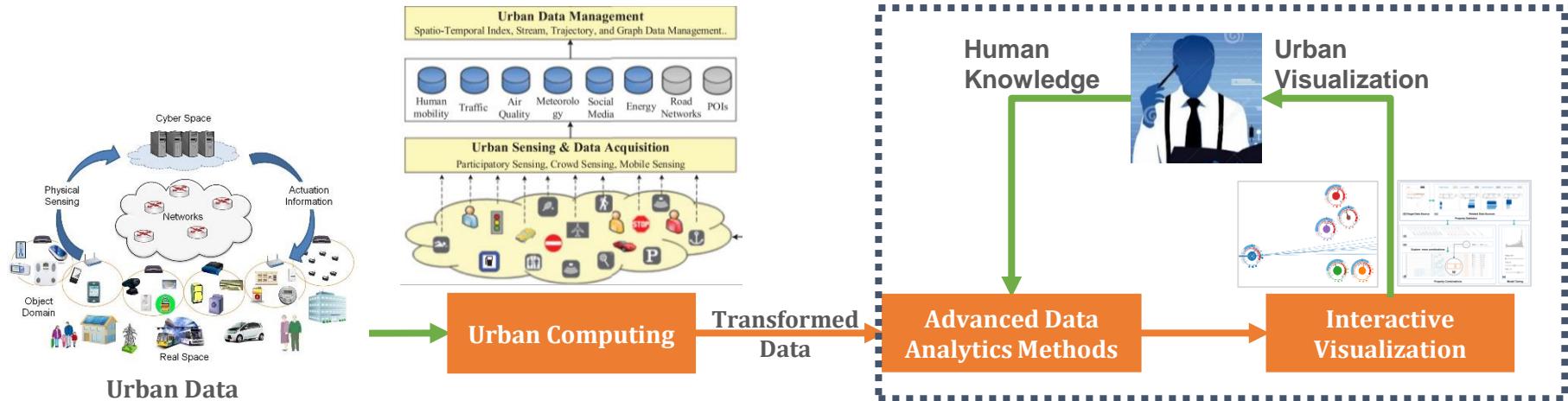


城市数据的可视分析



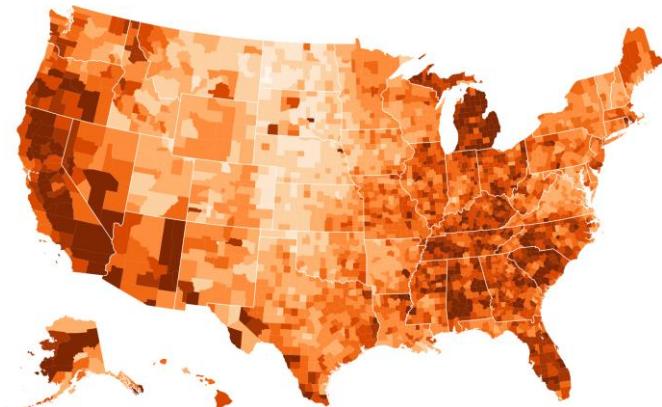
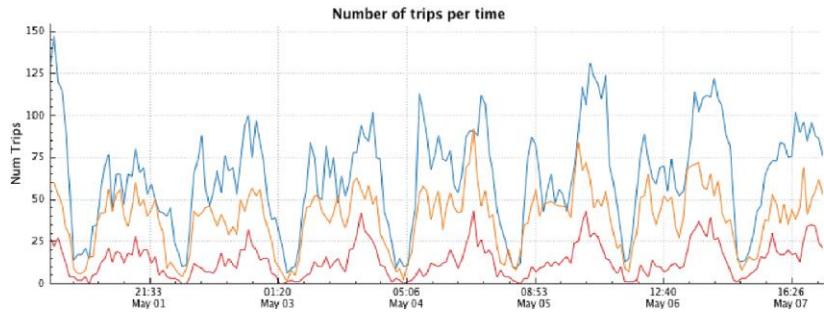
城市可视分析 结合先进的数据分析方法和交互式可视化

- 易于使用、性能高效的多层次可视化
- 通过交互式可视化多角度分析数据
- 叙述性地呈现可视化结果



如何可视化城市数据

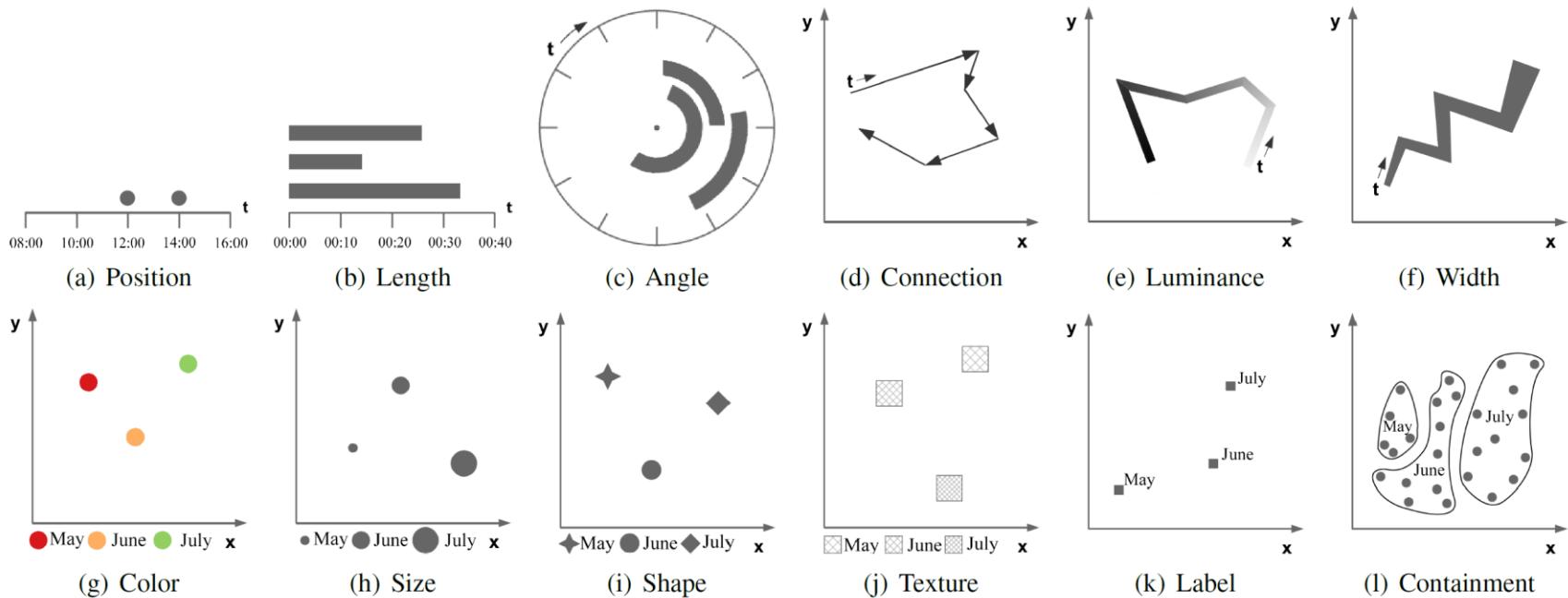
- 时间数据的可视化
- 空间数据的可视化
- 时空数据的可视化



时间数据的可视化



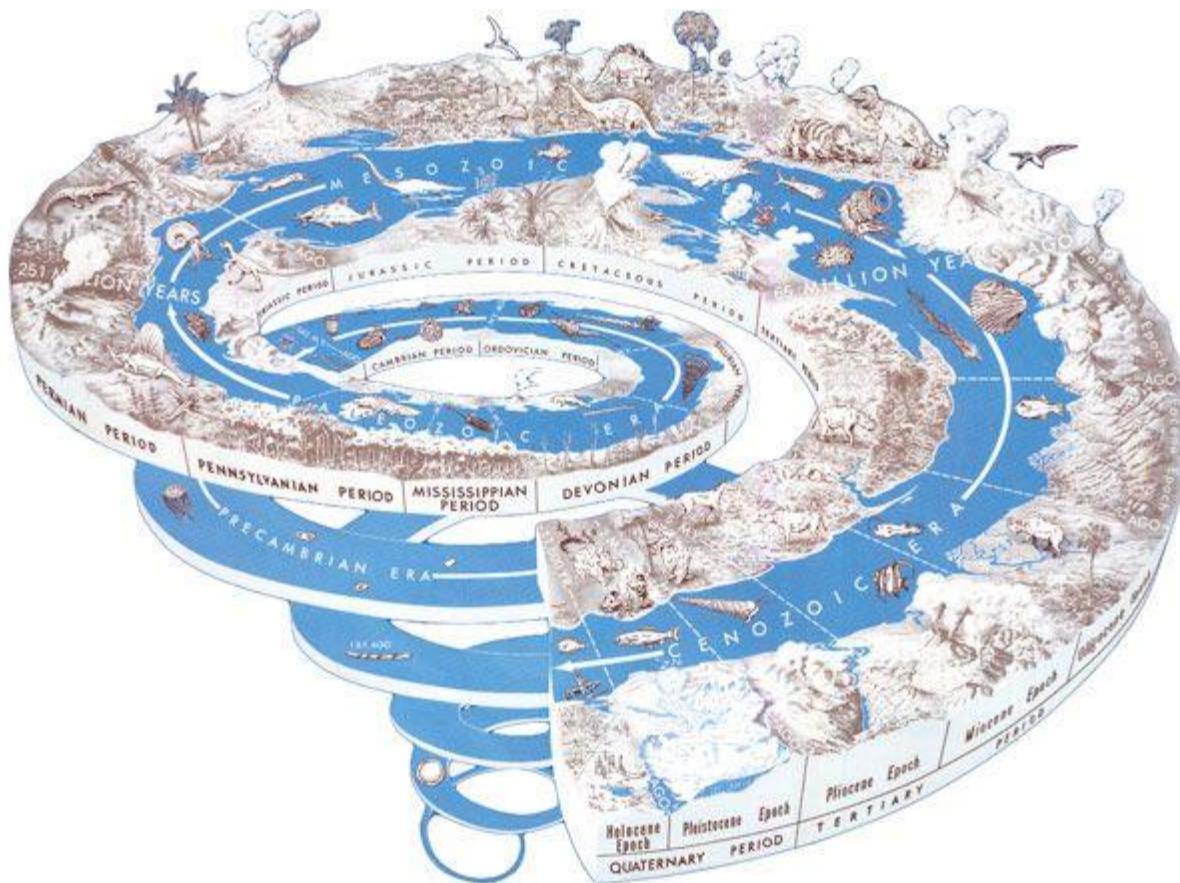
许多可视化技巧可被应用于可视化时间数据



时间数据的可视化

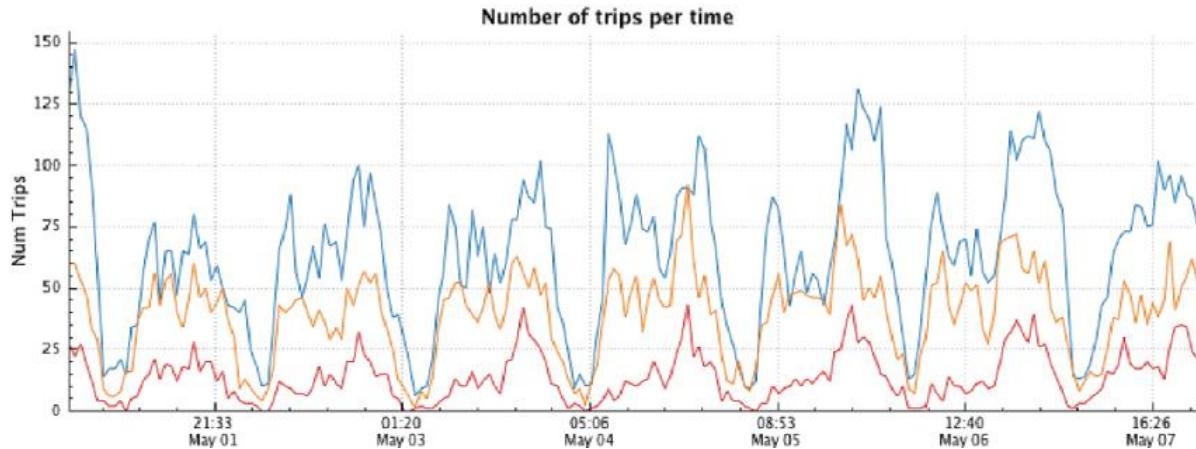
时间数据的两个重要特征

- 线性
- 循环



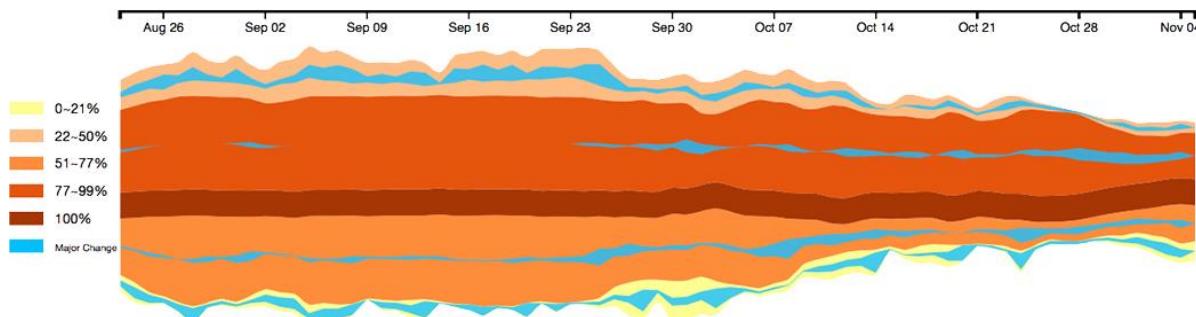
可视化线性的时间

使用图表准确地表示绝对的时间



简单的折线图可视化展示不同区域出租车载客数随时间的变化。

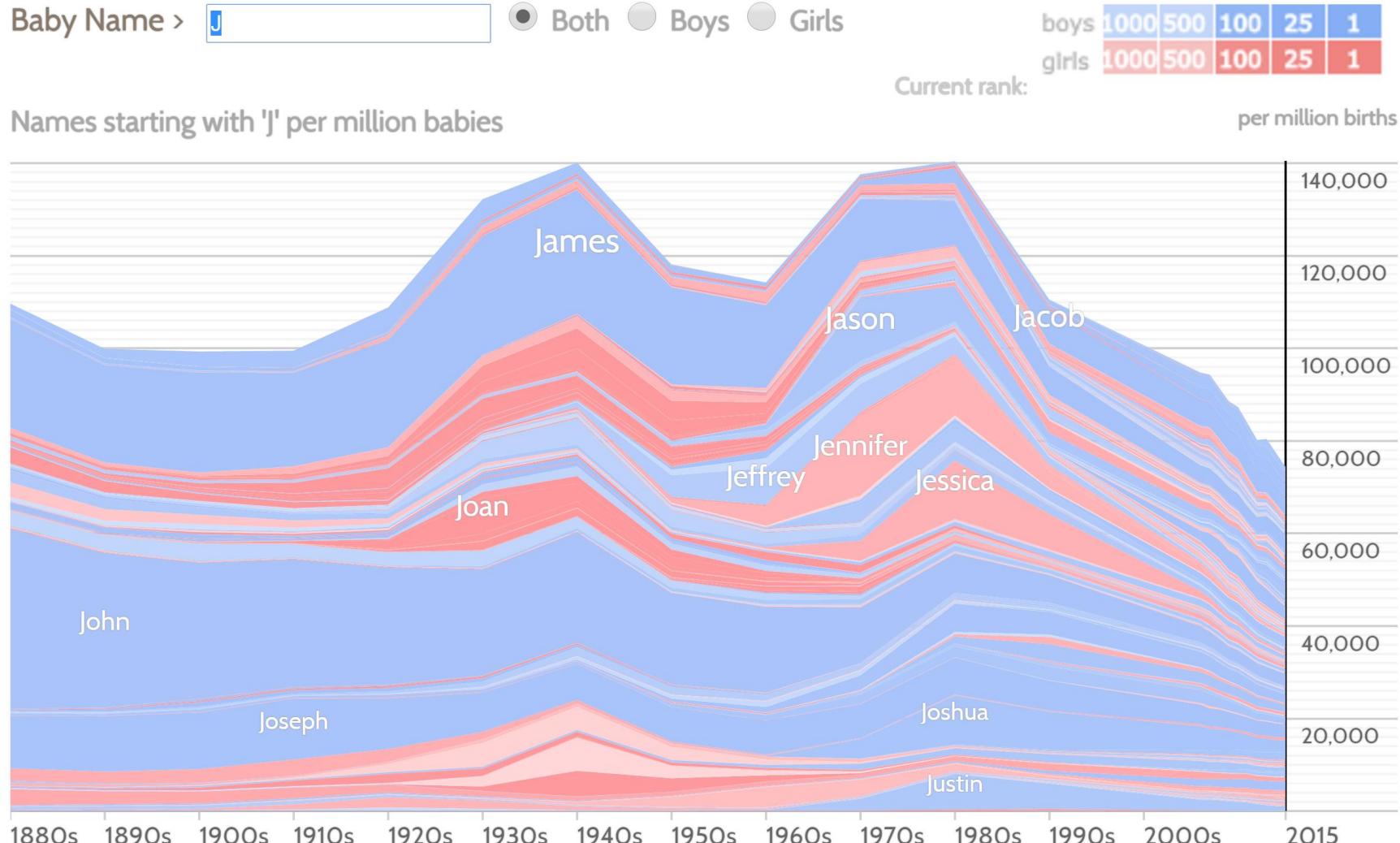
N. Ferreira, et. al, 2013



一个展示南极洲臭氧层空洞覆盖面积变化的堆叠图可视化。

W. Wu, et. al, 2014

可视化线性的时间 – Name Voyager





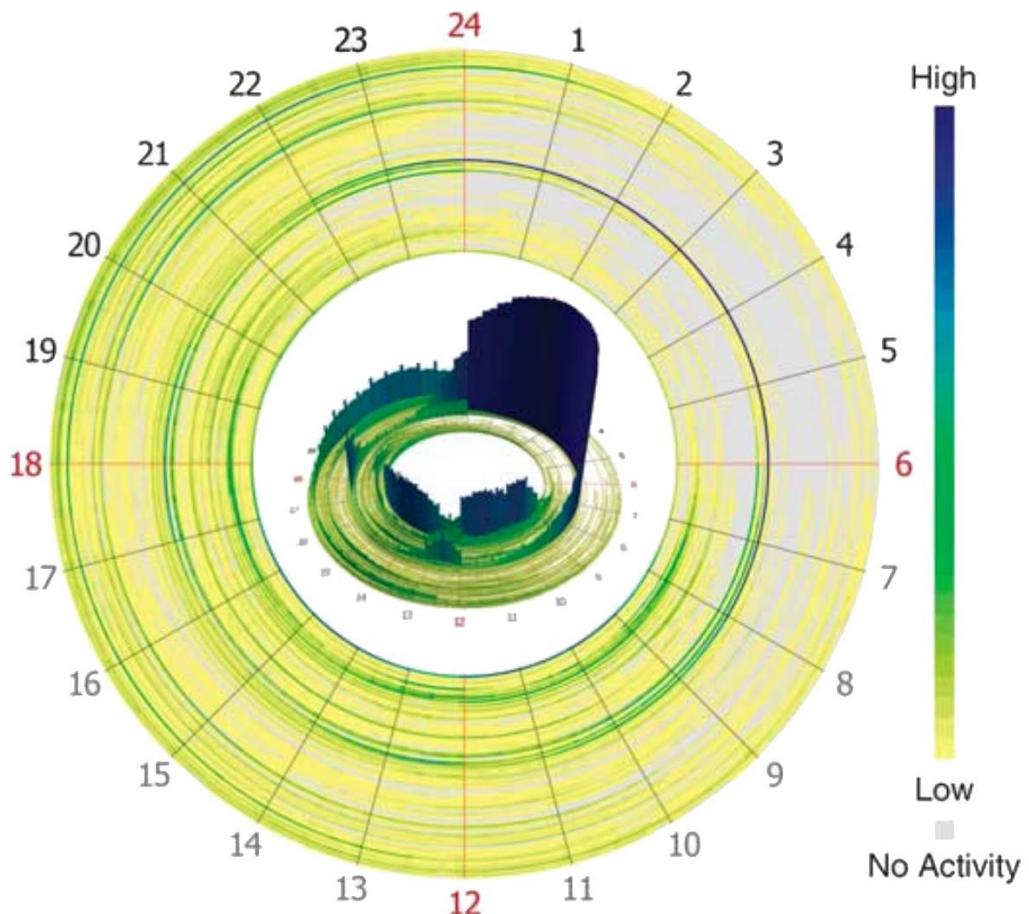
可视化线性的相对时间关系

颜色和连接可以用于相对地可视化线性时间

用颜色和连接可视化1884-2012年间全球气温的变化。B. Bach, et. al, 2016

可视化循环的时间

圆形时间轴经常用于强调时间的循环属性



可视化地利用圆形时间轴来展示96种人类活动在一天24小时中频率的变化。每个圆环代表一种活动，颜色编码着活动的频繁程度。

J. Zhao, et. al, 2008

可视化循环的时间 - KronoMiner



KronoMiner

Using Multi-Foci Navigation for the
Visual Exploration of Time-Series Data

Jian Zhao¹
Fanny Chevalier²
Ravin Balakrishnan¹

¹University of Toronto

²OCAD University

空间数据的可视化

- 基于点的可视化
- 基于区域的可视化
- 基于线条的可视化



7 am - 8 am



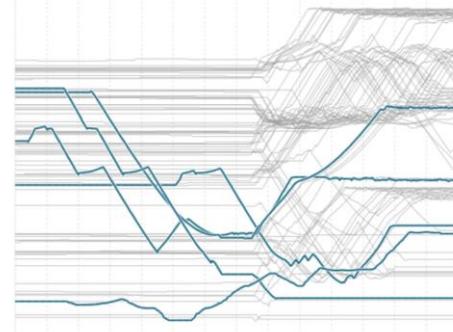
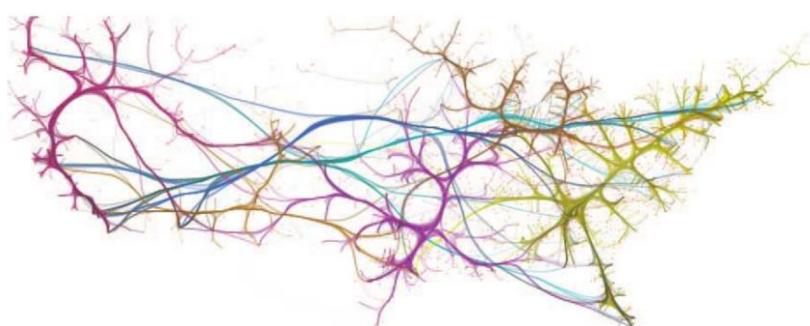
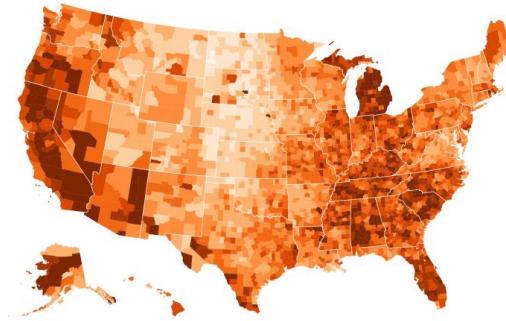
8 am - 9 am



9 am - 10 am



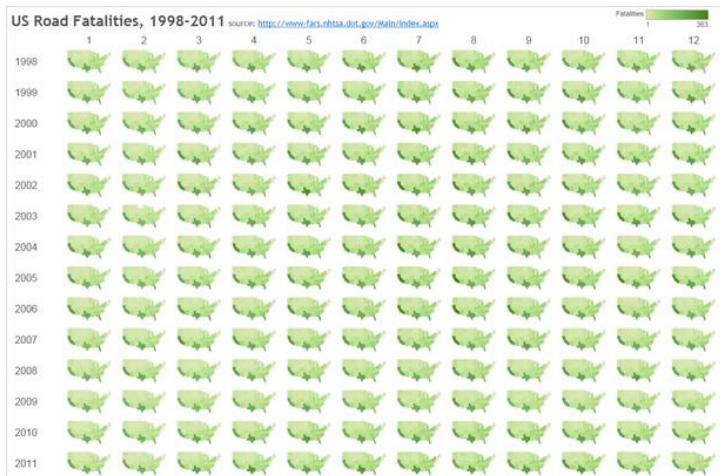
10 am - 11 am



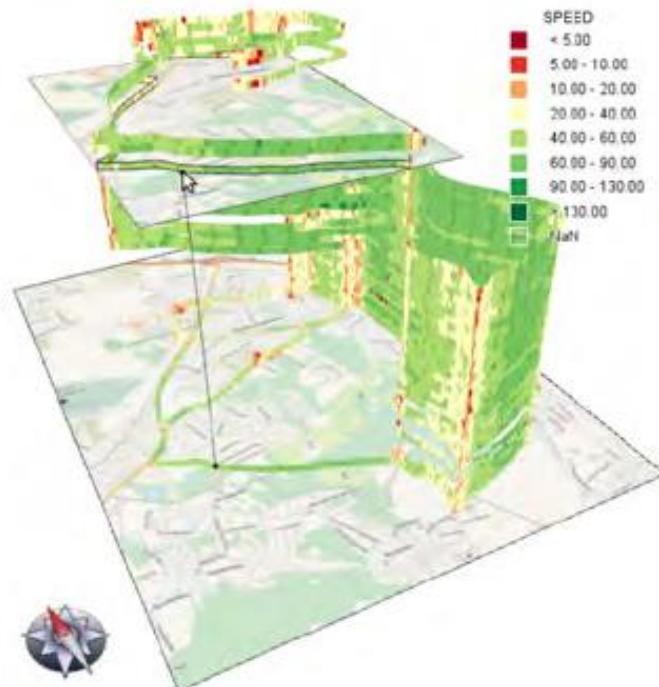
时空数据的可视化

同时考虑时间和空间属性的可视化方法

- Small multiples
- Space-time Cube



美国不同年份因道路事故
死亡人数的分布情况



立体的轨迹可视化方法Space-time
Cube，颜色编码车流速度。

C. Tominski, et. al, 2012



如何获取数据 – 国外数据



DATA TOPICS IMPACT APPLICATIONS DEVELOPERS CONTACT

The home of the U.S. Government's open data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and [more](#).

GET STARTED

SEARCH OVER 192,728 DATASETS

Monthly House Price Indexes



BROWSE TOPICS



Agriculture



Business



Climate



Consumer



Ecosystems



Education



Energy



Finance



Health



Local
Government



Manufacturing



Ocean



Public Safety



Science &
Research

如何获取数据 – 中国政府公开数据



中国政府公开信息整合服务平台
Chinese Government Public Information Online

收藏我们 帮助

高级检索

专题资源 →

最新资源

更多>>

- [宁夏分站]贺兰县超额完成2016年农田水利建设总任务 2016-11-17
- [宁夏分站]贺兰县免费为4000名老年人办理意外伤害保险 2016-11-17
- [宁夏分站]吴忠市加速推进东南部引提调水工程建设 2016-11-17
- [宁夏分站]青铜峡市农机购置补贴资金突破一千万元 2016-11-17
- [宁夏分站]领导干部生态环境损害责任追究实施细则(试行)印发 2016-11-17
- [宁夏分站]宁夏政府工作报告坚持“开门写” 2016-11-17
- [宁夏分站]平罗逾七成财政支出用于改善民生 2016-11-16
- [宁夏分站]彭阳入选全国休闲农业 和乡村旅游示范县 2016-11-16
- [宁夏分站]西部担保188亿融资担保助企渡寒冬 2016-11-16
- [宁夏分站]我区将从国家铁路网末梢一跃成重要枢纽 2016-11-16
- [宁夏分站]我区全面推进安全生产责任险工作 2016-11-16

分站导航 →

政府公报

主题分类 体裁分类 公文文种 更多>>

主题分类	体裁分类	公文文种
财政、金融、审计	城乡建设、环境保护	对外事务
港澳台侨工作	工业、交通	公安、安全、司法
国防	国民经济管理、国有资产监管	国土资源、能源
科技、教育	劳动、人事、监察	民政、扶贫、救灾
民族、宗教	农业、林业、水利	人口与计划生育、妇女儿童工
商贸、海关、旅游	市场监管、安全生产监管	卫生、体育
文化、广电、新闻出版	综合政务	组织机构

政府机构

友情链接： 中华人民共和国中央人民政府 | 国家图书馆 | 数字图书馆推广工程 | 中国政府公开信息整合服务项目

如何获取数据 – 中国国家数据

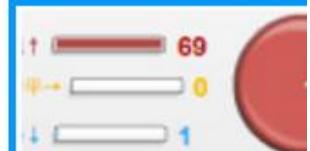


可视化图表

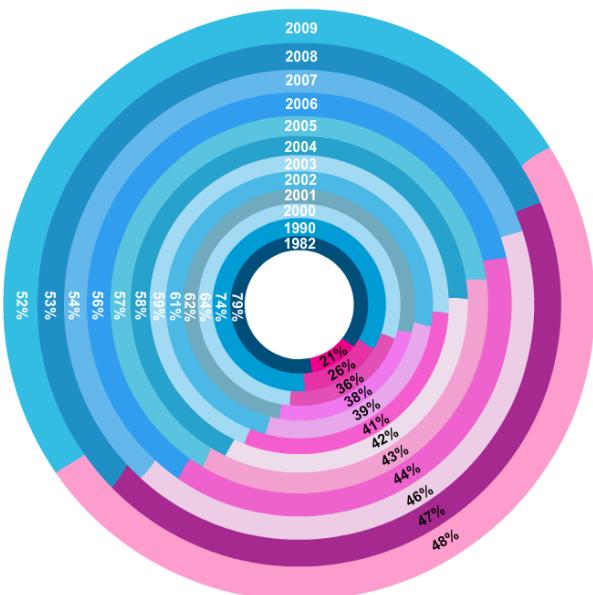
可视化统计图秉
数字，通过简单的图
图，您可更清楚了解



大中城市住宅销售



大中城市房价



▶ || 1982 1990 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012



中国制造业采购



房地产投资

如何获取数据 – 中国气象数据



欢迎进入中国气象数据网!

[登录](#) [注册](#) [我的数据筐](#) [联系我们](#)



国家气象信息中心
中国气象局气象数据



首页

数据服务

实况资料

气候背景

在线展示

用户支持

关于我们

search



重要资料推荐

地面气象资料

[中国地面逐小时资料](#) | [全球地面定时资料](#)

高空气象资料

[中国高空定时资料](#) | [全球高空定时资料](#)

卫星探测资料

[中国静止气象卫星](#) | [中国极轨气象卫星](#)

天气雷达探测资料

[雷达基数据](#) | [基本反射率](#)

数值预报模式产品

[T639模式产品](#) | [GRAPES模式产品](#)

地面观测

高空观测

卫星云图

雷达拼图

模式产品

国家气象信息中心
中国气象局气象数据

全国逐小时降水量

11月19日00时-01时 BJT



用户支持



会员注册流程



如何成为会员



专业词汇解释



如何获取数据



我的收藏



常见问题



电话: 86-10-68407499



邮箱: data@cma.gov.cn





如何获取数据 – 地方政府的数据

上海市政府
数据服务网

首页 数据 应用 接口 移动应用 地理信息 互动交流 用户登录

请输入数据/应用/移动应用名称关键词... 全站搜索 更多 ▾

数据 Dataset
通过多种分类，找到您想要的数据....

数据领域

您可以点击下方链接进入相应的分类

经济建设	资源环境	教育科技	道路交通	社会发展	公共安全
文化休闲	卫生健康	民生服务	机构团体	城市建设	信用服务

如何清洗数据 – Google Refine



Google refine 

如何清洗数据 – Trifacta



如何可视化 - Tableau



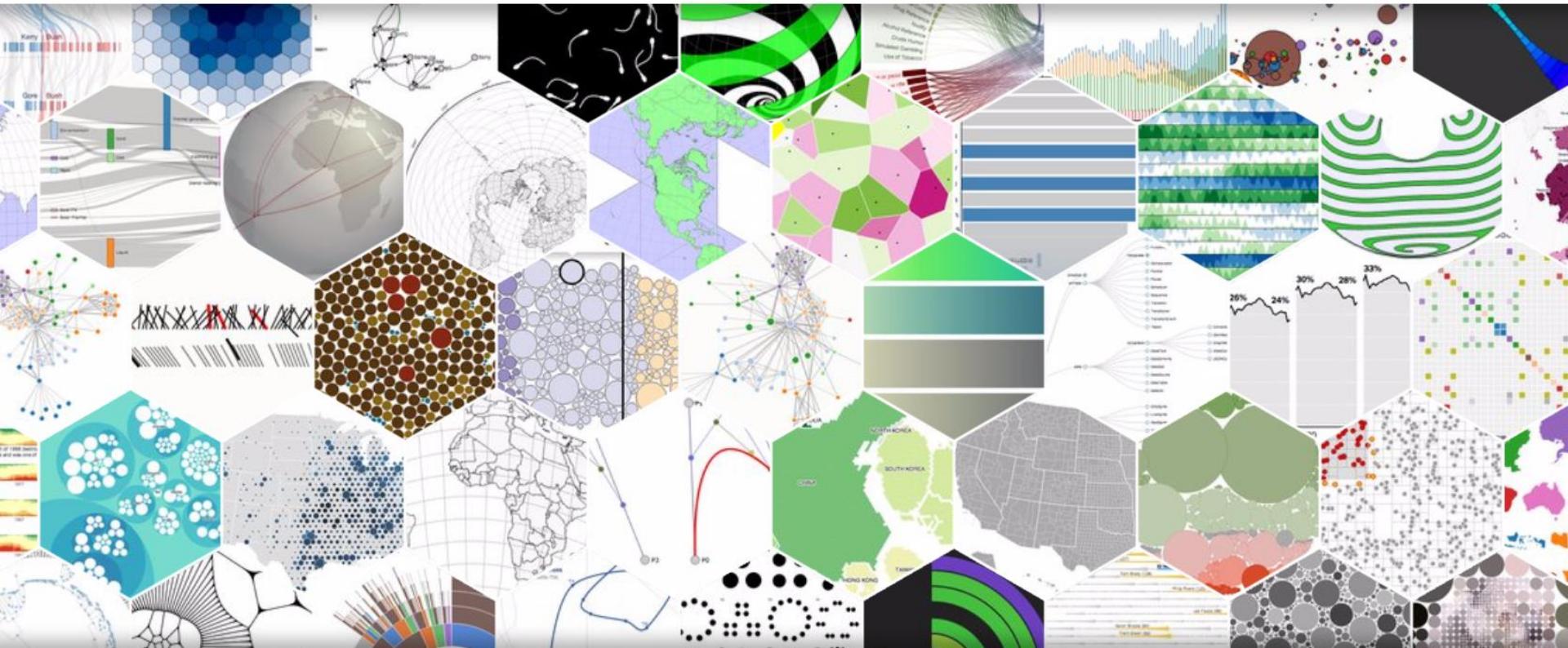


如何可视化 – Power BI

如何可视化 – D3



Data-Driven Documents

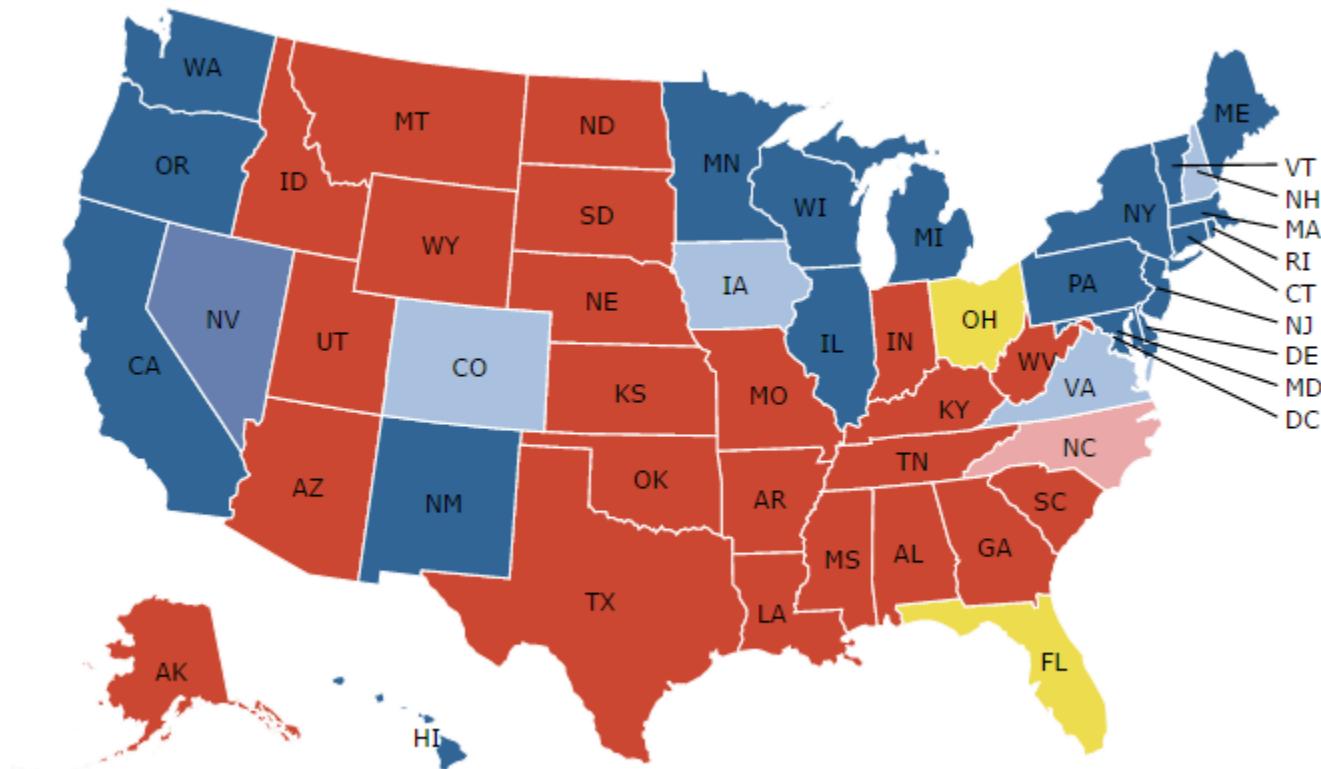


<https://d3js.org/>

如何可视化 – 基于D3的地理可视化库

DataMaps

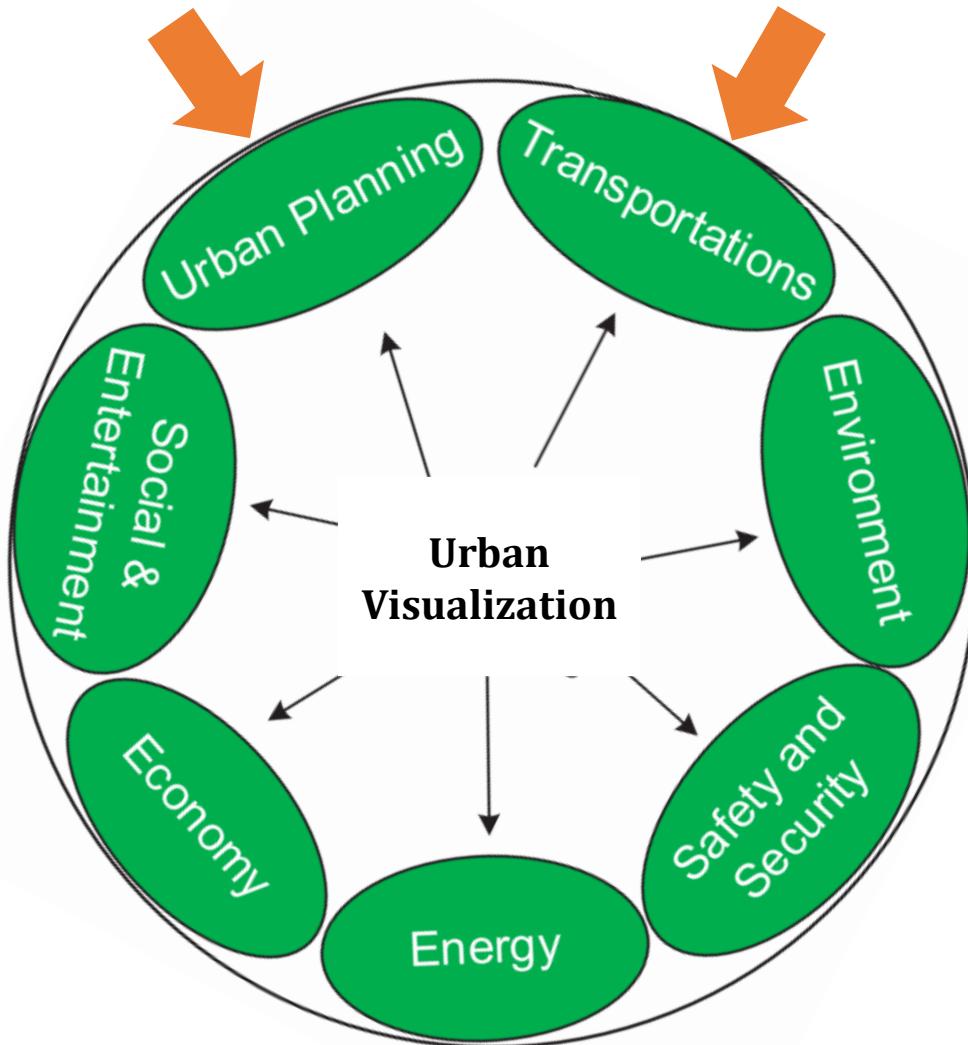
Customizable SVG map visualizations for the web in a single Javascript file using D3.js



<https://datamaps.github.io/>



城市数据可视化分析应用



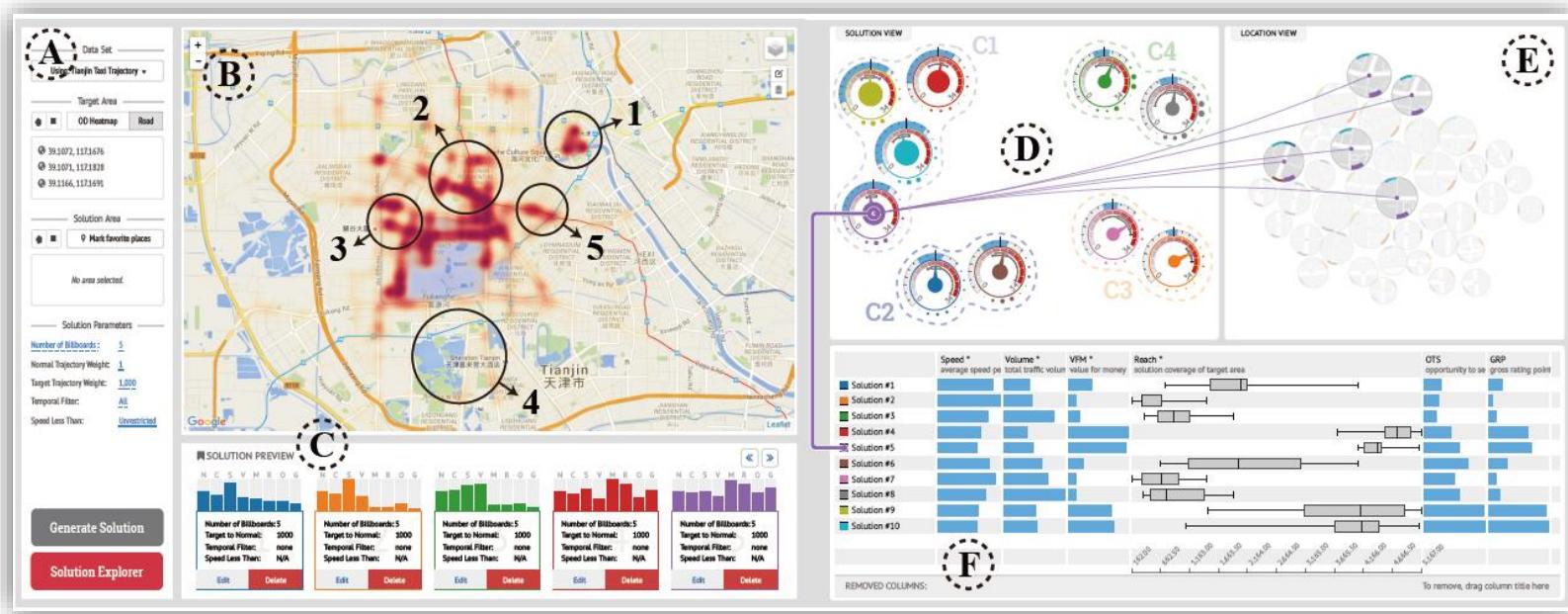


案例分析

基于大数据的广告牌位置选取问题

基于道路放大技术的交通流量分析

城市规划



SmartAdP: Visual Analytics of Large-scale Taxi Trajectories for Selecting Billboard Locations

户外大型广告牌

由于户外大型广告牌非常容易被看到，它成为了一种流行的营销手段。



沿路的户外大型广告牌

广告牌放置的地点



地点，
地点，
地点。



适当的放置地点



不适当的放置地点



影响广告牌放置地点的因素

我们应该考虑哪些因素呢？



影响广告牌放置地点的因素

交通流量

车流速度

轨迹OD

周边环境



影响广告牌放置地点的因素

交通流量

车流速度

轨迹OD

周边环境



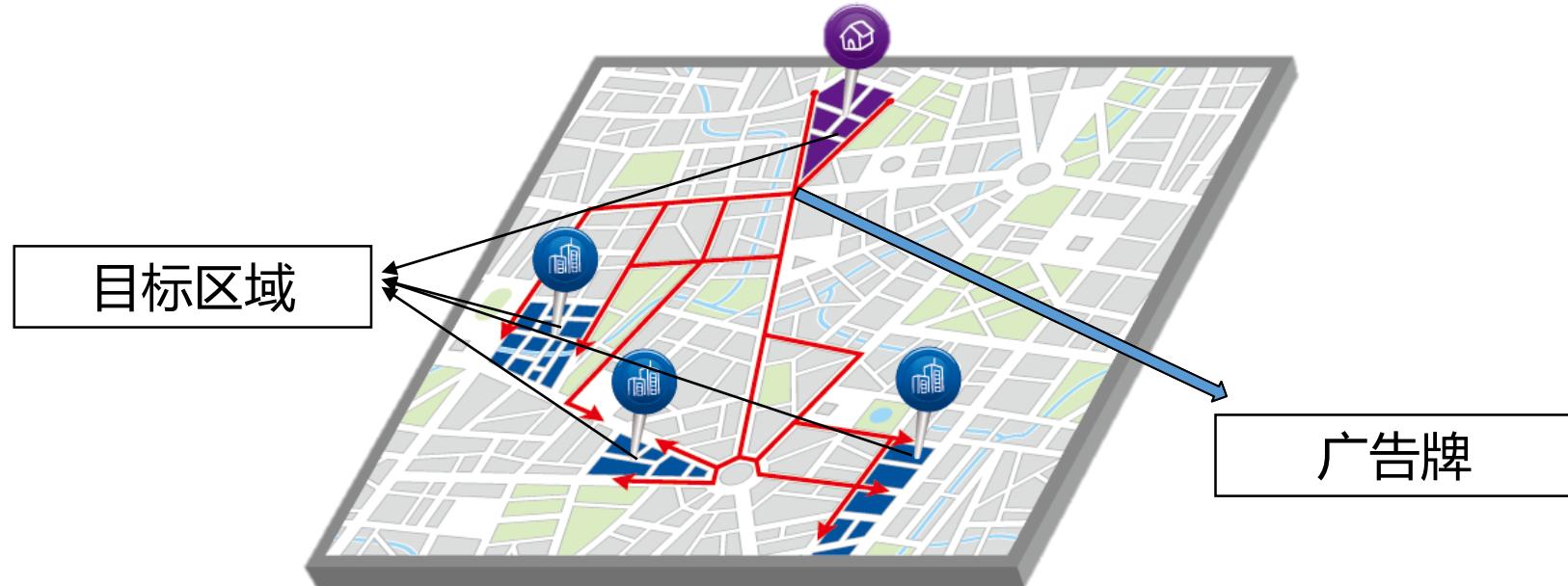
影响广告牌放置地点的因素

交通流量

车流速度

轨迹OD

周边环境



影响广告牌放置地点的因素

交通流量

车流速度

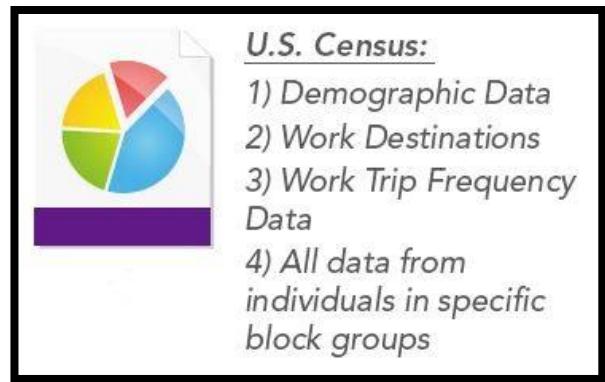
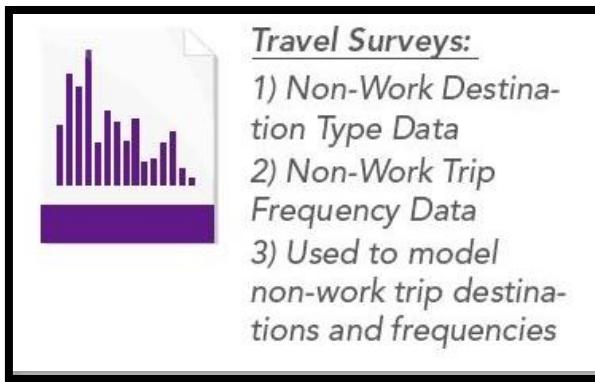
轨迹OD

周边环境



关键挑战 (1/3)

- 如何以较少的代价来获取详细的数据?
- 传统方法不仅昂贵而且耗时耗力



容易采集

揭示隐藏的城市
流量特征

关键挑战 (2/3)

- 如何快速生成一个好的广告牌放置方案
 - 庞大的解决方案空间
 - 存在难以建模的因素
 - 对于最优解的定义众口难调



计算能力

+



领域知识

=

性价比最优的
广告方案

关键挑战 (3/3)

- 如何可视化地比较多个放置方案并确定最优的一个
 - 哪种类型的广告策略是最好的...
 - 如果我多花上50,000元...





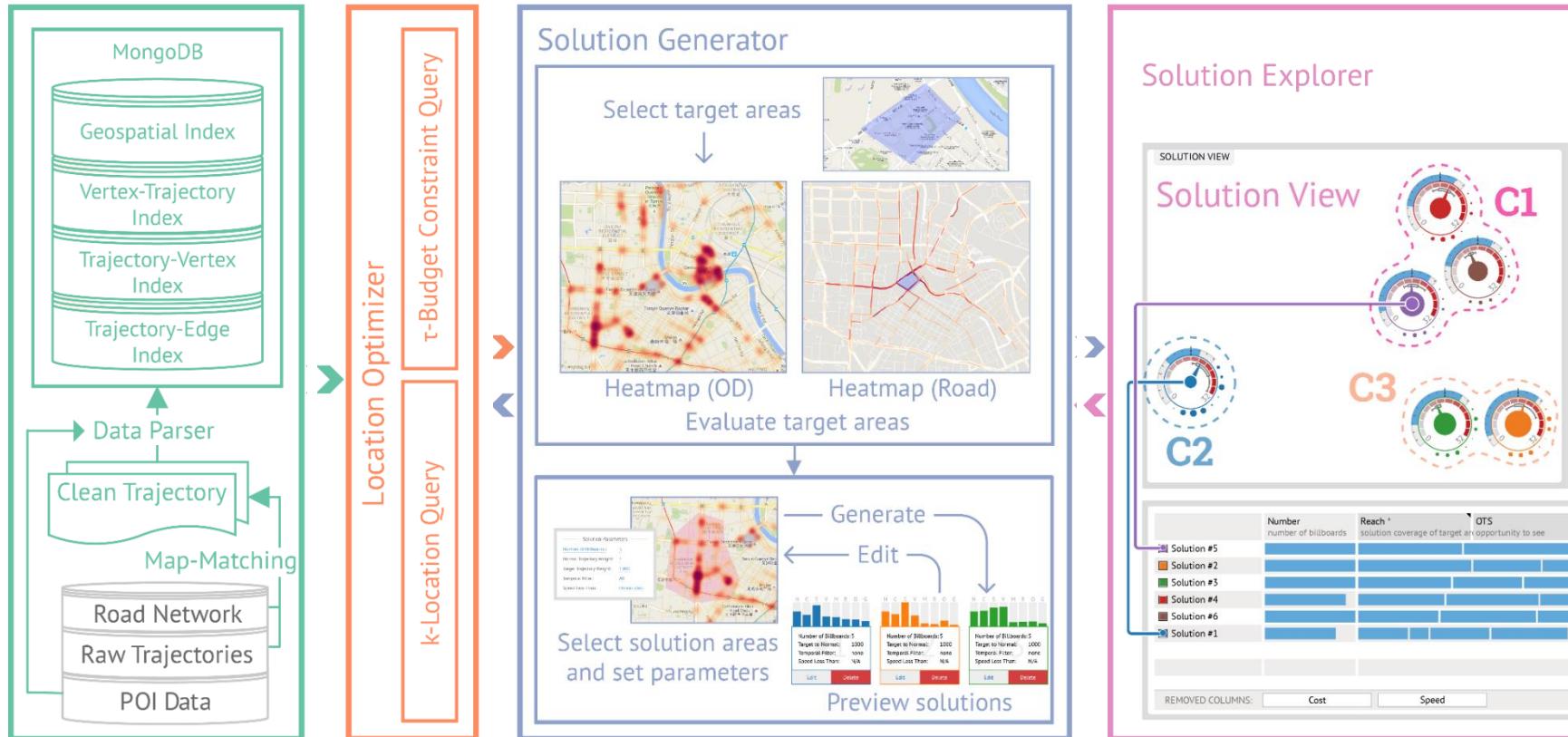
设计目标

- 从大规模出租车轨迹中**提取**出与广告牌放置位置相关的信息
- 一套综合计算能力和领域知识来**生成**放置方案的系统
- 一种准确且清晰易懂的可视表达用于帮助用户**比较**不同的放置方案

解决方案



结合先进数据挖掘算法和交互式可视化技术的一套
可视分析系统





分析任务

Location
Optimizer

Solution
Generator

Solution
Explorer

哪些区域对放置户外广告牌比较有利？

哪里是最佳广告牌放置地点？

一个广告牌放置地点/方案的性能如何？

方案之间的差异和相似性在哪里？

根据相似性方案可以被分成几组？

方案之间的排名是怎样的？哪个方案最好？

Solution Generator的交互



Solution Generator



数据结构的建立

为了从大规模出租车轨迹中提取数据并加速数据挖掘算法，我们建立了三套索引：

Trajectory

Edge (road segment)	e_{i1}	e_{i2}	\dots			
Tr_i						

Trajectory

Potential billboard location Vertex (road intersection)	v_{i1}	v_{i2}	\dots			
Tr_i						

Vertex
 v_i

Trajectory	Tr_{i1}	Tr_{i2}	\dots			
v_i						

数据挖掘的算法

- 问题: 寻找一组地点, 使得它们可以覆盖最多的轨迹

- 每条轨迹有一个权重

- 每个地点有一个耗费

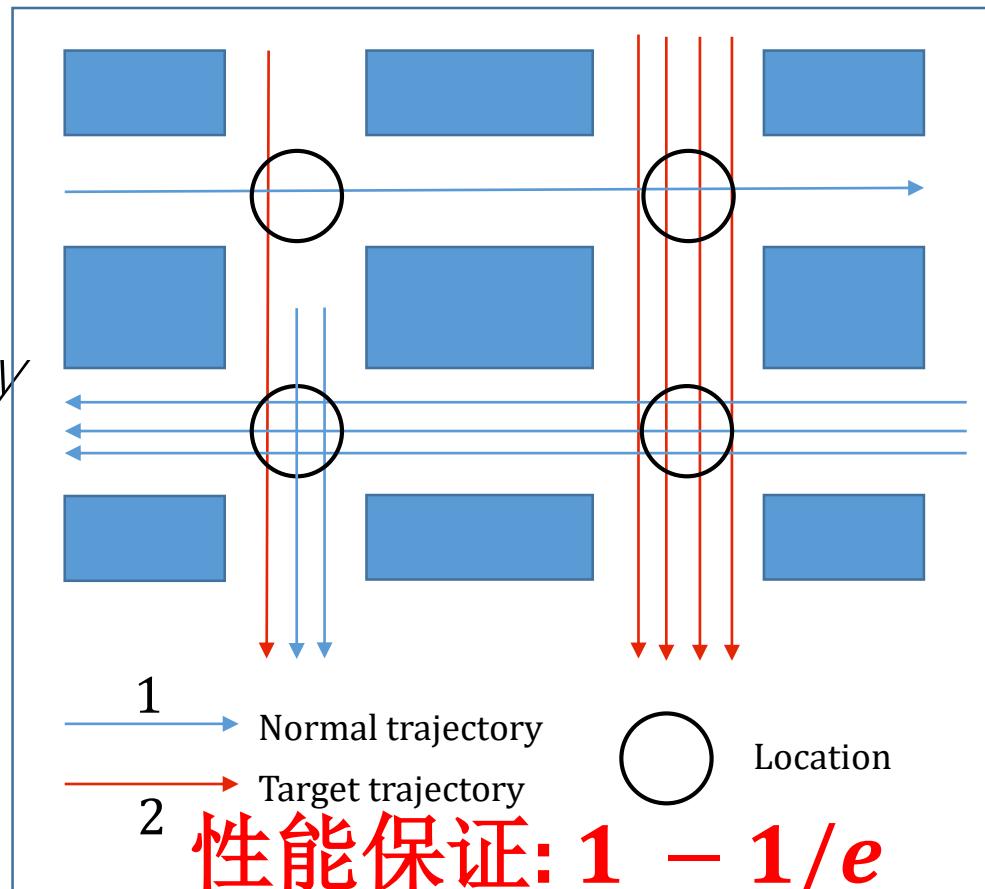
- *K-Location Query*

- 选择与更新

- *τ -Budget Constraint Query*

- 利用率:

- $$U(v) = \frac{\text{current coverage value}}{\text{cost}}$$





分析任务

Location
Optimizer

Solution
Generator

Solution
Explorer

哪些区域对放置户外广告牌比较有利？

哪里是最佳广告牌放置地点？

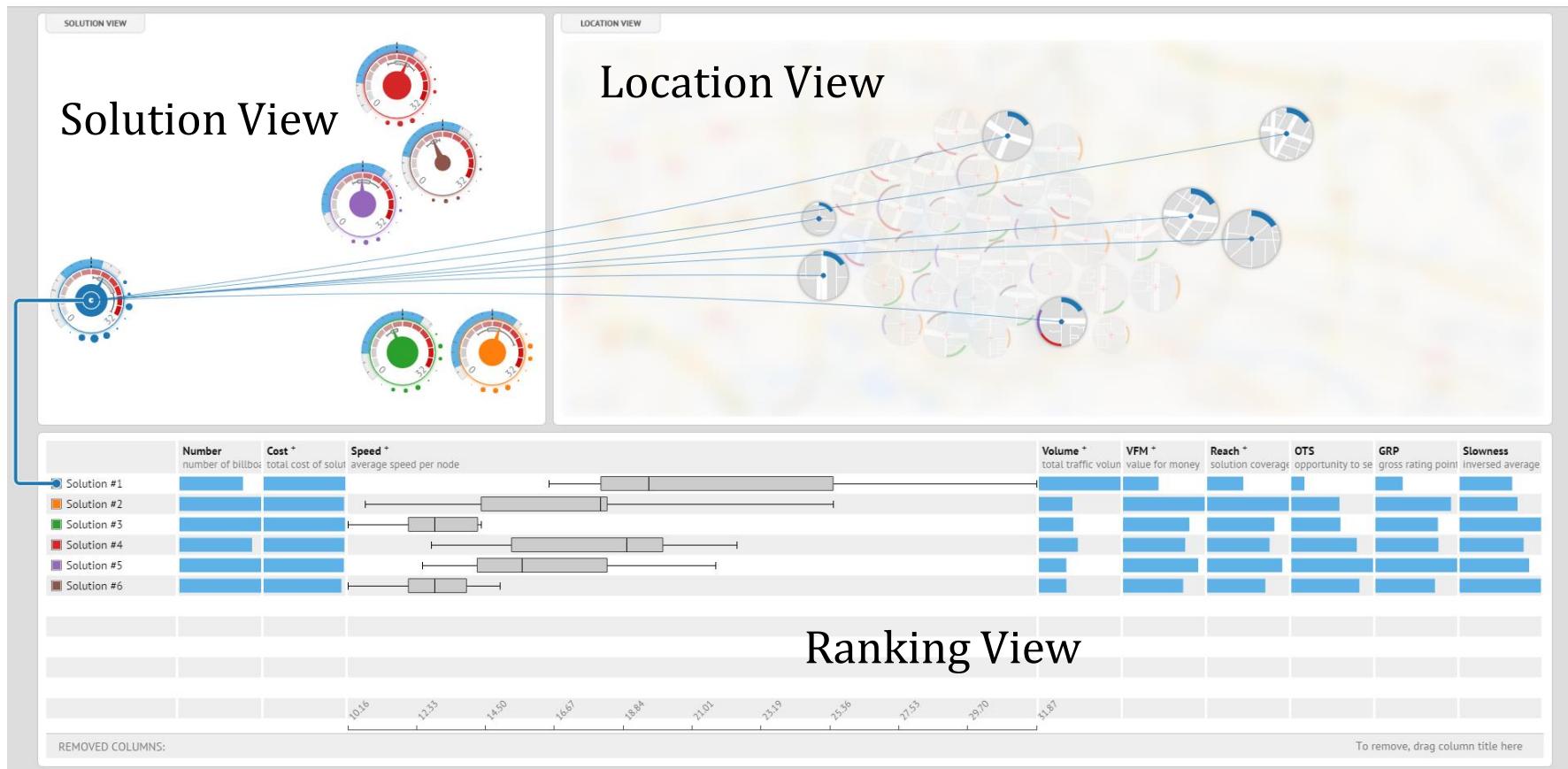
一个广告牌放置地点/方案的性能如何？

方案之间的差异和相似性在哪里？

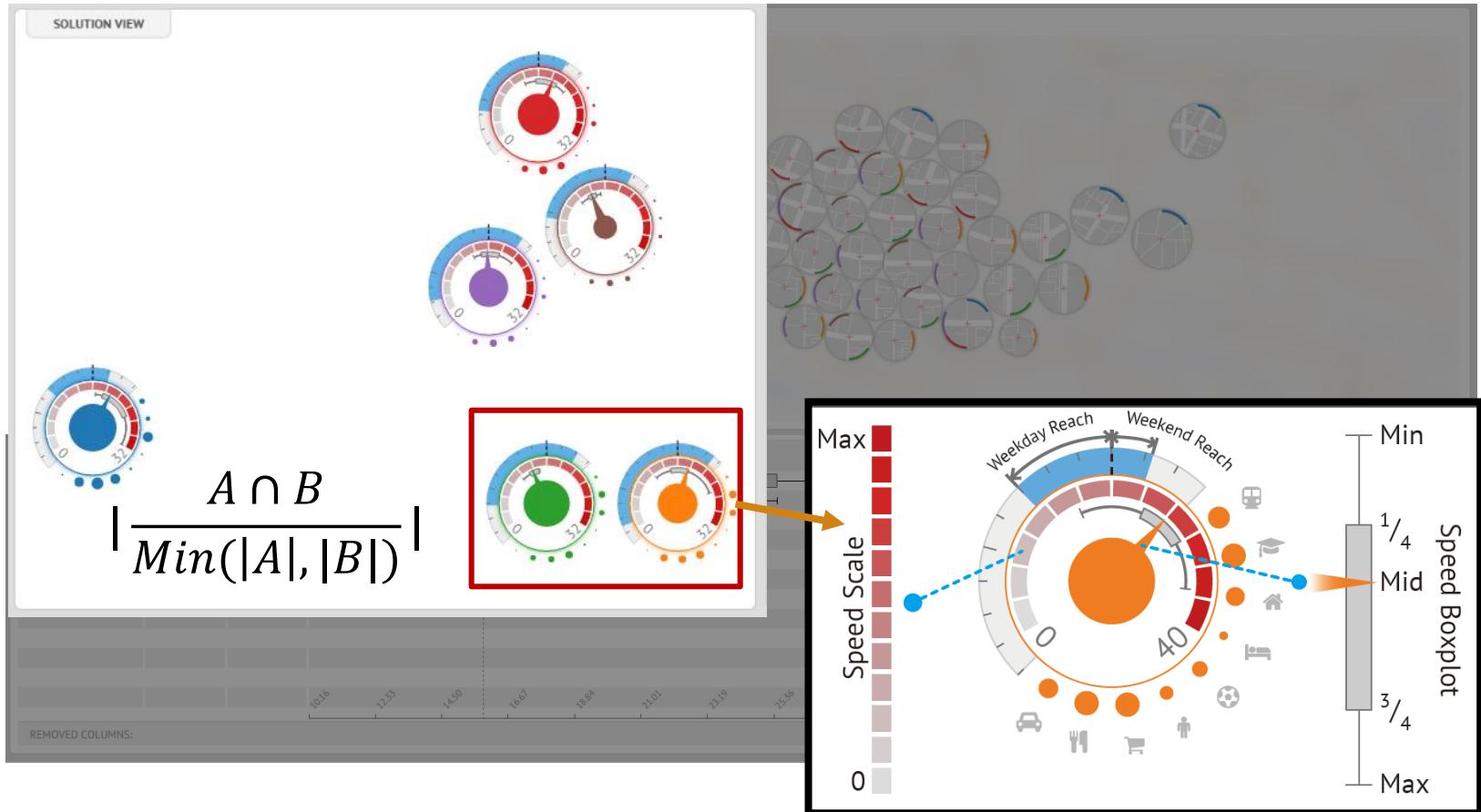
根据相似性方案可以被分成几组？

方案之间的排名是怎样的？哪个方案最好？

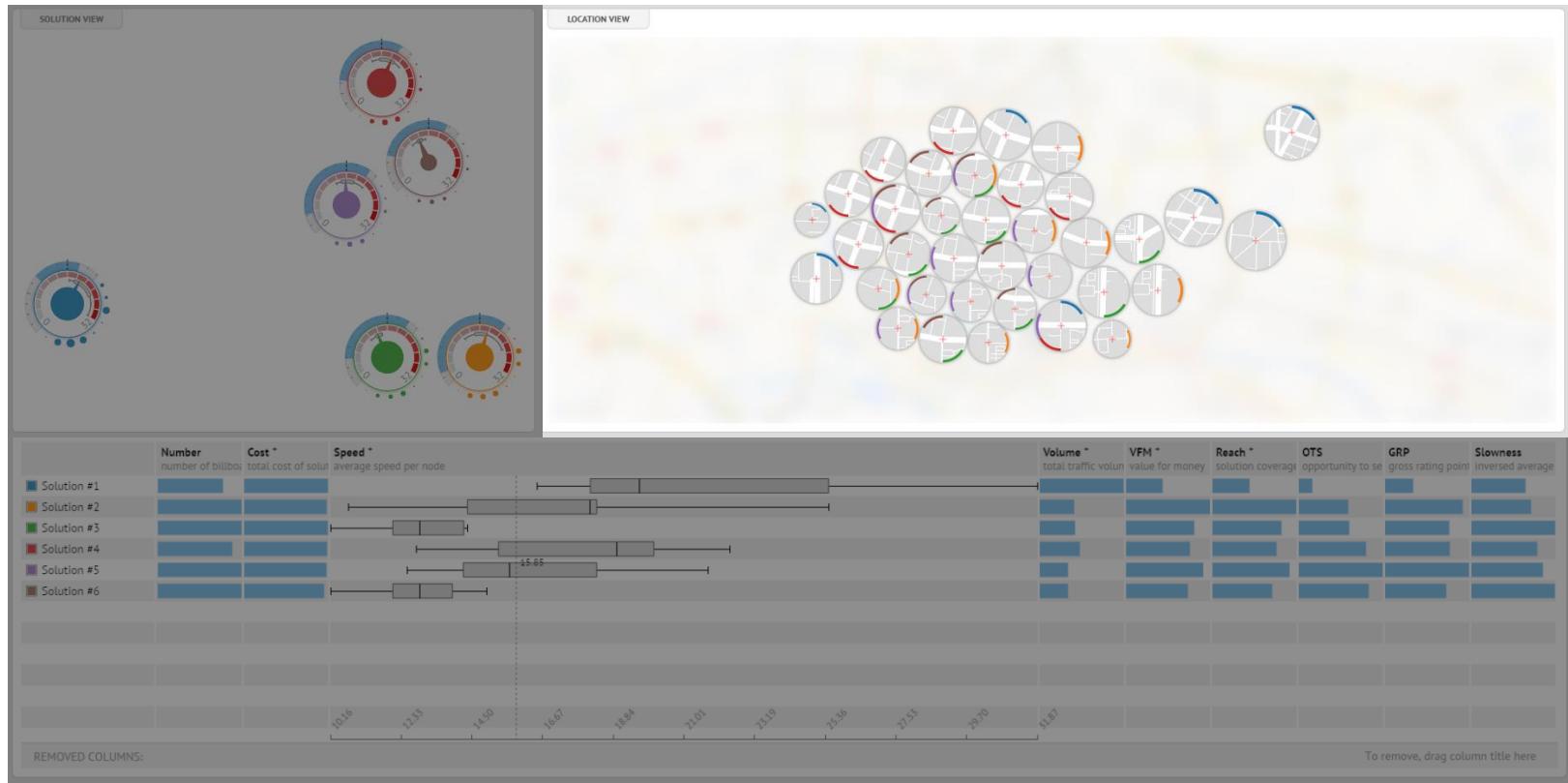
Solution Explorer



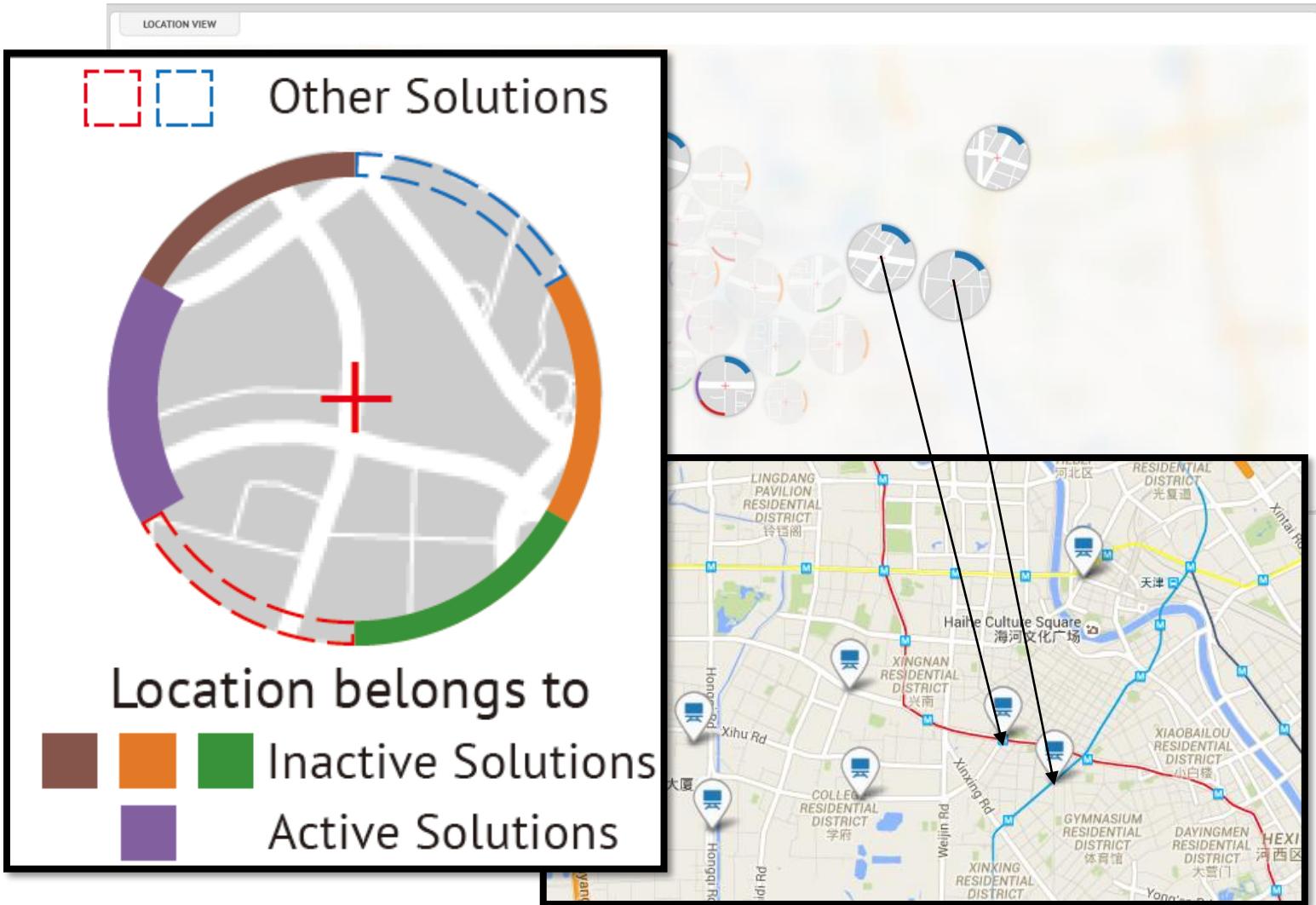
Solution View



Location View



Location Glyph



Solution Explorer的交互





案例

- 基于天津数据集的案例分析
 - 路网数据：133,726条道路片段，99,007个顶点
 - GPS轨迹数据：400万条出租车轨迹
 - POI数据：15万个坐标

Case Video



Case Studies

SmartAdp云系统



SmartAdp Yingcai

smartadp.chinacloudapp.cn

SmartAdP Save Load

Data Set
Using: Tianjin Taxi Trajectory

Target Area
OD Heatmap Road
39.1072, 117.1676 Click to remove
39.1071, 117.1828 Click to remove
39.1166, 117.1691 Click to remove

Solution Area
Mark favorite places
39.1086, 117.1532 Click to remove
39.1126, 117.1571 Click to remove
39.1085, 117.1597 Click to remove
39.1112, 117.1495 Click to remove
39.1111, 117.1562 Click to remove
39.1128, 117.1598 Click to remove

Solution Parameters
Number of Billboards: 5

Generate Solution

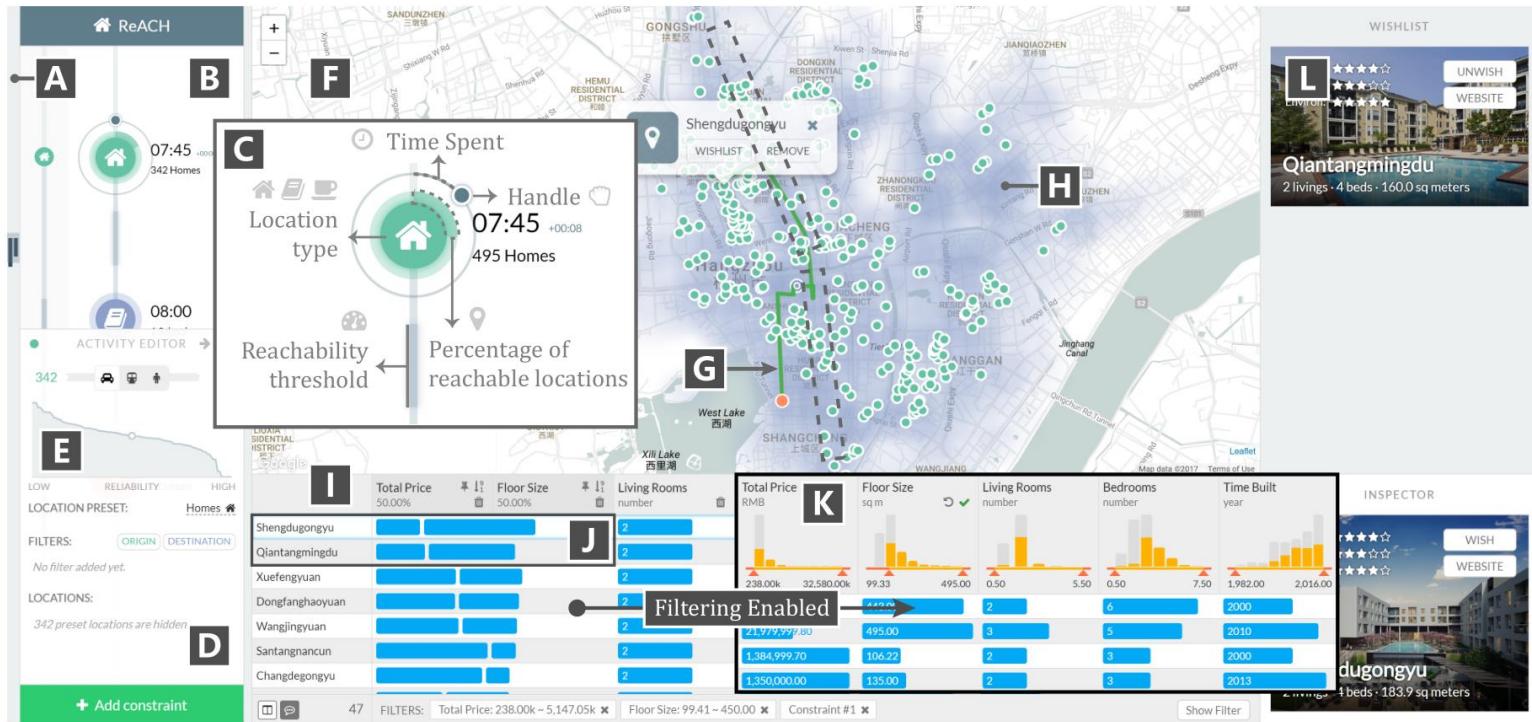
Solution Explorer

SOLUTION PREVIEW

N	C	S	V	M	R	O	G
Number of Billboards: 5	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 5	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 5	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 5	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 5	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 8	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		
Number of Billboards: 8	Target to Normal: 1000	Temporal Filter: none	Speed Less Than: N/A	Edit	Delete		

The map displays the city of Tianjin with various residential districts labeled: Residential District, XINGYELICUN, XINGNAN Residential District, COLLEGE Residential District, XINXING Residential District, GYMNASIUM Residential District, DAYINGMEN Residential District, TAOFUAN Residential District, YUEXIU ROAD Residential District, and HEPING. A red line traces a taxi trajectory from the west through several districts. Blue and pink shaded areas represent target and solution regions. Below the map, seven bar charts show the distribution of billboards across categories N, C, S, V, M, R, O, and G for different solutions.

<http://smartadp.chinacloudapp.cn/>



HomeFinder Revisited: Finding Ideal Homes with Reachability-centric Multi-Criteria Decision Making

ACM SIGCHI 2018



浙江
大学
ZHEJIANG UNIVERSITY

Microsoft
Research
微软亚洲研究院

住房选址



选择合适的住房需要综合考虑诸多因素



地点



价格



尺寸、布局

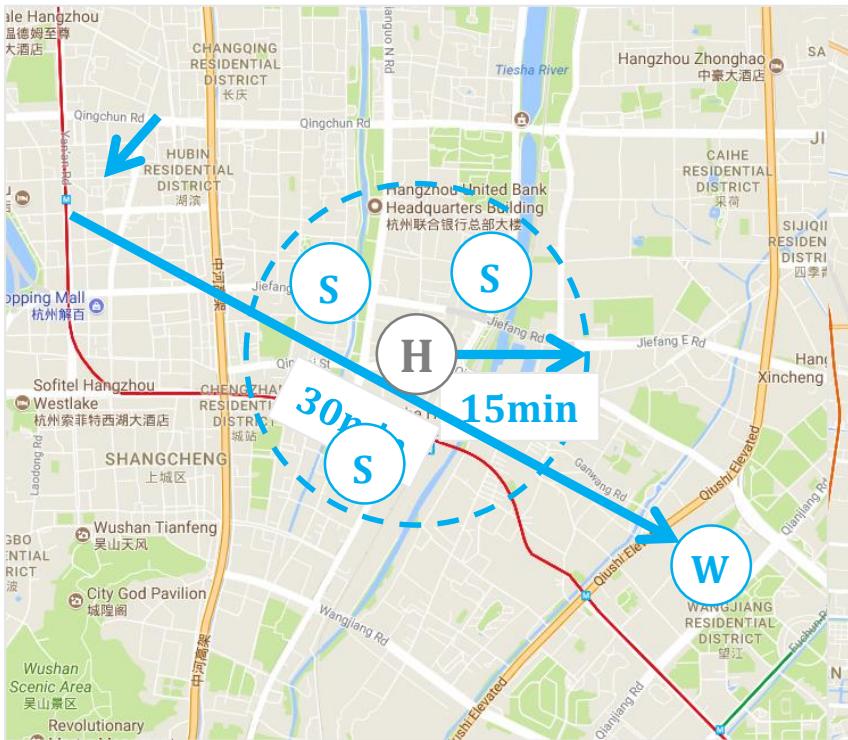
地点是其中最重要的因素之一

地点，地点，地点



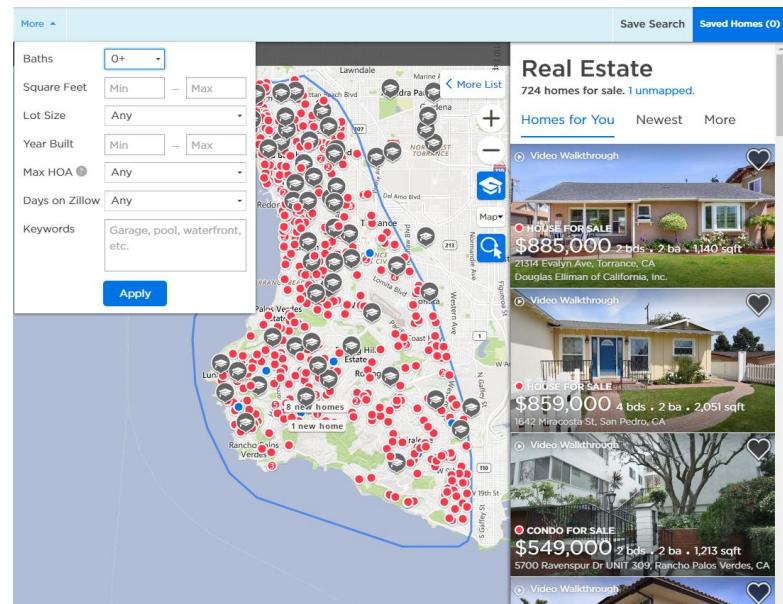
帮我找一处房子让我能够.....

- ✓ 在 8:15 离开家
- ✓ 8:30 前把孩子送到学校
- ✓ 9:00 前到公司



现有住房选址系统

- 国内链家、搜房，国外Zillow、Hubzu等
- 地理搜索功能较为简单
- 大多数过滤局限于使用地理距离

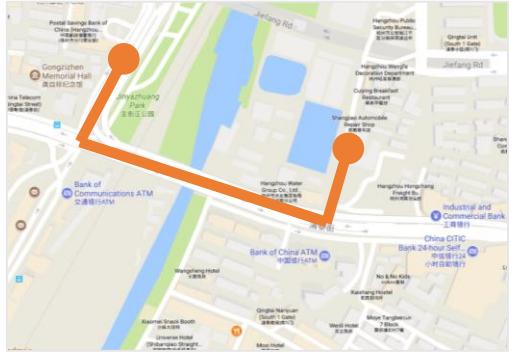


从地理距离到可达性



基于距离

- 使用欧氏距离
- 无视空间上下文
- 速度快，但是严重不准确



基于路网

- 使用道路距离
- 无视交通路况
- 不准确



车辆轨迹

基于可达性

- 使用历史轨迹数据
- 基于海量轨迹推断给定时间内所需的旅行时间
- 准确、缓慢



主要挑战

从包括可达性的多属性住房数据中，发现满足用户日常行程和个人偏好的住宅地点。

主要挑战

可达性计算的效率

日常行程的表示

个人偏好的融合

主要挑战

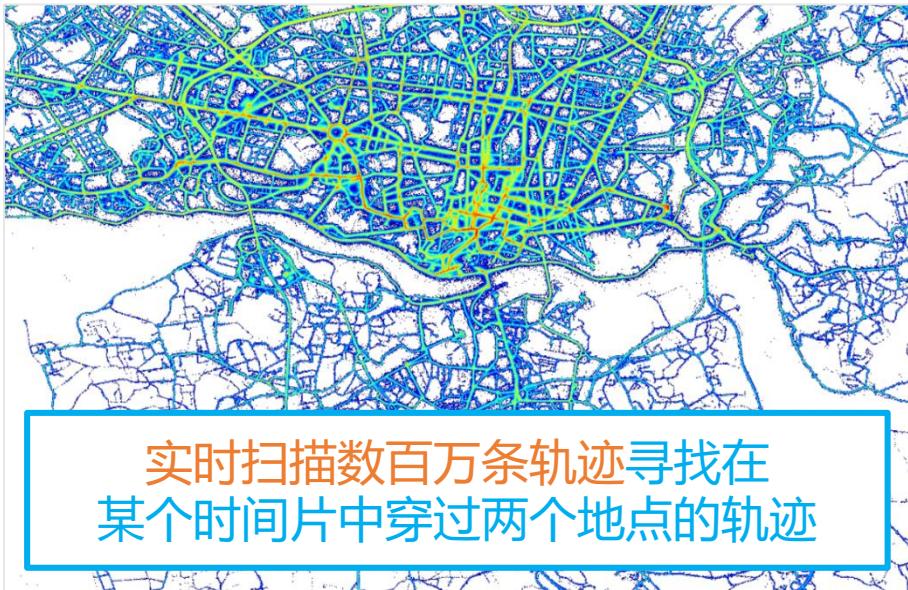
从包括可达性的多属性住房数据中，发现满足用户日常行程和个人偏好的住宅地点。

主要挑战

可达性计算的效率

日常行程的表示

个人偏好的融合



主要挑战

从包括可达性的多属性住房数据中，发现满足用户日常行程和个人偏好的住宅地点。

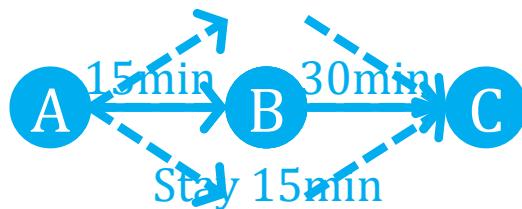
主要挑战

可达性计算的效率

日常行程的表示

个人偏好的融合

1. 多阶段的行程规划



任何在(A)到(B)之间停留15分钟的路径(B)都可以

主要挑战

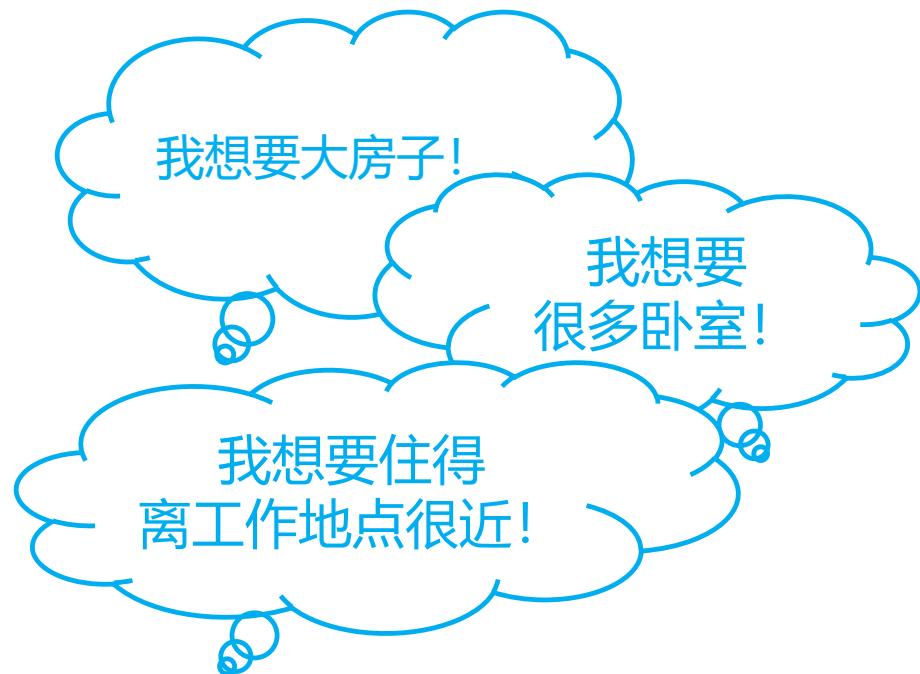
从包括可达性的多属性住房数据中，发现满足用户日常行程和个人偏好的住宅地点。

主要挑战

可达性计算的效率

日常行程的表示

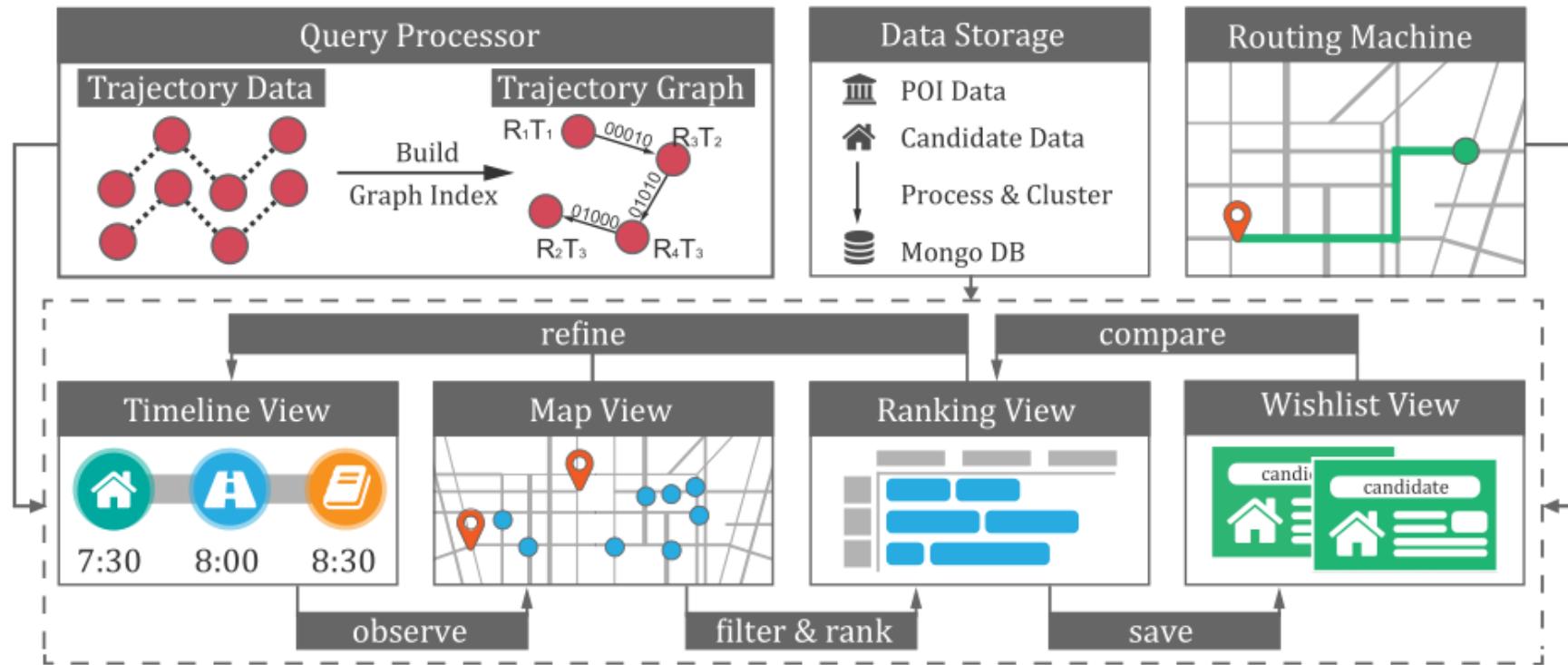
个人偏好的融合



住宅选址的可视分析



借助先进的数据挖掘模型计算可达概率，帮助用户基于模型计算结果交互地筛选和分析候选住宅。

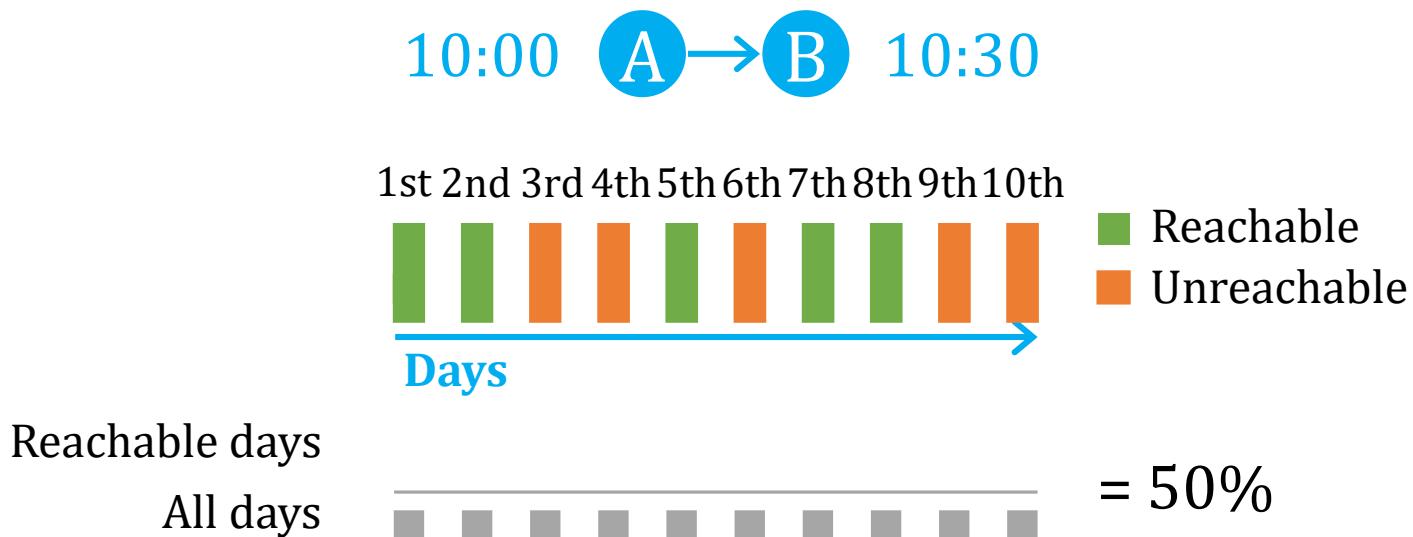


HomeFinder的可视分析框架

可达概率

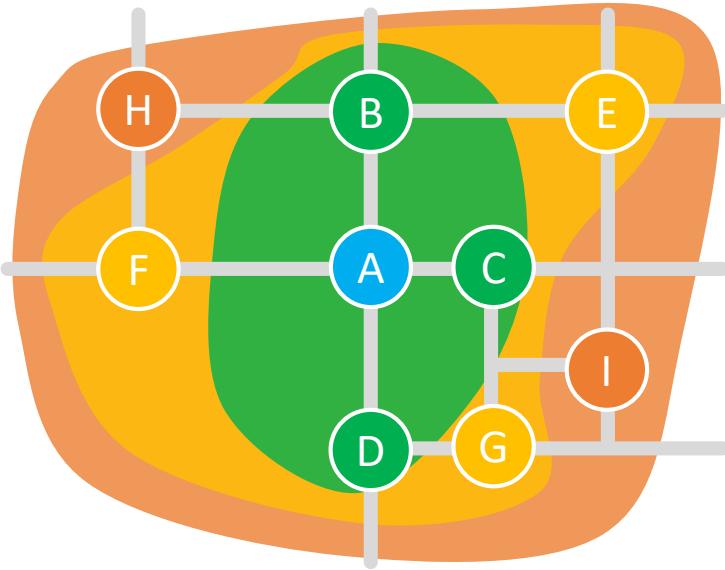


Compute reachable probability **empirically** from taxi trajectories with spatio-temporal indices.



可达区域

The probability of reaching any location inside the region from A is greater than a specified **reachability threshold t** .



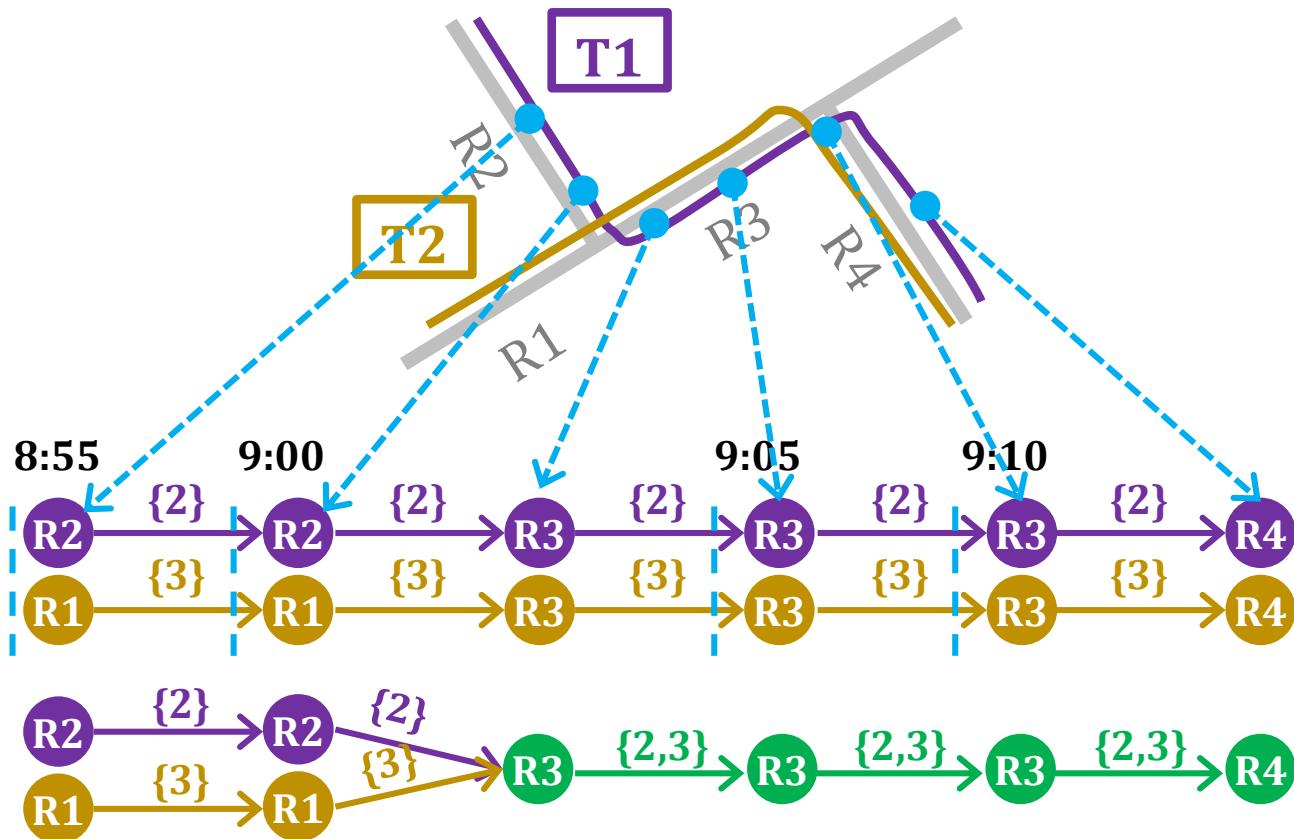


可达性计算的瓶颈

- Computation is slow and inefficient
 - takes 20 seconds to query a 20-minute reachable region
- Reachability is computed only for partial roads
 - relies on predefined thresholds to stop searching early
- Only Individual trajectories are considered
 - fails to alleviate data sparsity by joining consecutive trajectories

利用图计算可达性

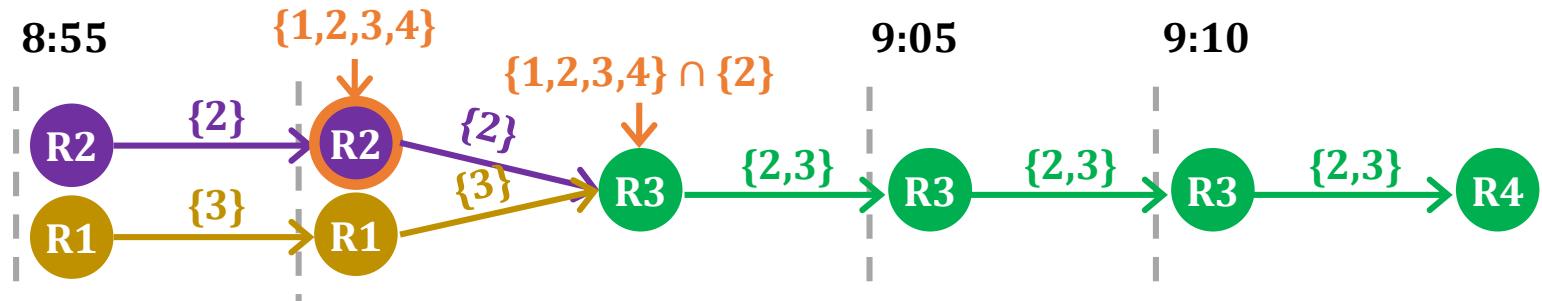
许多轨迹之间时空上共享部分片段。



实时查询



Perform the **breath-first search** starting from the specified origin, assigning the set of reachable days of adjacent nodes with the **intersection** of the sets **in the current node** and **on the edge**.



Details on the paper!



模型性能

- 使用一座大型城市两个月的出租车轨迹数据集测试
 - 大约 20亿 GPS坐标点
 - 89.4% 生成的边被压缩移除
 - 将20分钟可达区域的查询时间从三节点集群需要20秒缩短至单工作站运行3-4秒
- 理论时间和空间复杂度提升
 - 空间复杂度 $\Theta(\beta(T)|M||T| + |M||E|)$. $\rightarrow \Theta(\gamma(T)\beta(T)|M||T|)$.
 - 时间复杂度 $\Theta(\beta(T)|M|^2|T|)$. $\rightarrow \Theta(|M||E| + \gamma(T)\beta(T)|M||T|)$.

用户需求



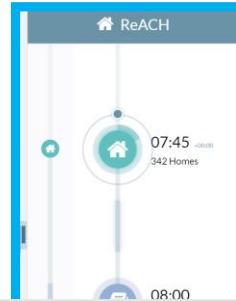
任务一：交互式地构建可达性限制条件

任务二：可视化地探索可达性计算结果

任务三：迭代地优化可达性限制条件

任务四：高效地过滤并排序候选住宅

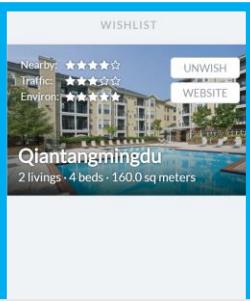
可视化系统设计



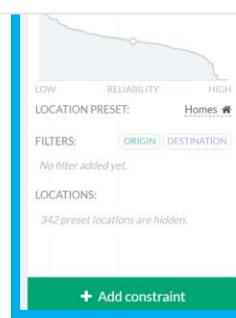
时间轴视图



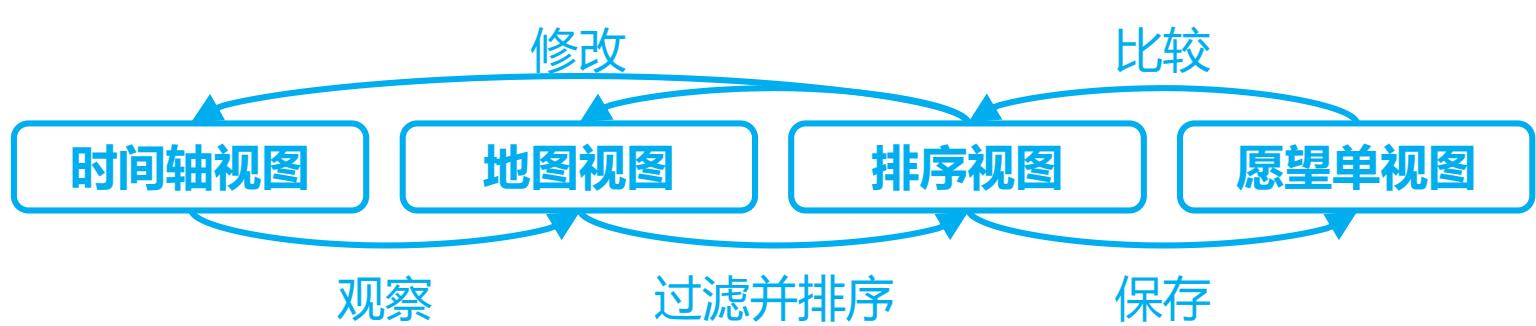
地图视图



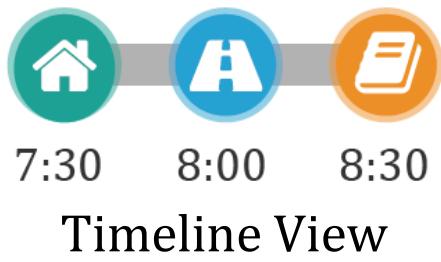
愿望单视图



排序视图



时间轴视图



Task #1: 交互式地构建可达性限制条件

Task #3: 迭代地优化可达性限制条件

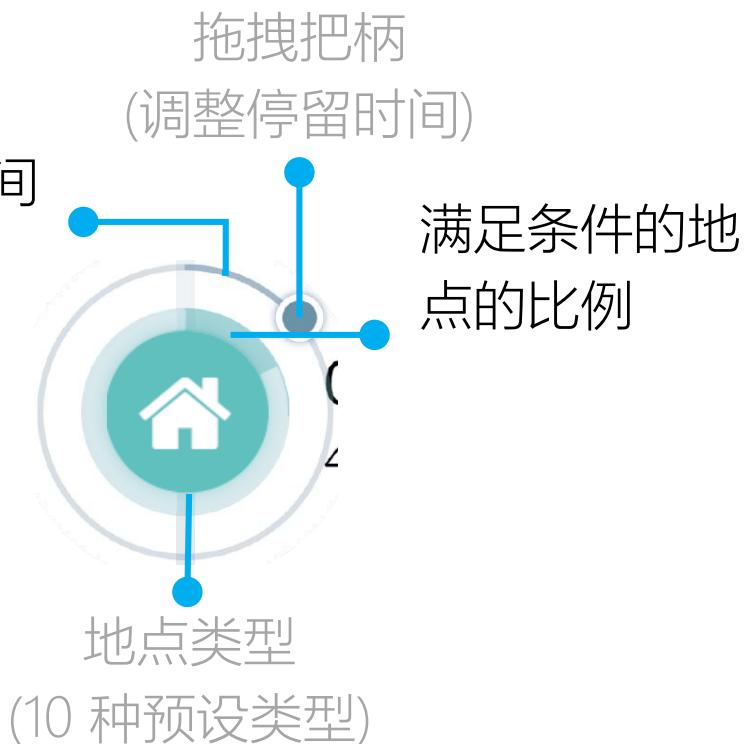
时间轴视图



计划停留时间

可达性阈值
(最低可达概率)

可达性限制条件



拖拽把柄

(调整停留时间)

满足条件的地
点的比例

地点类型

(10 种预设类型)

住宅可达性的交互设置



The figure consists of three side-by-side screenshots of a mobile application called "ReACH".

- Left Screenshot:** Shows a search interface with a map and a vertical timeline on the left. A blue callout points to the "ACTIVITY EDITOR" button at the top right of the map area. Another blue callout points to the "FILTERS" section below it, which includes "ORIGIN" and "DESTINATION" buttons and a message "No filter added yet.". A third blue callout points to the "LOCATIONS" section, which says "342 preset locations are hidden.".
- Middle Screenshot:** Shows a map with several green and orange location markers. A blue callout points to the top right corner where icons for car, bus, and pedestrian are displayed. Another blue callout points to the center of the screen where a list of locations is shown with columns for "Total Price RMB", "Floor Size sqm", and "Link".
- Right Screenshot:** Shows a vertical timeline on the left with markers for "08:00 591 Homes", "08:20 69 School", and "08:40 1 Custom". A blue callout points to the "08:20" entry. The main area shows a map with a red route line and a list of locations below it.

自适应缩放

直观时间拖拽

设置停留时间

符合条件的地点数目

切换不同的交通方式
修改可达性阈值

可达性阈值

筛选所选地点

地图视图



Task #2: 可可视化地探索可达性计算结果

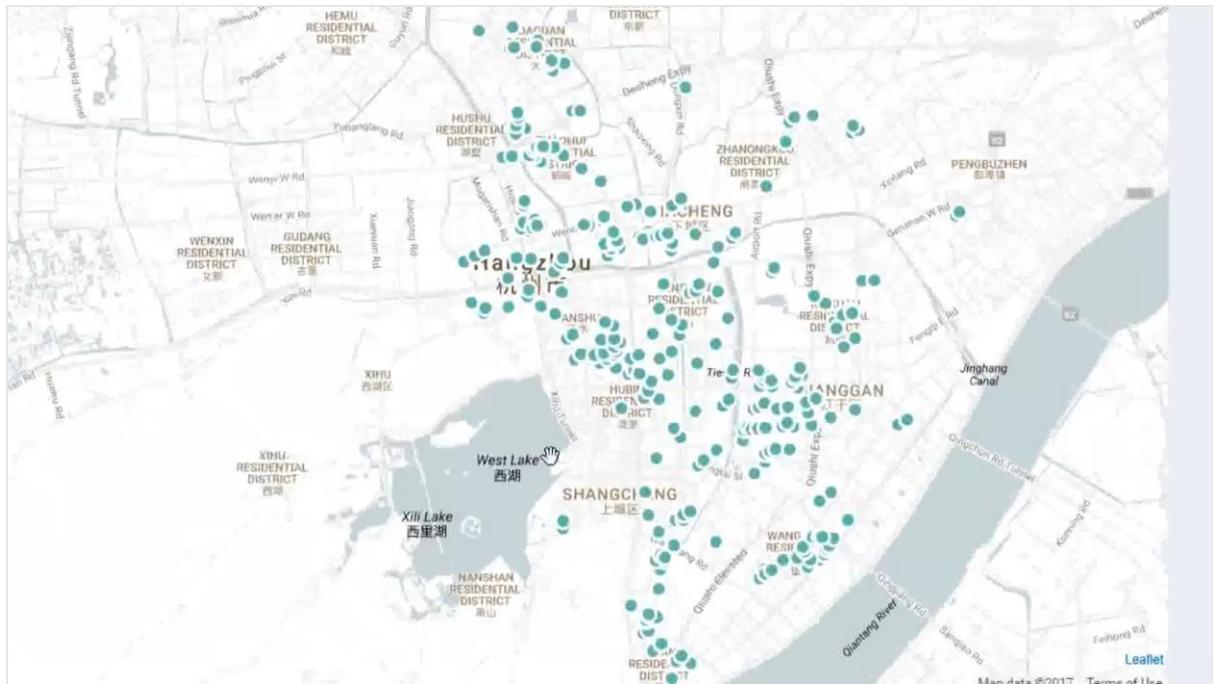
地图视图



Map-centred
exploratory
approach

Candidate
Layer

Constraint
Layer



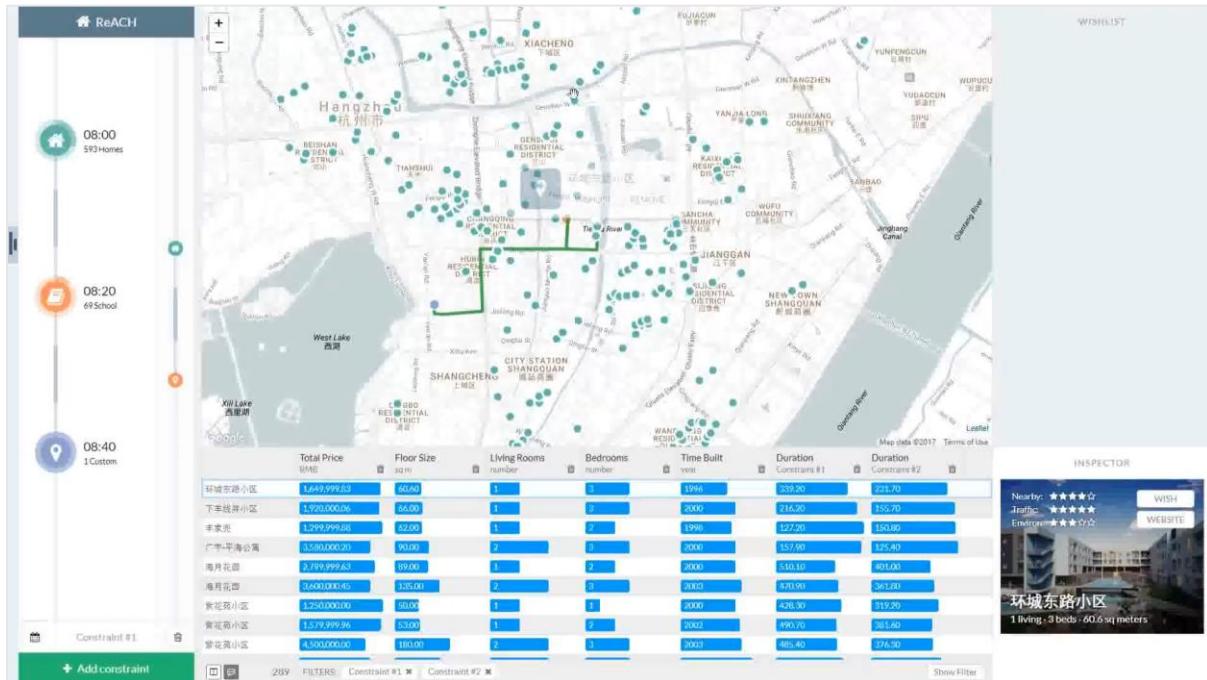
地图视图



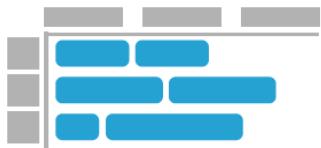
Map-centred
exploratory
approach

Candidate
Layer

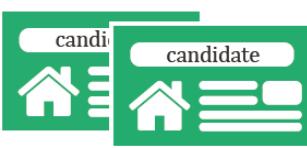
Constraint
Layer



排序视图与愿望单视图



Ranking View



Wishlist View

Task #4: 高效地过滤并排序候选住宅

住宅的可视对比和筛选

属性过滤

合并排序

保存比较



住宅的可视对比和筛选

属性过滤

合并排序

保存比较



住宅的可视对比和筛选



属性过滤

合并排序

保存比较



评估与实验



- 模型评估与分析
- 两个使用案例
 - Orchestration of Constraints
 - Comparison with Reachability and Ranking (details on the paper)
- 基于任务的用户调研分析



模型评估与分析

- Evaluated with two-month taxi trajectories in a large city
 - nearly 2 billion GPS locations
 - 89.4% edges compressed and removed
 - the execution time of a 20-minute reachable region query shortened to 3-4 seconds on a single workstation (previously 20 seconds)
- Space and time t' $\Theta(\beta(T)|M||T| + |M||E|)$. \rightarrow $\Theta(\gamma(T)\beta(T)|M||T|)$.
paper) $\Theta(\beta(T)|M|^2|T|)$. \rightarrow $\Theta(|M||E| + \gamma(T)\beta(T)|M||T|)$.
 - space complexity:
 - time complexity:

Orchestration of Constraints



- Quick demonstration of usability
 - interactively orchestrate constraints
 - refine daily routines based on the reachability
- John, a father of two kids, looking for an apartment

For the full demonstration of advanced features,
please find the video at <https://youtu.be/WcjMXAfuygg>
or search "HomeFinder Revisited" at YouTube

Orchestration of Constraints



ReACH

Map date: 8/2017 · Terms of Use

	Total Price RMB	Floor Size sqm	Living Rooms number	Bedrooms number	Time Built year
江南水乡	14,000,000.53	443.00	2	6	2000
江南水乡	21,979,999.80	495.00	3	5	2010
荷韵江南苑	1,350,000.35	95.00	2	3	2009
绿都御景蓝湾	572,688.96	70.99	2	2	2014
绿都御景蓝湾	905,454.55	84.00	2	3	2015
绿都御景蓝湾	1,361,999.66	97.20	2	4	2012
金城兰庭	1,299,999.70	83.96	2	2	2014
金城兰庭	1,325,000.02	89.36	2	3	2013
东方文化园	1,025,000.09	89.00	2	3	2015

No constraint yet

+ Add constraint

1927 FILTERS: Show Filter

INSPECTOR

No candidate selected.

总结与展望



- 城市数据可视分析通过可视化交互，结合数据挖掘与领域知识，完成城市数据的理解、推理和决策等任务
- 城市数据可视分析仍然存在许多挑战
 - 数据复杂：异构性、稀疏性、动态变化、不确定性等
 - 任务多样：规划选址、关联分析、异常检测等
- 人类和计算机协同合作
 - Human-machine intelligence



谢谢!

巫英才

www.ycwu.org

ycwu@zju.edu.cn

Acknowledgements



- 科技部重点研发专项(2018YFB1004300), 国家自然科学基金 (U1609217, 61502416, 61761136020), MSRA Grant
- Yu Zheng, Jie Bao, Huamin Qu, Wei Chen, Tai-Quan Peng
Jian-Hua Zhu, David Gotz, Daniel Keim, Kwan-Liu

