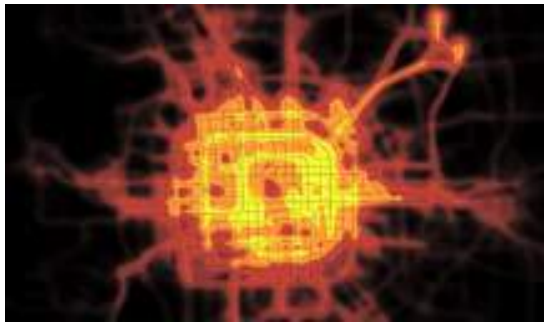
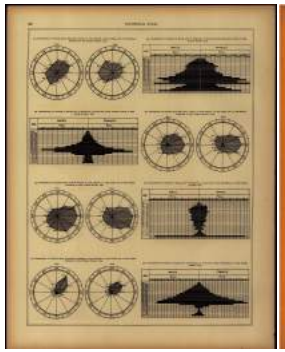
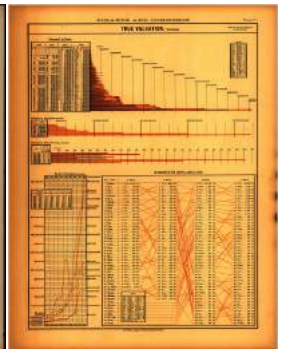
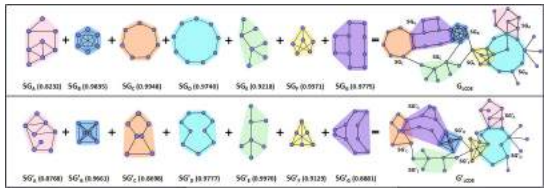
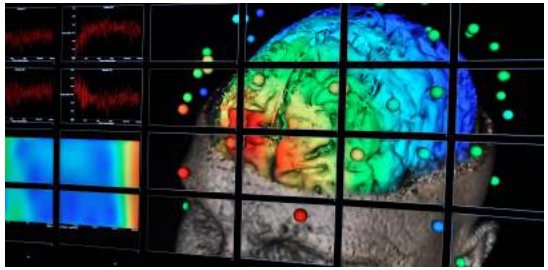










城市数据可视分析

袁晓如

信息科学中心
机器感知与智能教育部重点实验室
北京大学 信息科学技术学院
2016.07.29

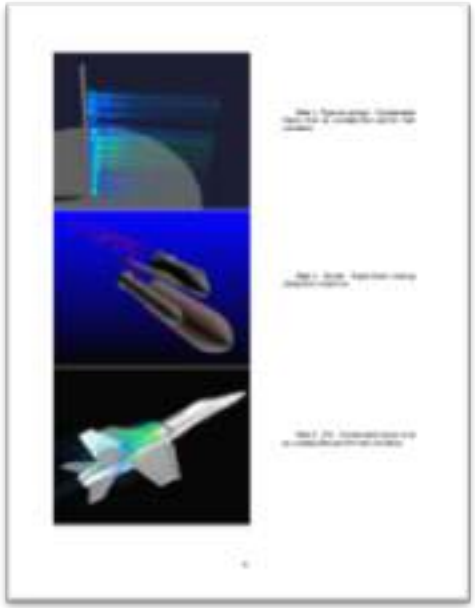
@晓如微博

China 2016 西安论坛

第一篇提到“大数据”的论文
First Paper on “Big Data”

- Michael Cox and David Ellsworth, “Application-Controlled Demand Paging for Out-of-Core Visualization.” Proceedings of the 8th IEEEE conference on Visualization ’97. 1997
- NASA Ames Research Center

ChinaVis 2016 西安论坛



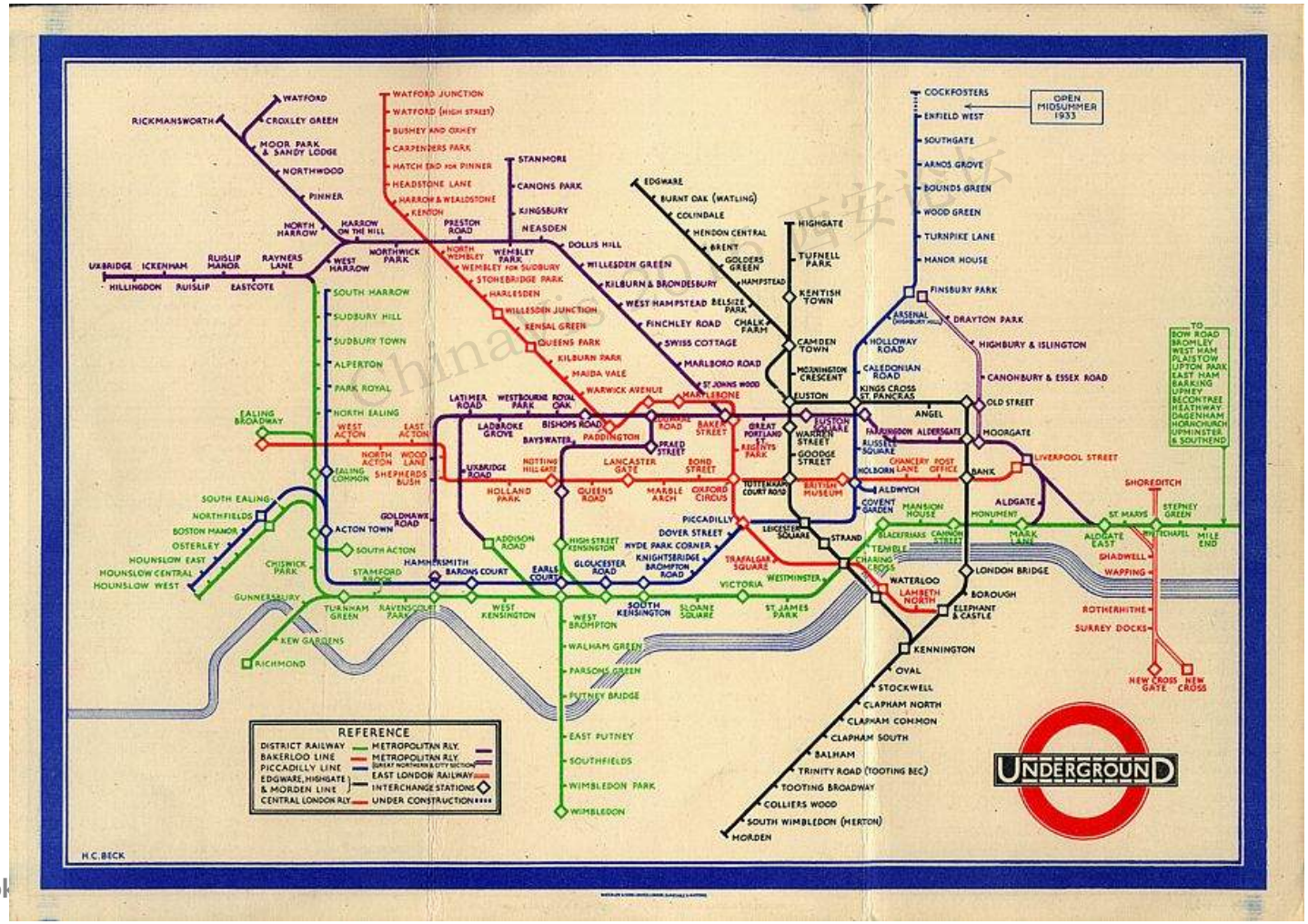
第一篇提到“大数据”的论文
First Paper on “Big Data”

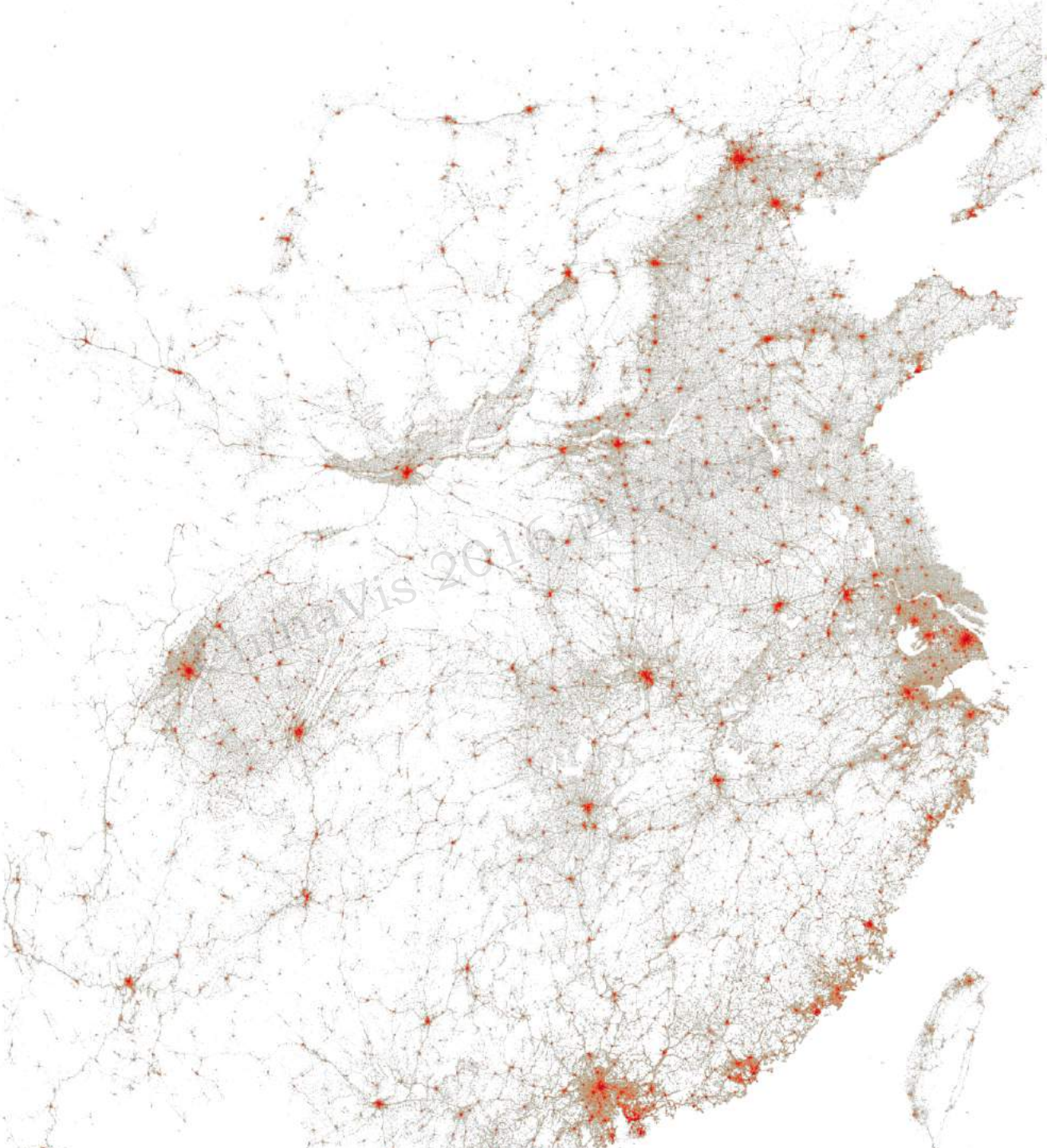
1 Introduction

Visualization provides an interesting challenge for computer systems: data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of *big data*. When data sets do not fit in main memory (*in core*), or when they do not fit even on local disk, the most common solution is to acquire more resources. This *write-a-check* algorithm has two drawbacks. First, if visualization algorithms and tools are worth developing, then they are worth deploying to more production-oriented scientists and engineers who may have on their

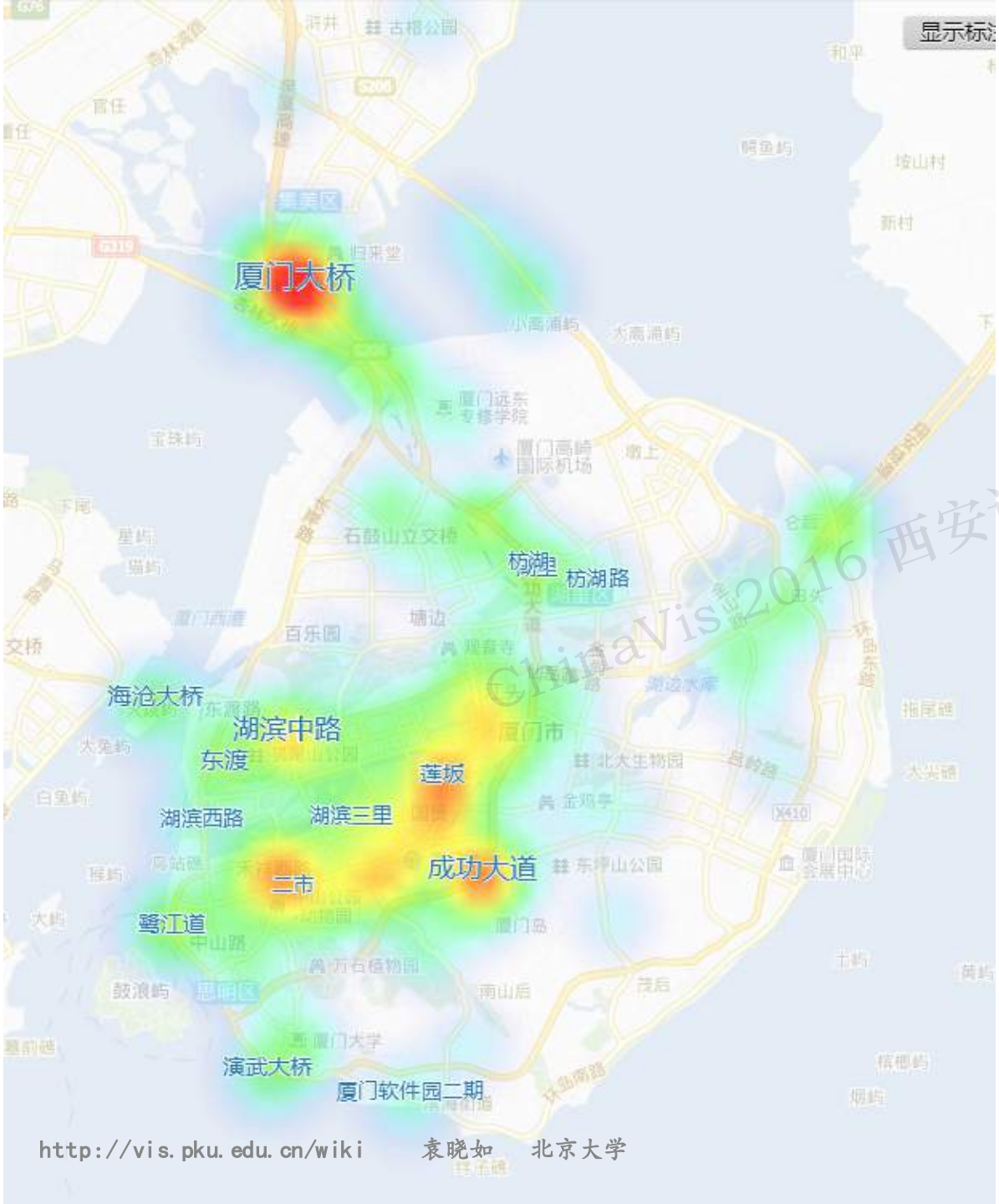
Michael Cox and David Ellsworth, “Application-Controlled Demand Paging for Out-of-Core Visualization.” Proceedings of the 8th IEEEE conference on Visualization ’97. 1997

1933 伦敦地铁图

- 
OPEN MIDSUMMER 1933
REFERENCE
DISTRICT RAILWAY
BAKERLOO LINE
PICCADILLY LINE
EDGWARE, HIGHGATE & MORDEN LINE
CENTRAL LONDON RLY
METROPOLITAN RLY.
METROPOLITAN RLY. (GREAT NORTHERN & CITY SECTION)
EAST LONDON RAILWAY
INTERCHANGE STATIONS
UNDER CONSTRUCTION
UNDERGROUND
H.C.BECK



ChinaVis 2016



关键词:

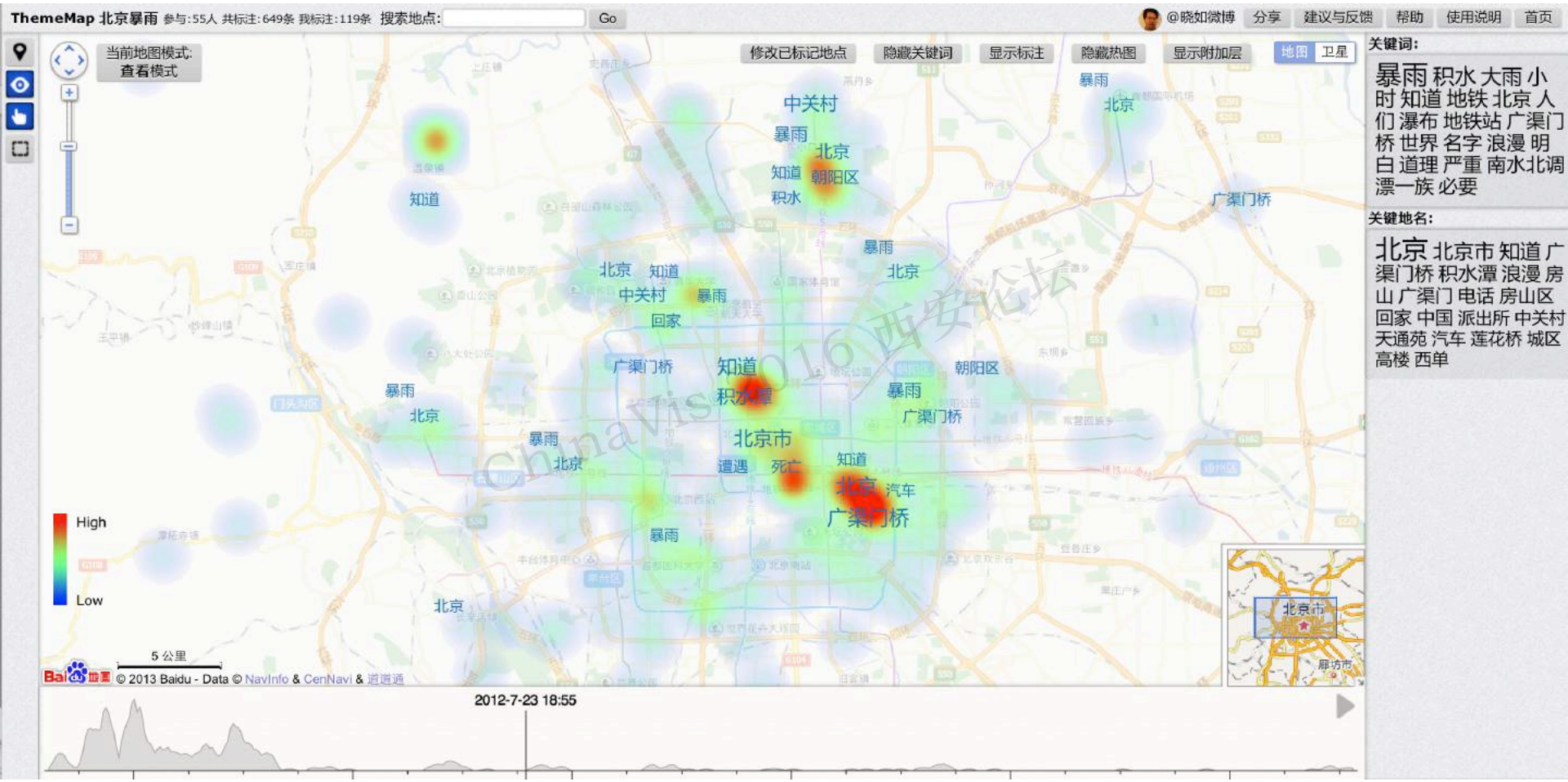
大道 严重 隧道 红绿灯 仙岳 路口 火车站 司机 路段 海沧 时间 上班 知道 轮渡 东渡 方向 中路 开车 公园 车道

关键地名:

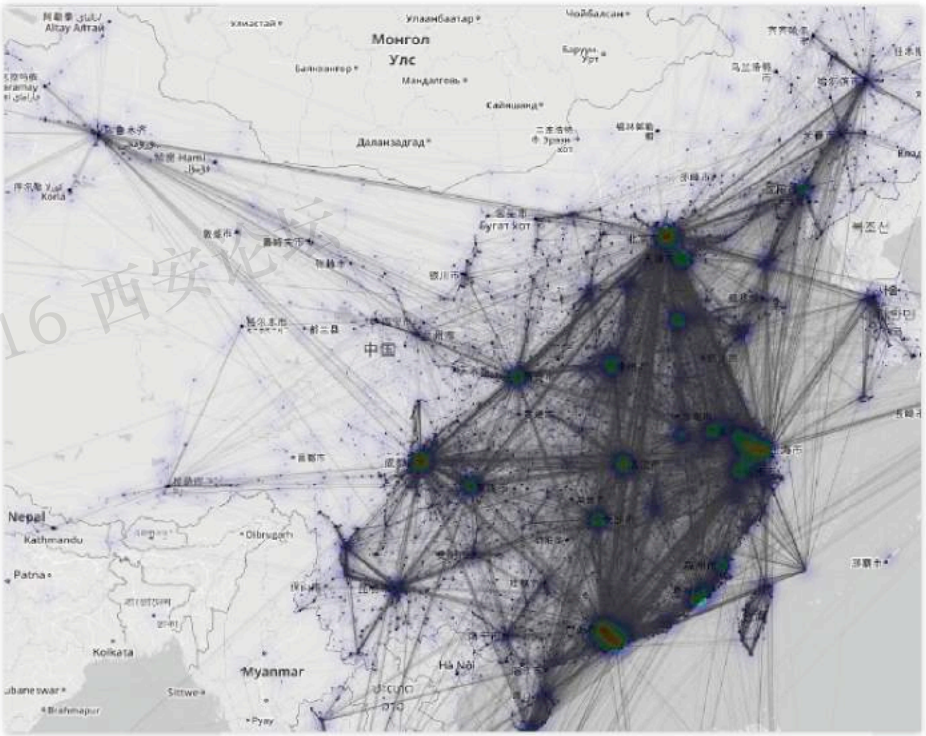
厦门大桥 成功大道 湖滨中路 桥上 地下停车场 东渡 海沧大桥 演武大桥 莲坂 鹭江道 集美 南湖公园 湖滨三里 湖里 厦门软件园二期 二楼 二市 枋湖 湖滨西路 枋湖路

ChinaVis 2016 西安论坛

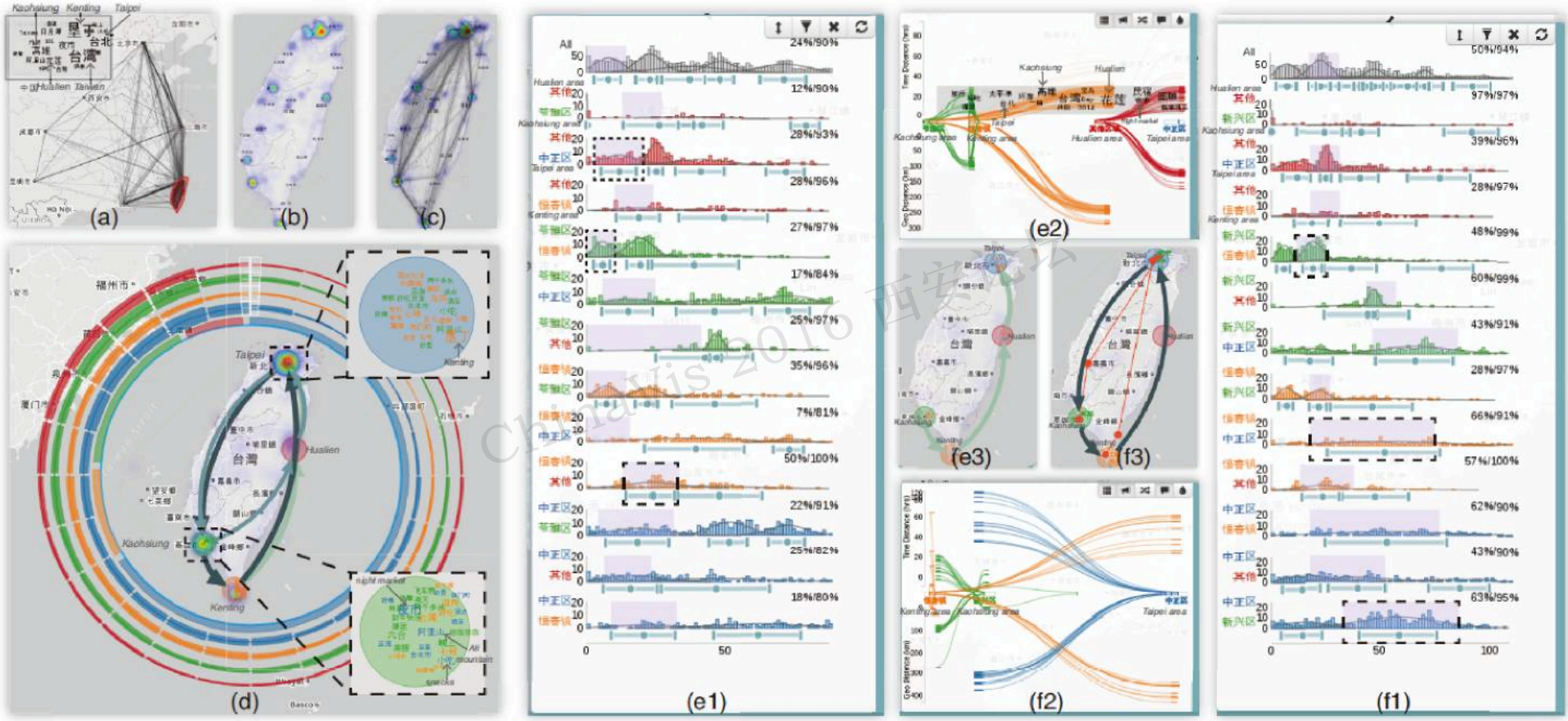
ChinaVis 2016 西安论坛



ChinaVis 2016 西安论坛

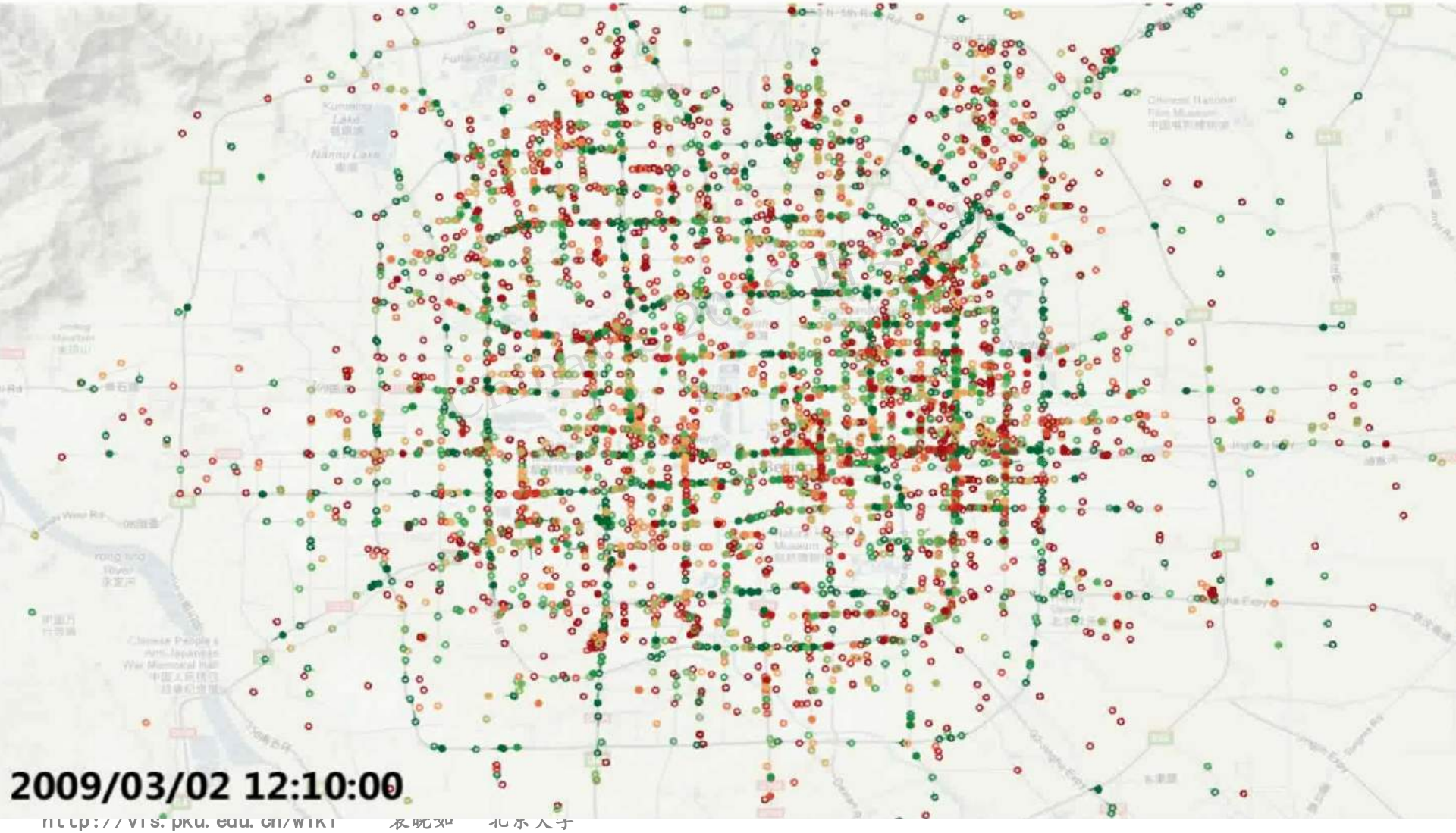


Vis 2016 西安论坛



ChinaVis 2015 西安论坛

北京出租车GPS数据



北京市出租车行驶轨迹的可视化

北京大学可视化与可视分析研究组

2012. 08. 12

VisPKU

出租车轨迹数据

- GPS设备采集
- 空间、时间分布广

- 北京市
- 28,519辆
- 2009年3月2日~25日
- 每30秒采样
- 约34.5 GB

ChinaVis 2016 西安论坛

出租车数据来自数据堂 Datatang.com

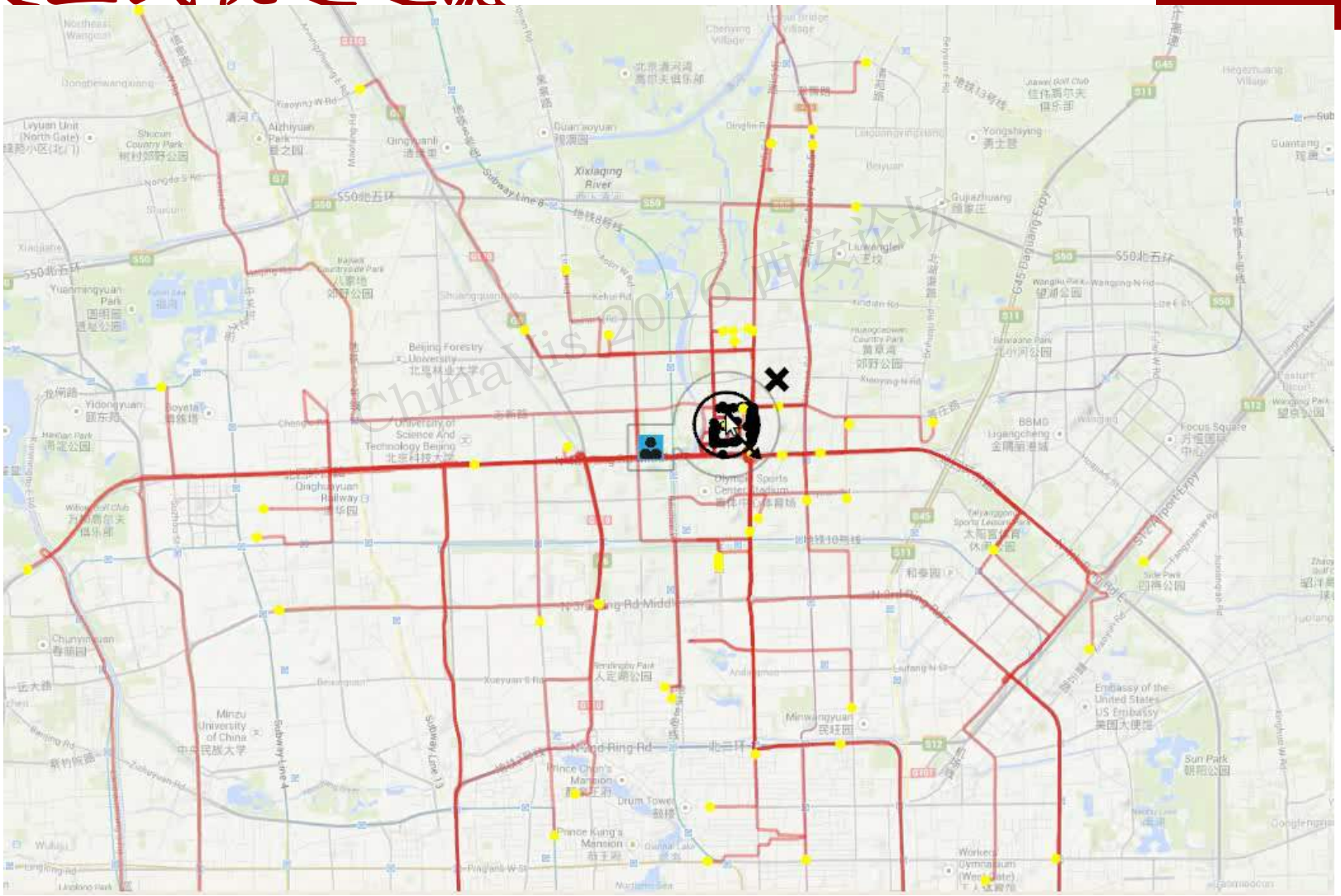
ChinaVis 2015 西安论坛

知微探秘：区域

```
SELECT * FROM T in
Area Where EXISTS
(intersect(A.region,
(SELECT C.polygon
FROM C in Beijing
Where C.center =
'BICC')));
```

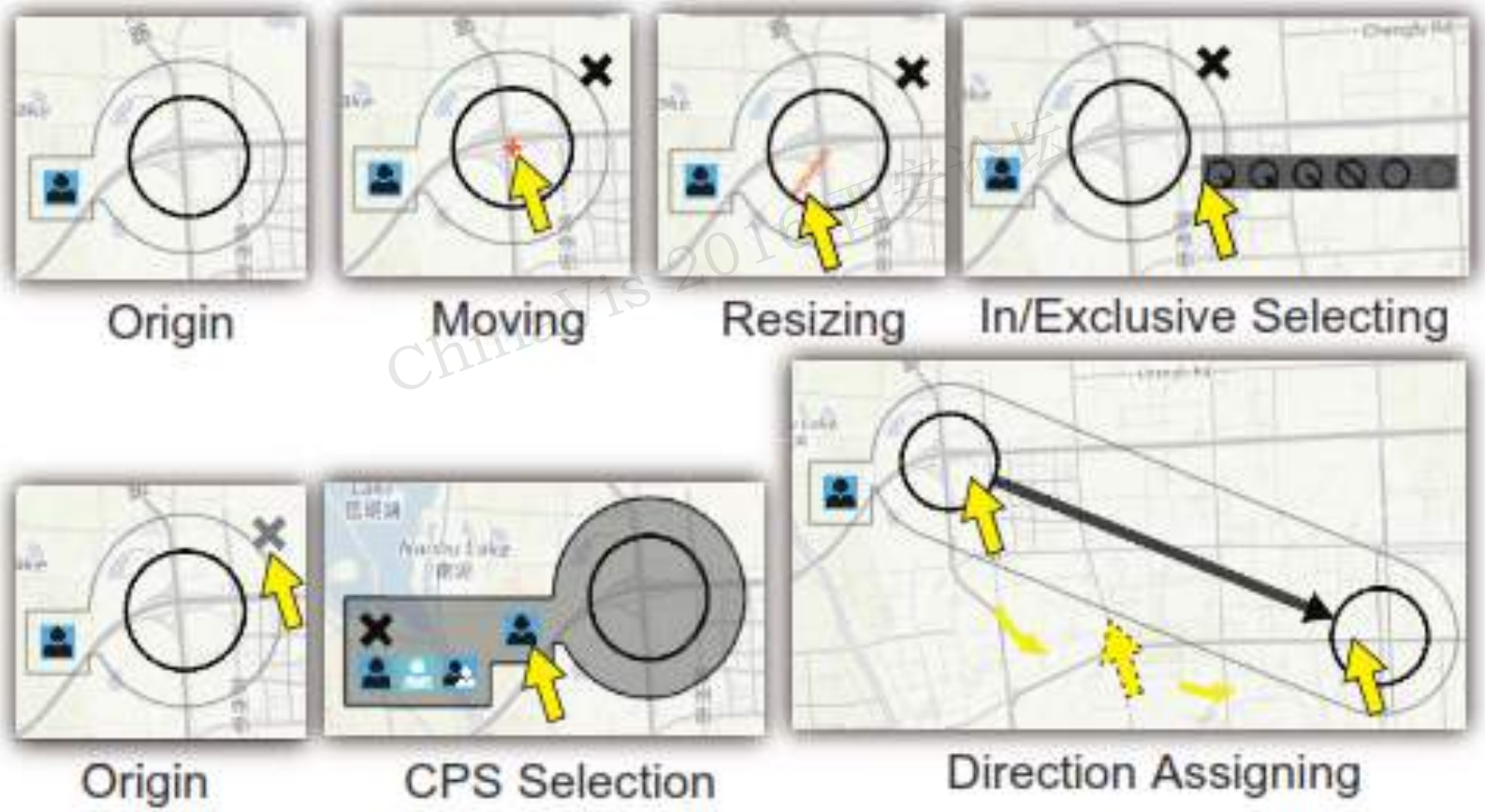
ChinaVis 2015 西安论坛

交互式轨迹过滤



chinaVis 2016 西安论坛

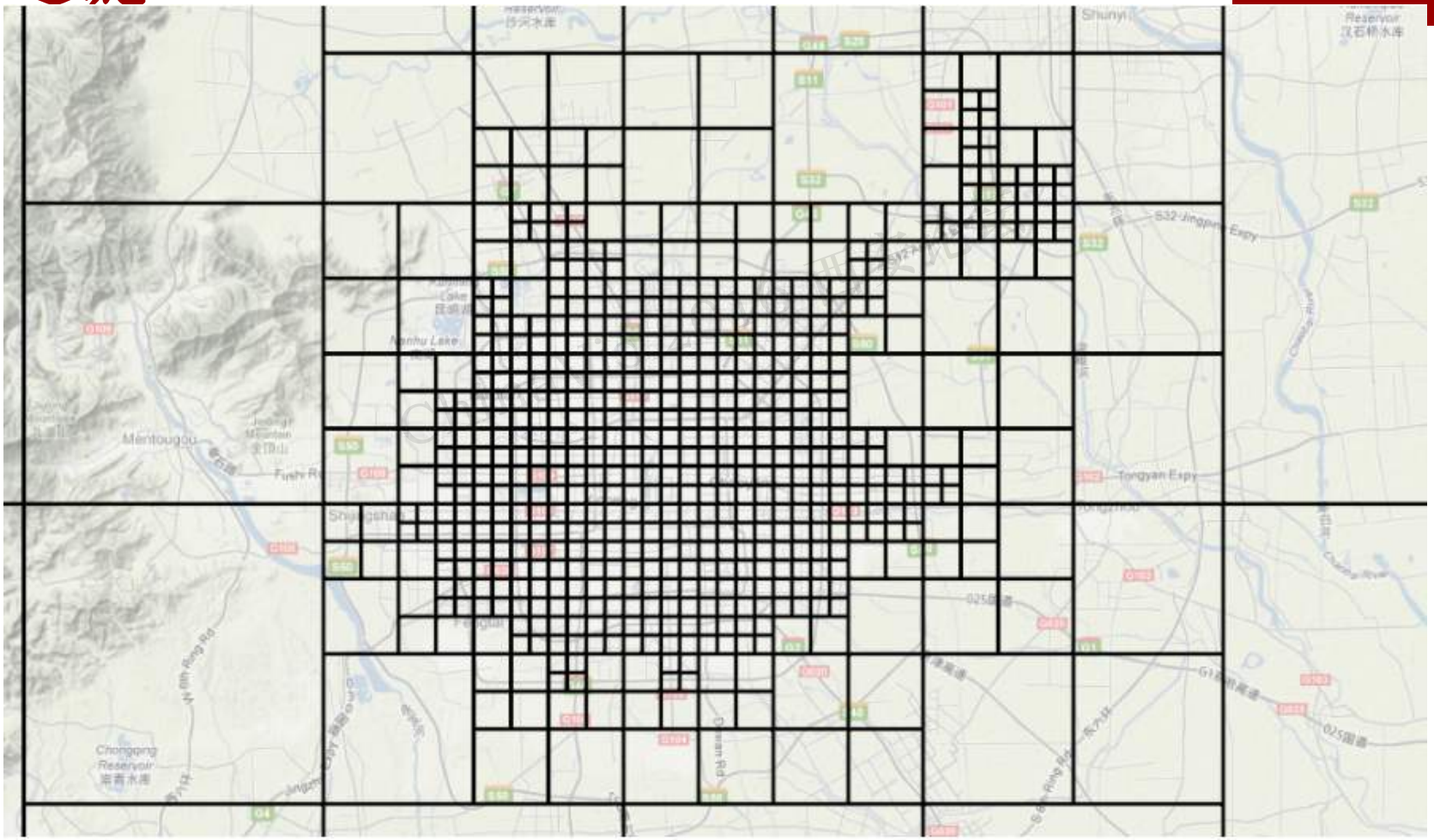
过滤的交互设计



ChinaVis 2016西安论坛

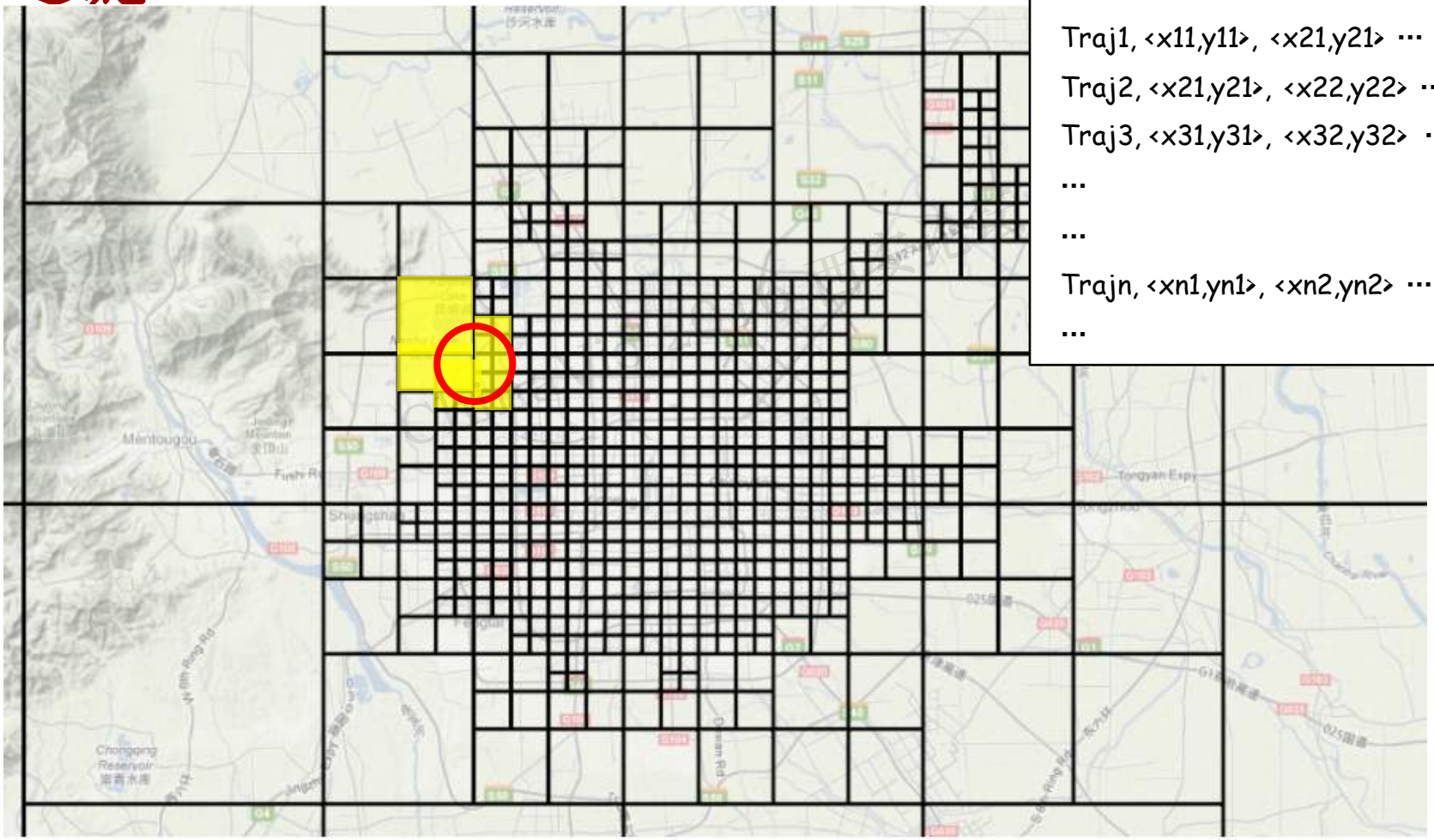
ChinaVis 2016 西安论坛

过滤


Shunyi
Reservoir
S32 Jingpin Expy
Nanhu Lake
Mentougou
Tongyan Expy
Shuangshu
Fengtai
Chongqing
Reservoir
N 6th Ring Rd
S 6th Ring Rd
Chaoba River

空间四叉树索引

过滤



Traj1, <x11,y11>, <x21,y21> …

Traj2, <x21,y21>, <x22,y22> …

Traj3, <x31,y31>, <x32,y32> …

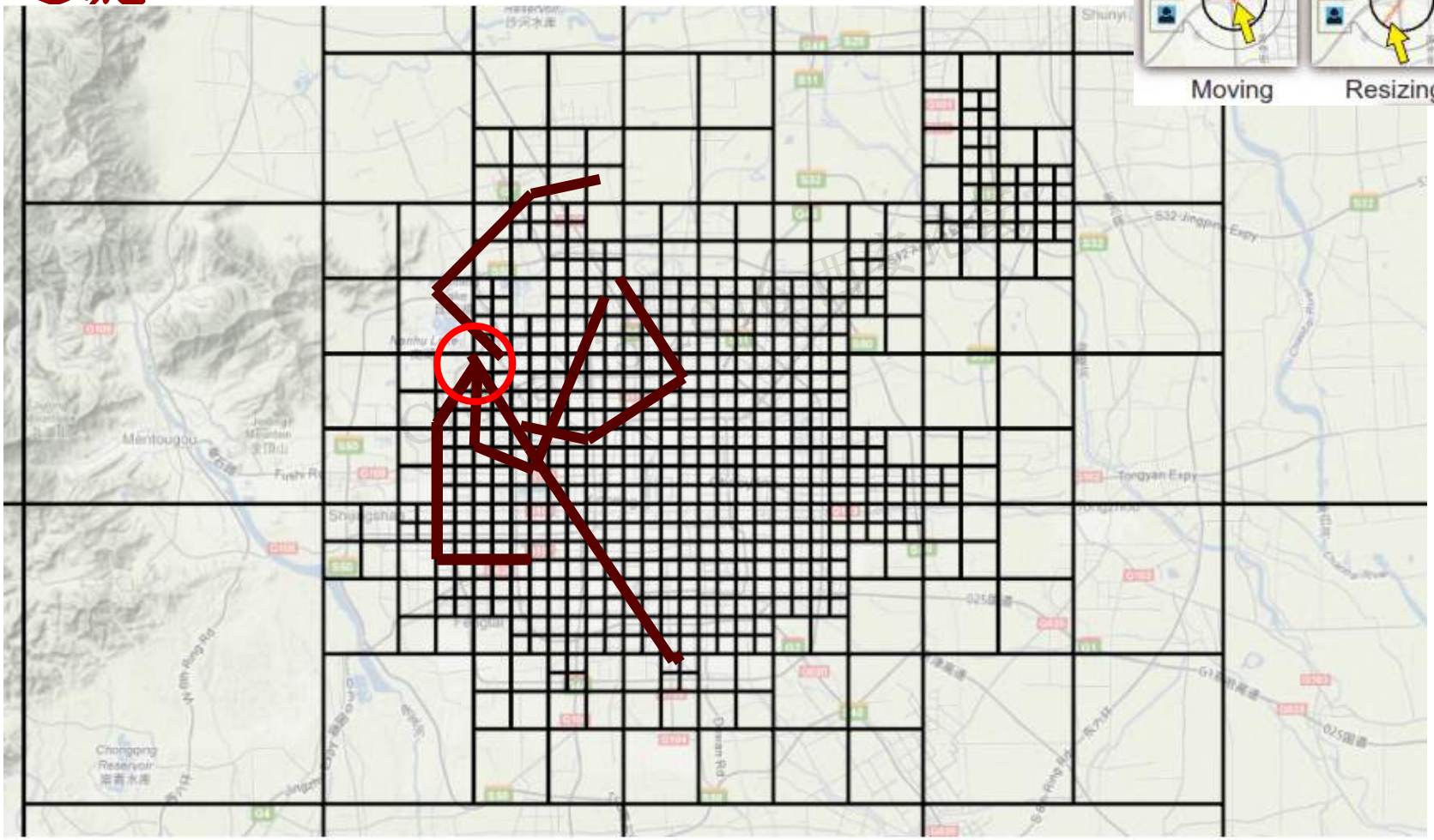
…

…

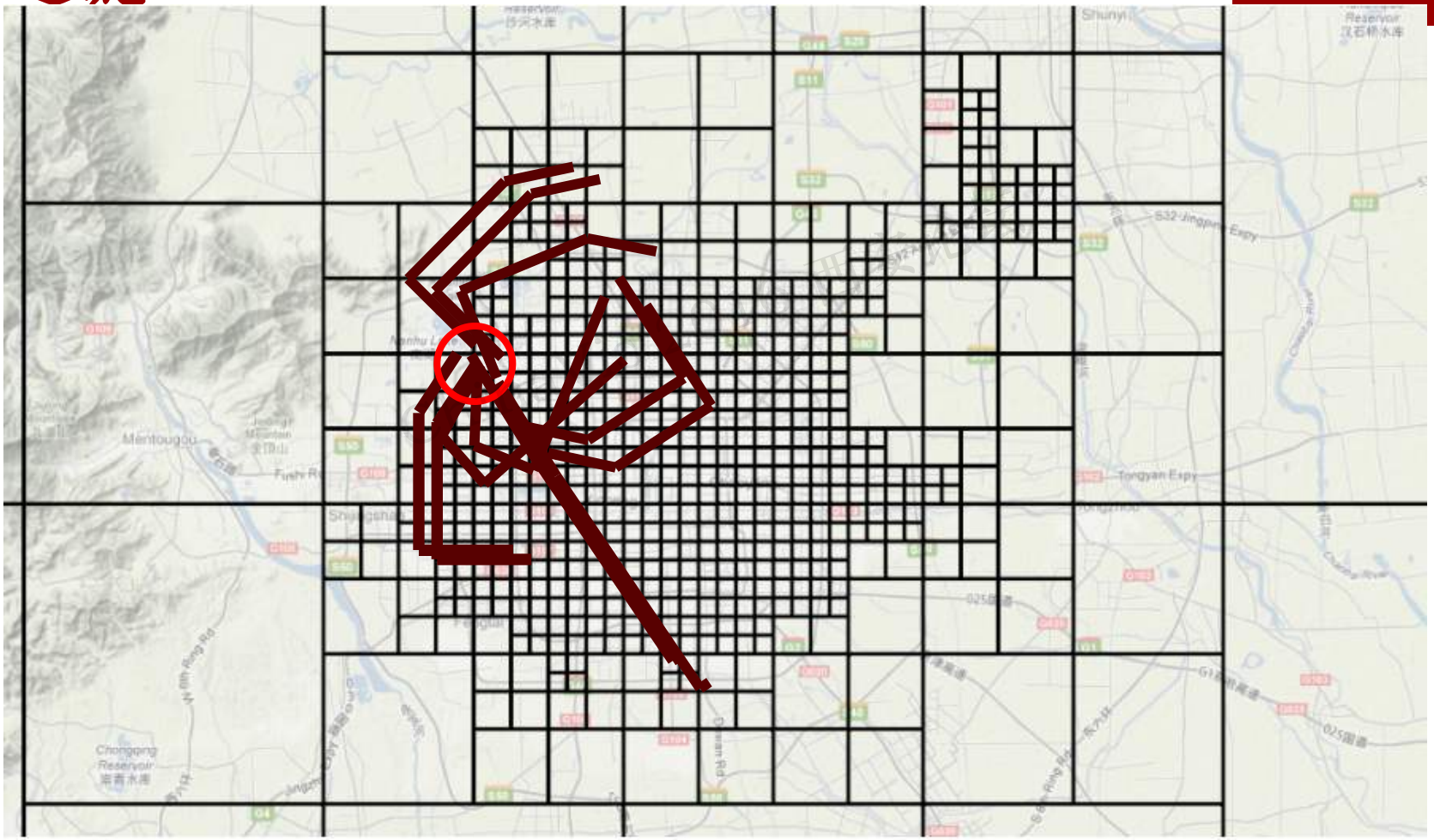
Trajn, <xn1,yn1>, <xn2,yn2> …

…

过滤



过滤

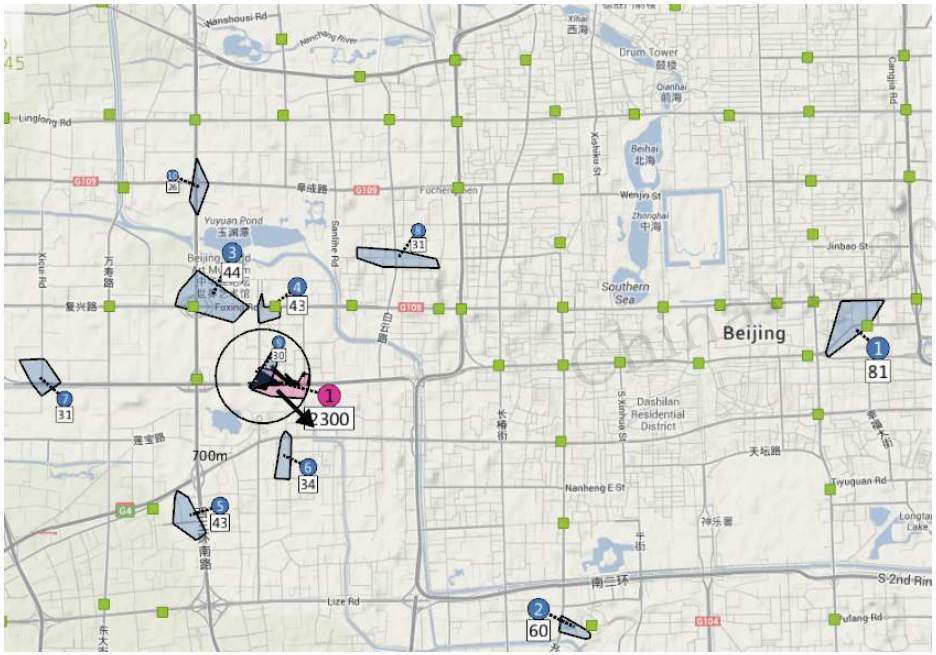

Shunyi
Reservoir
汉石桥水库
S32 Jingpin Expy
Mentougou
Tongyan Expy
Chaobai River
Chongqing
Reservoir
南青水库
N 6th Ring Rd
S 6th Ring Rd
Qiuyan Rd

流量统计

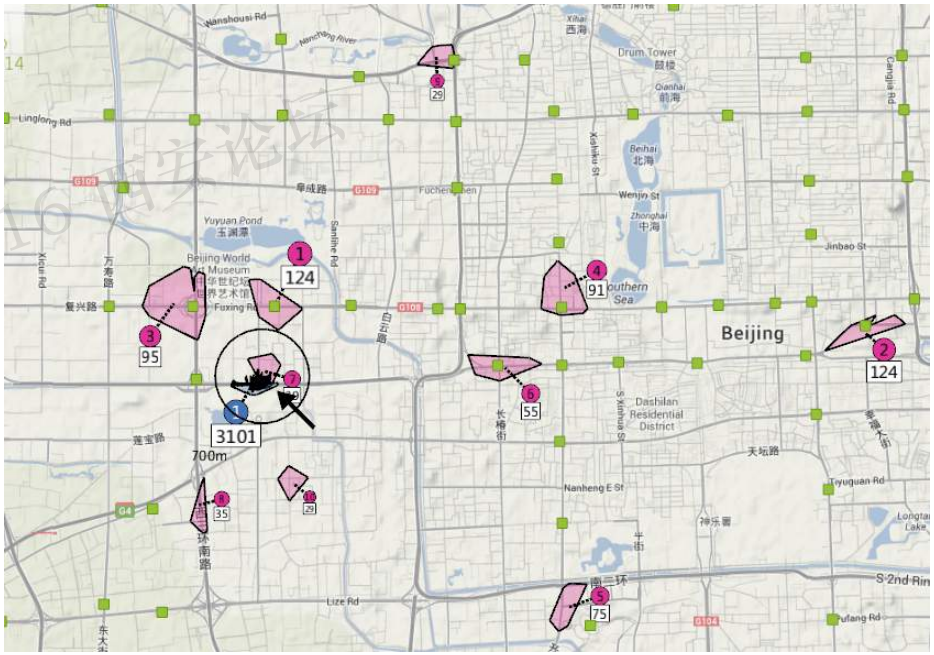
到达 #All Traj:786

出发 #All Traj:861

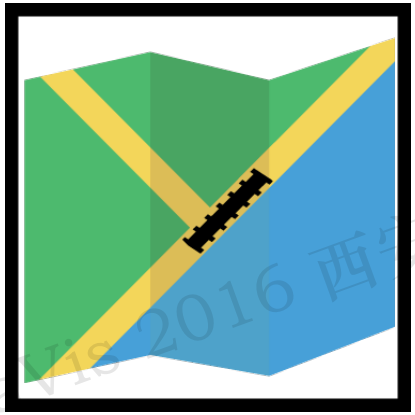
源-目的地



从北京西站出发的轨迹

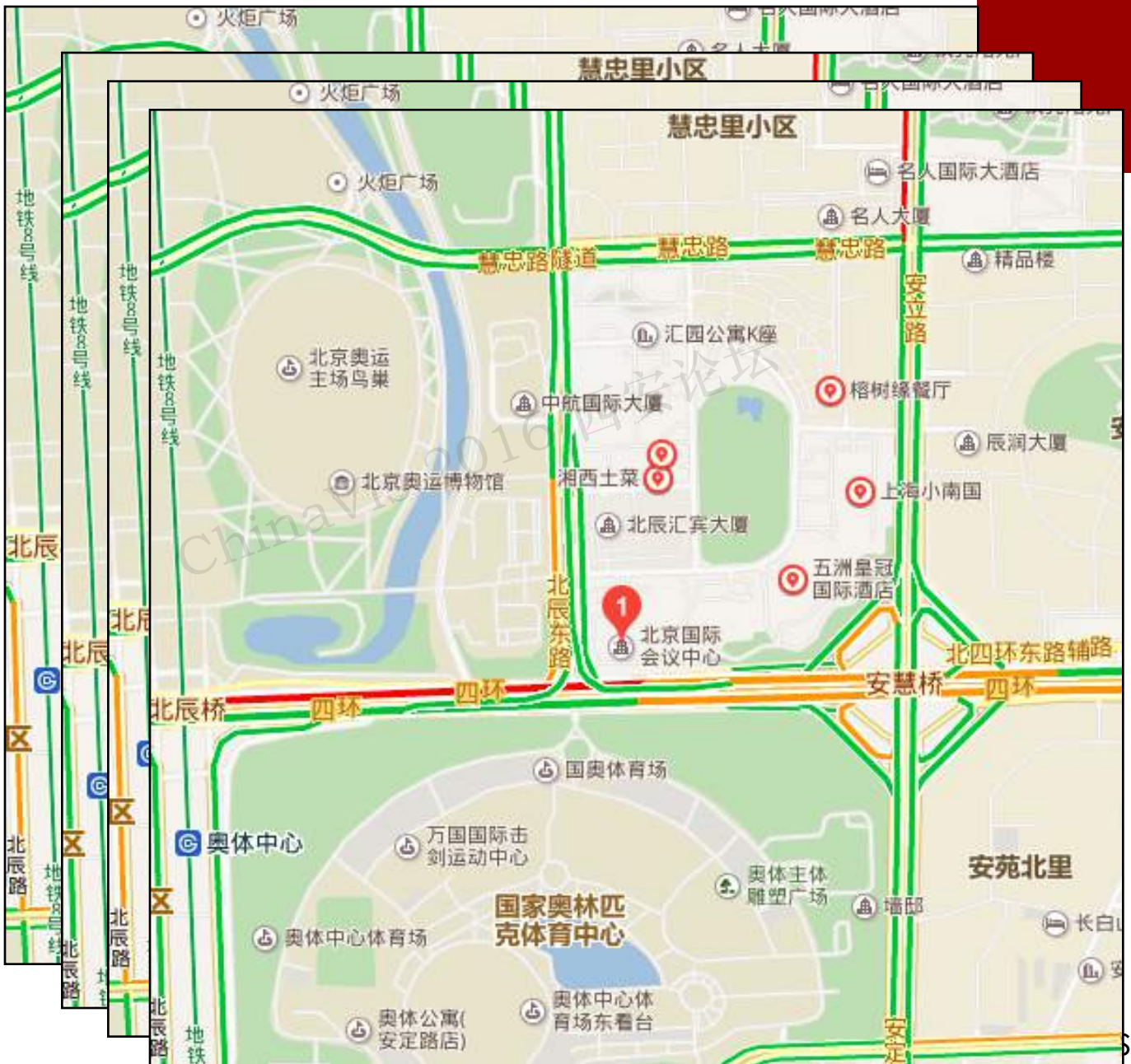


达到北京西站的轨迹



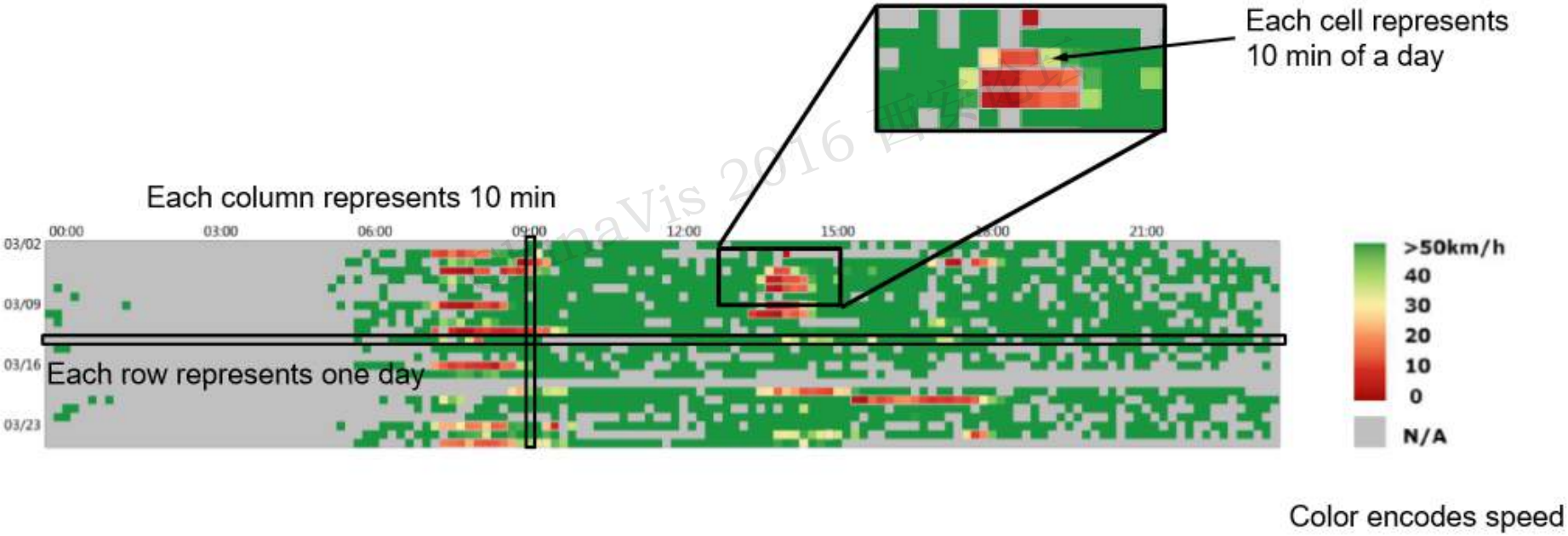
ChinaVis 2016 西安论坛

知微探秘：路段

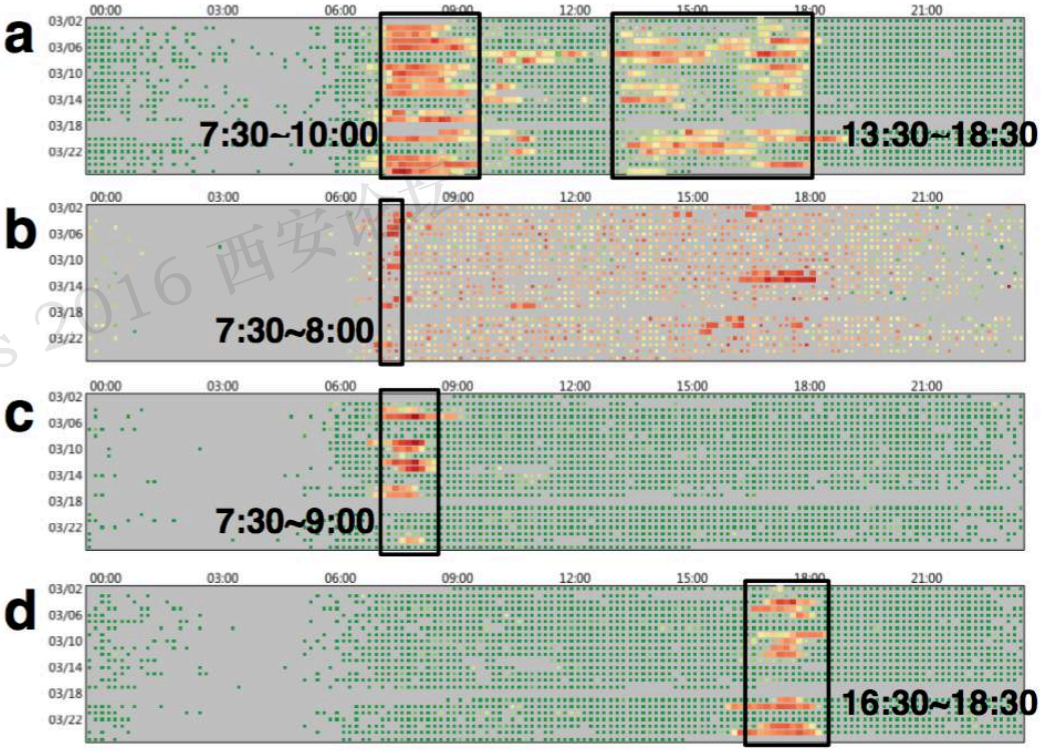


ChinaVis 2016西安论坛

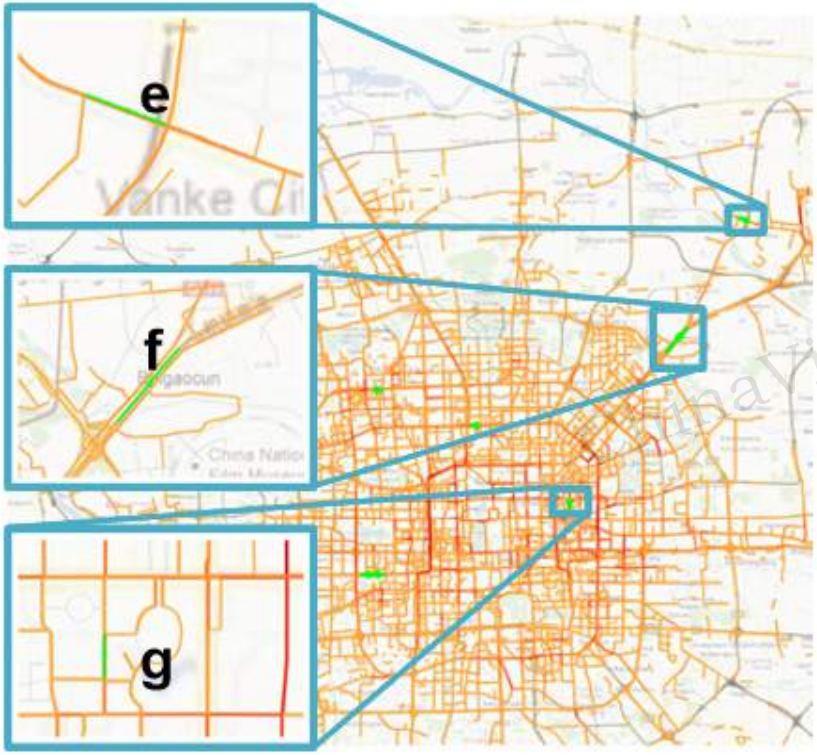
路段速度可视化

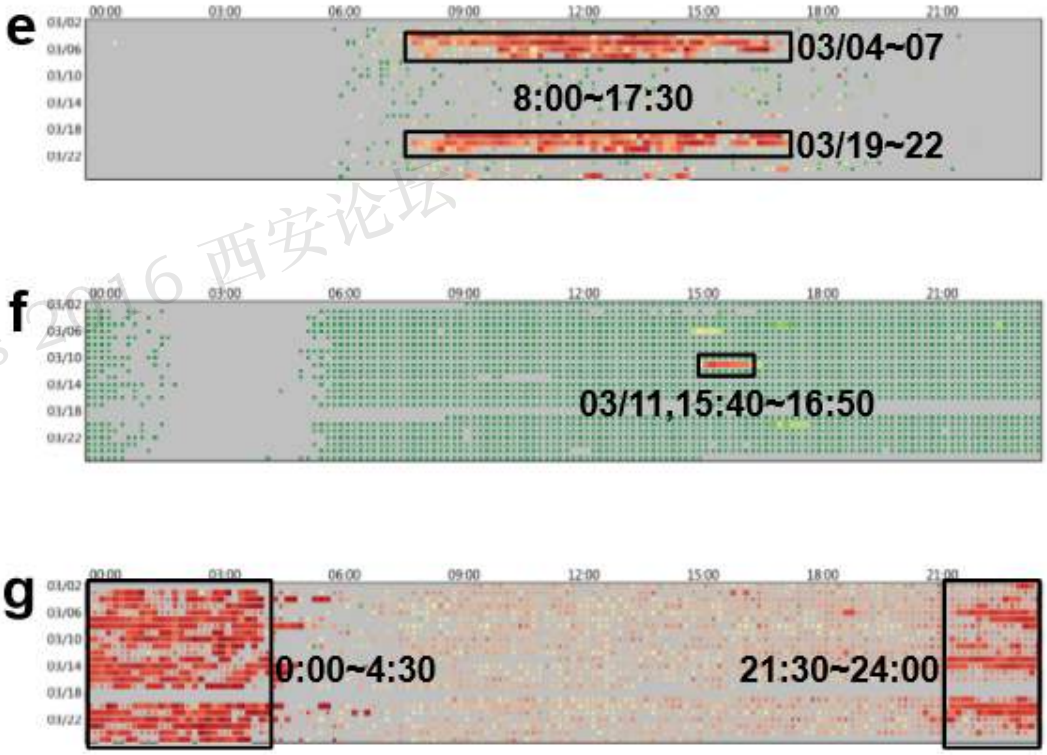


ChinaVis 2016 西安



ChinaVis 2016 西安论坛

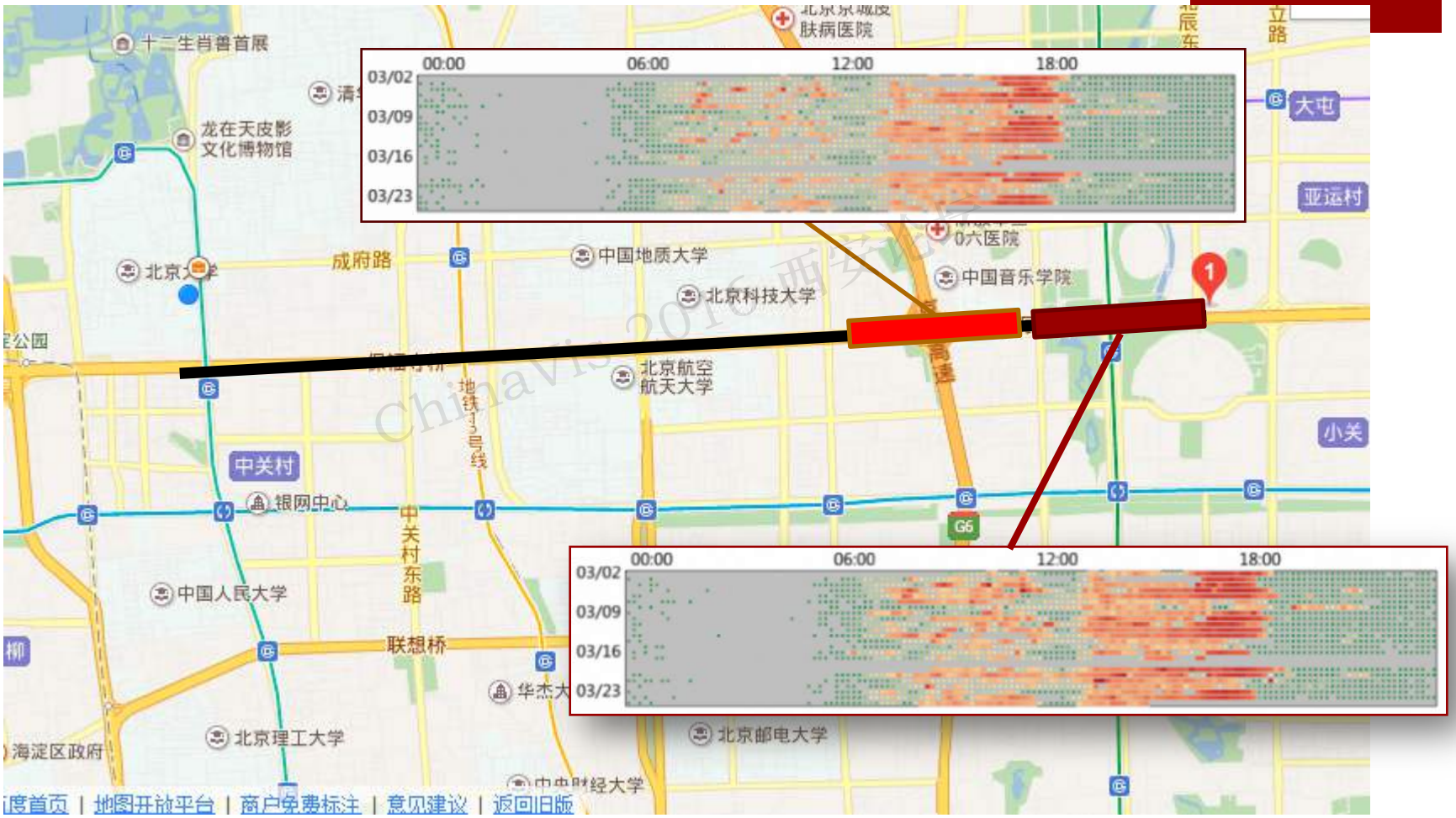




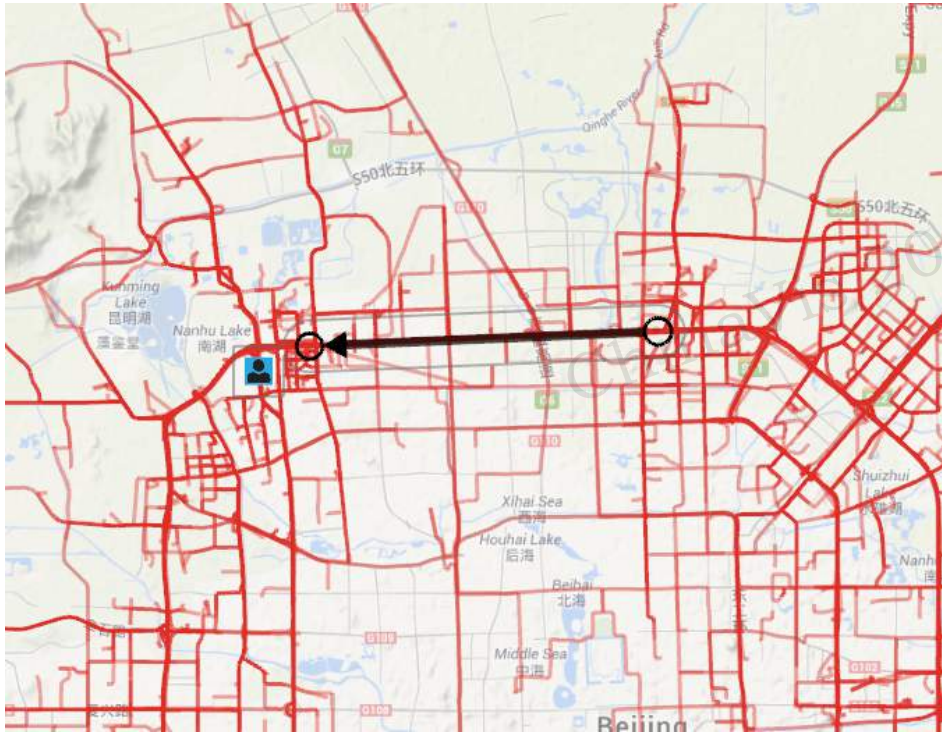
ChinaVis 2016 西安论坛

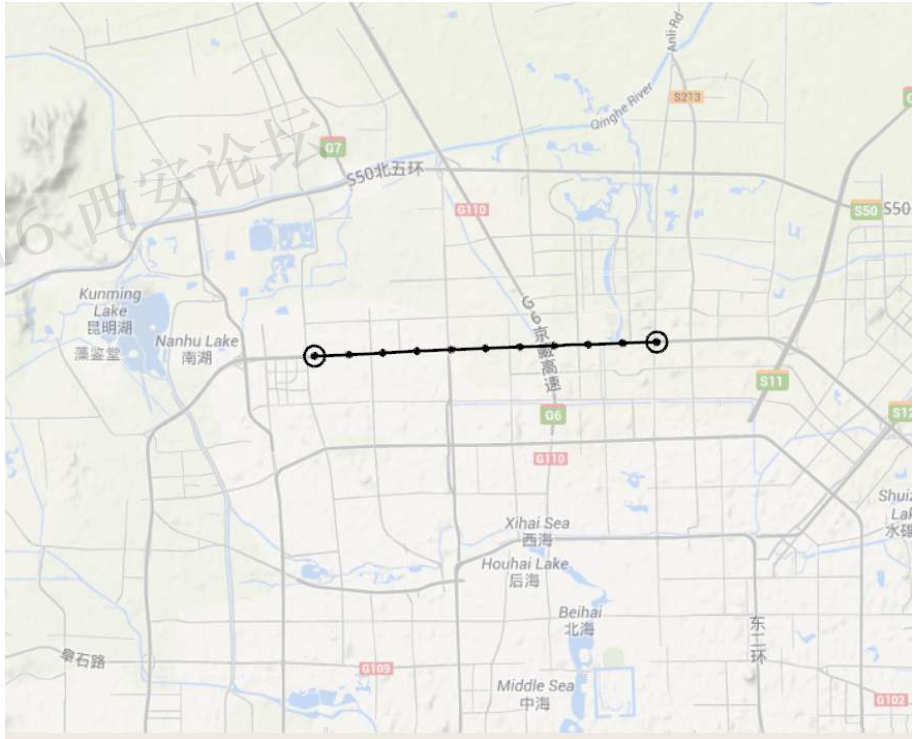
ChinaVis 2016 西安论坛

知微探秘：路径



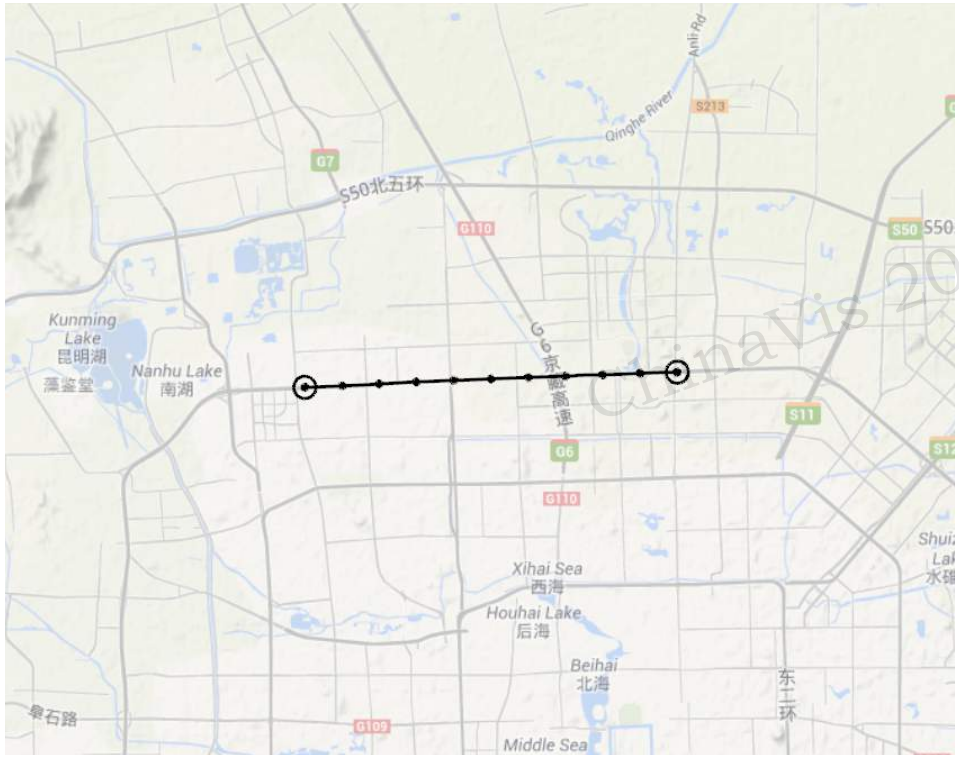
ChinaVis 2016 西安论坛

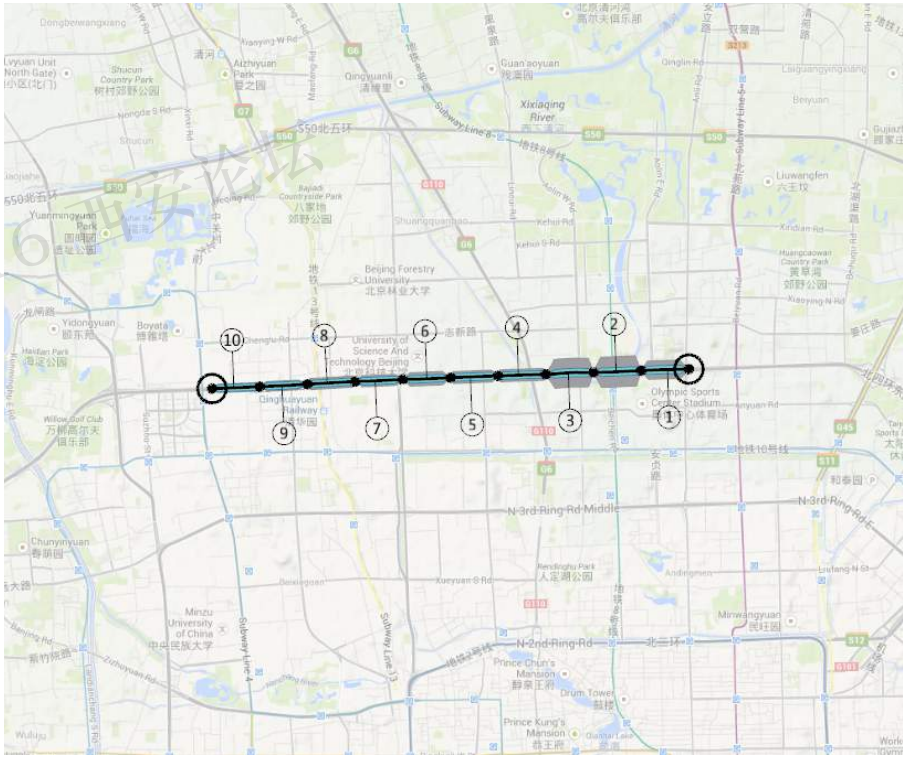

S50北五环
Kunming Lake
昆明湖
Nanhu Lake
南湖
Xihai Sea
西海
Houhai Lake
后海
Beihai
北海
Middle Sea
中海
Shuizhui Lake
Beijing


S50北五环
Kunming Lake
昆明湖
Nanhu Lake
南湖
Xihai Sea
西海
Houhai Lake
后海
Beihai
北海
Middle Sea
中海
京藏高速
东三环

2016 西安论坛

路径通行时间

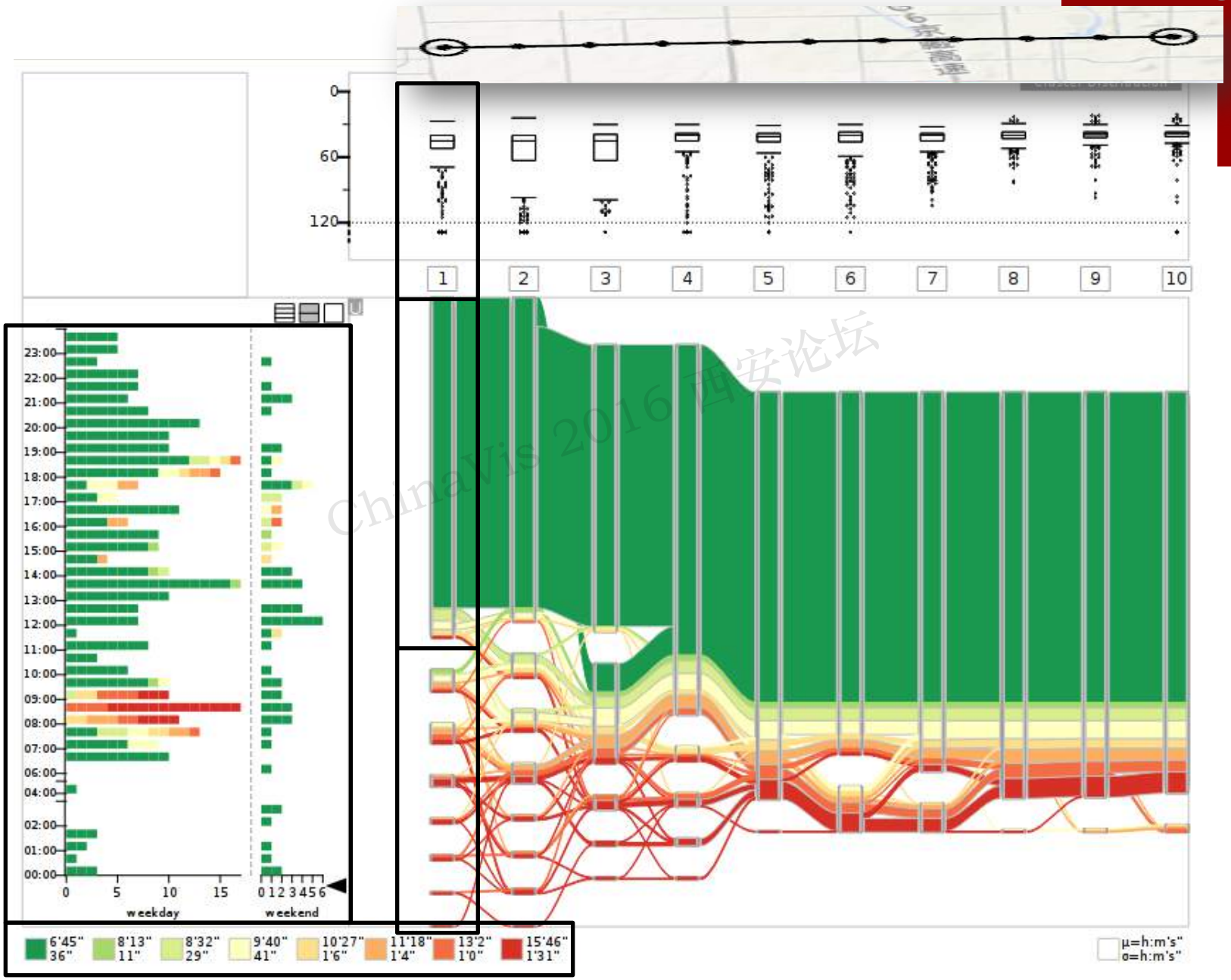




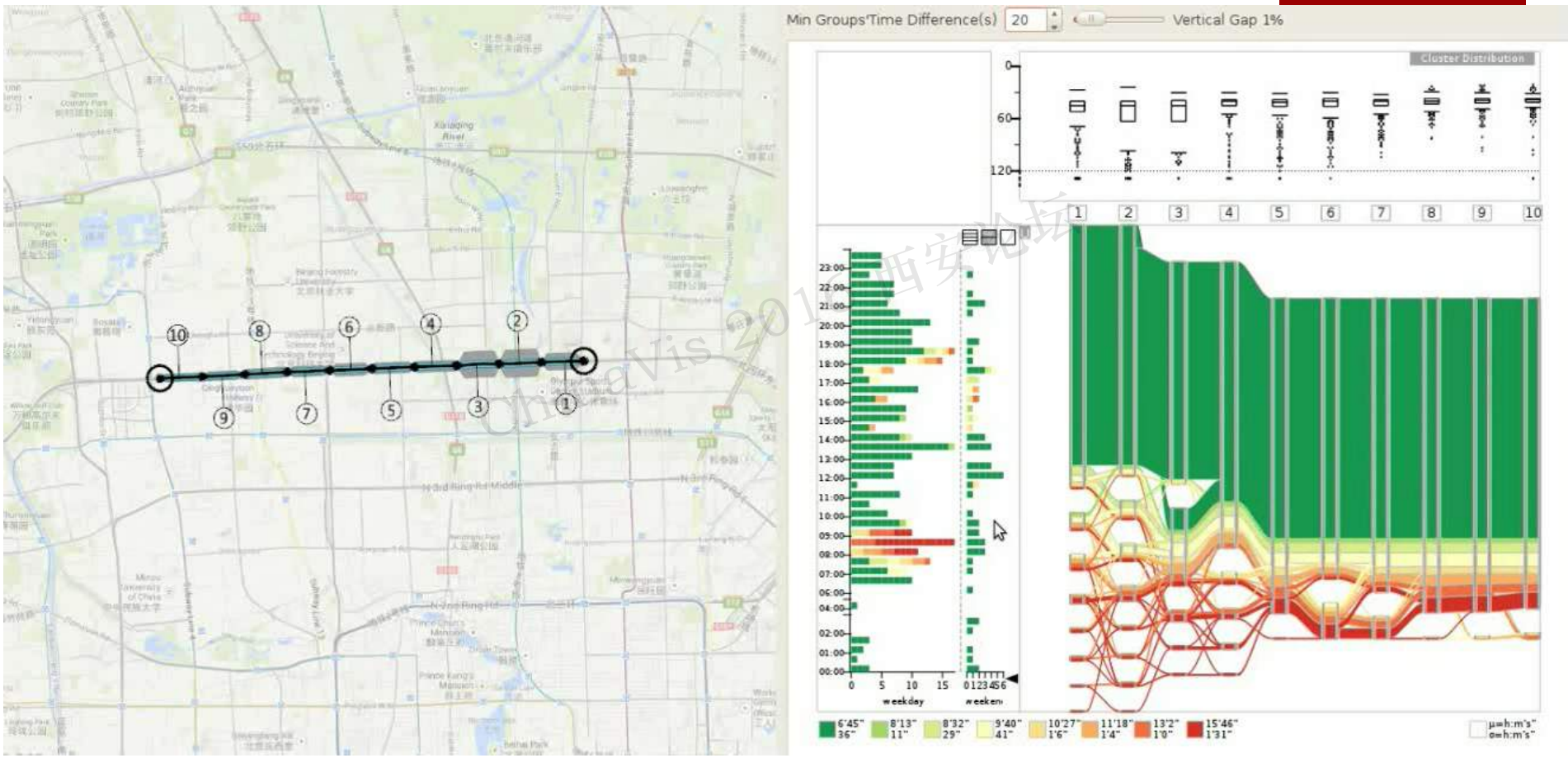
ChinaVis 2016 北京论坛

通行时间的方差
通行时间的平均值
1
2
3
4
5
6
7
8
9
10
Beijing Forestry University
北京林业大学
Olympic Sports Center Stadium
N-3rd-Ring-Rd-Middle
Yuanmingyuan Park
Yidongyuan
Boyata
Willow Golf Club
Chunyinyuan
Qinghuayuan Railway
University of Science And Technology Beijing

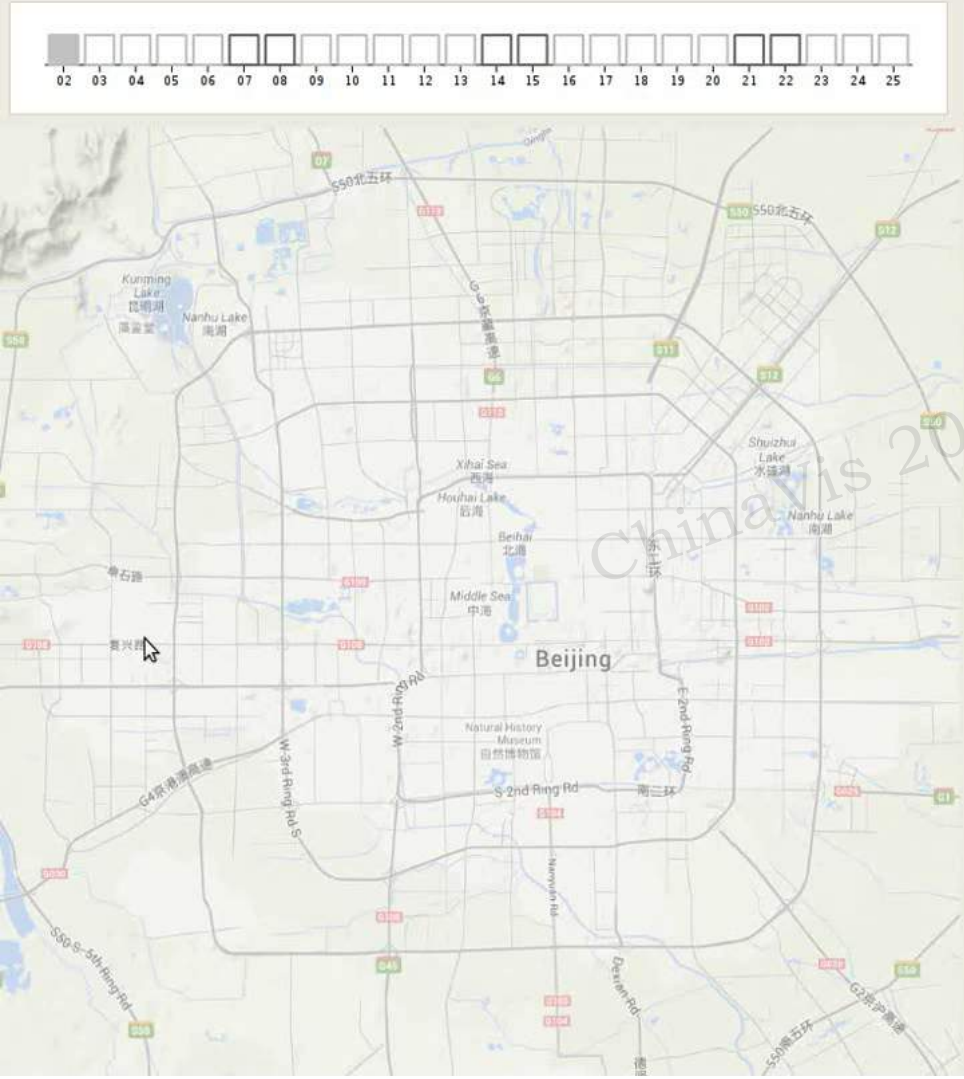
2016 厦门论坛



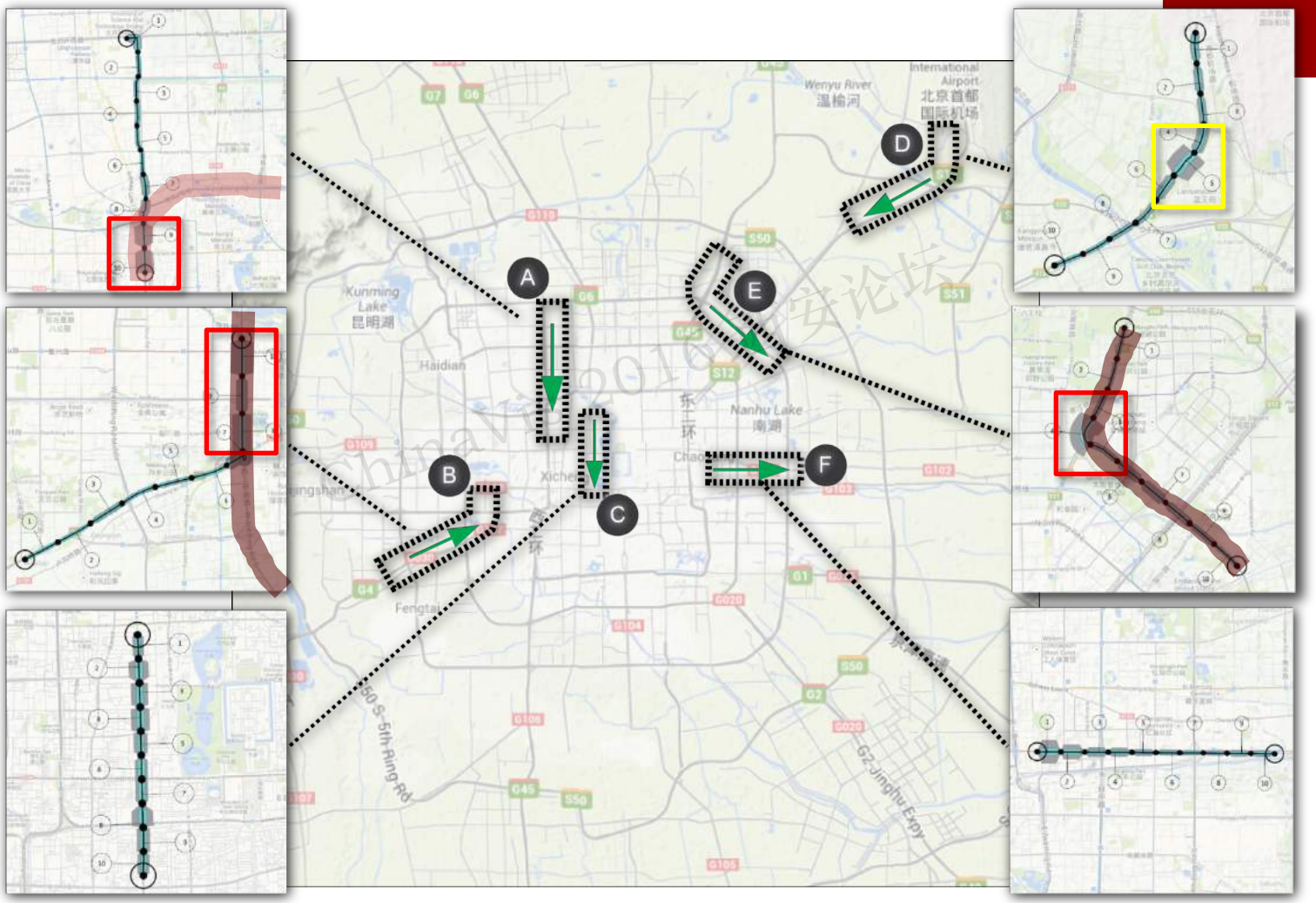
ChinaVis 2016 西安论坛

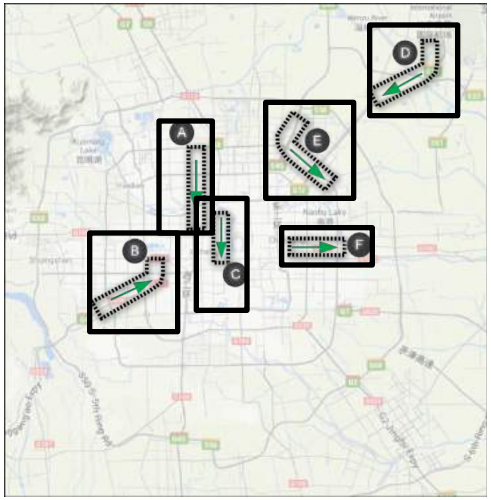


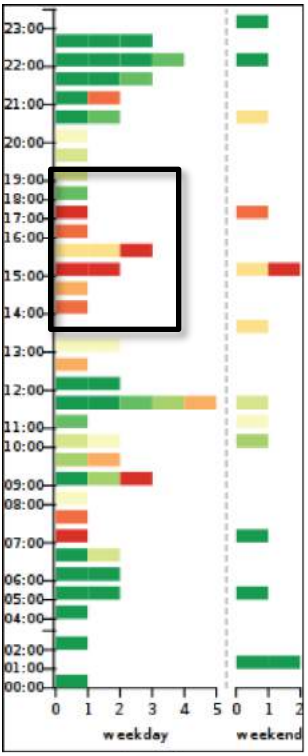
ChinaVis 2016 西安论坛

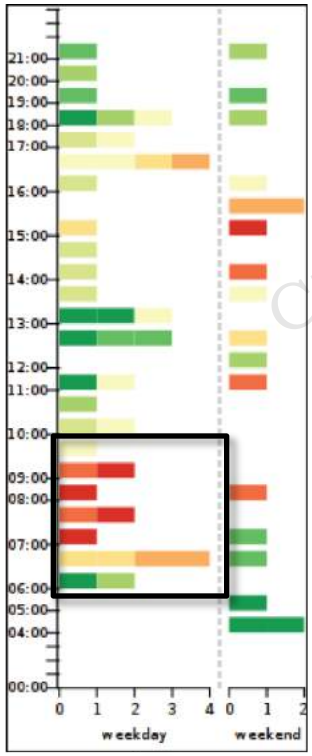


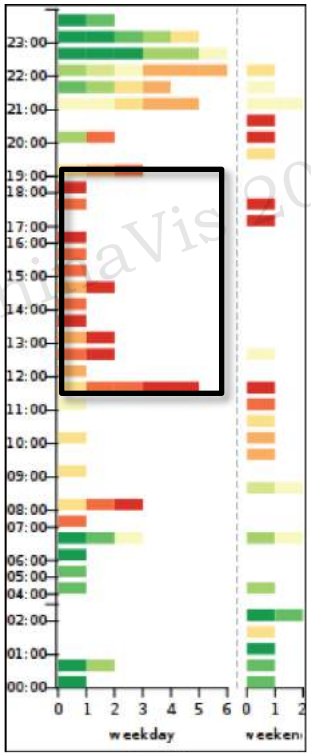
ChinaVis 2016 西安论坛

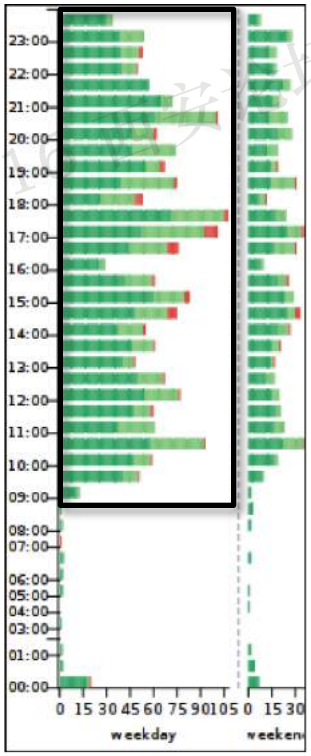

A
B
C
D
E
F
Wenyu River
温榆河
International Airport
北京首都国际机场
Kunming Lake
昆明湖
Haidian
Nanhu Lake
南湖
Fengtai
S-5th Ring Rd
G2-Jinghu Expy

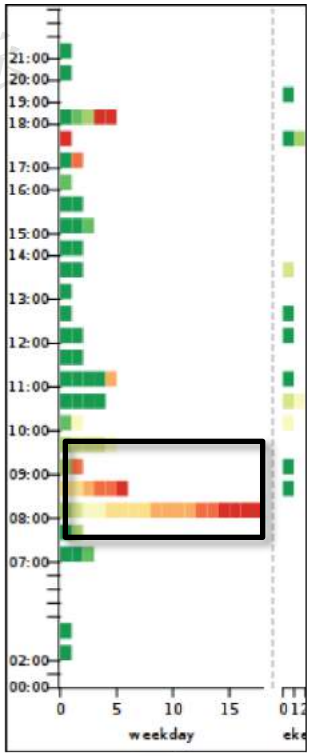


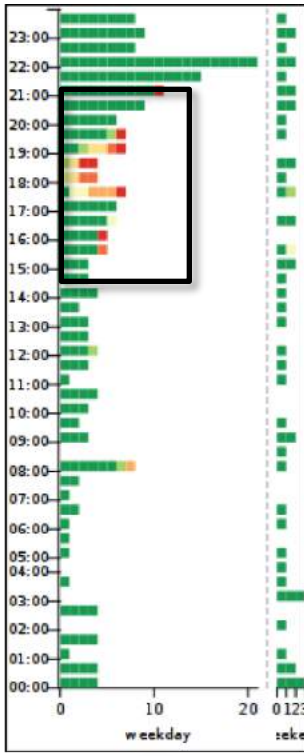








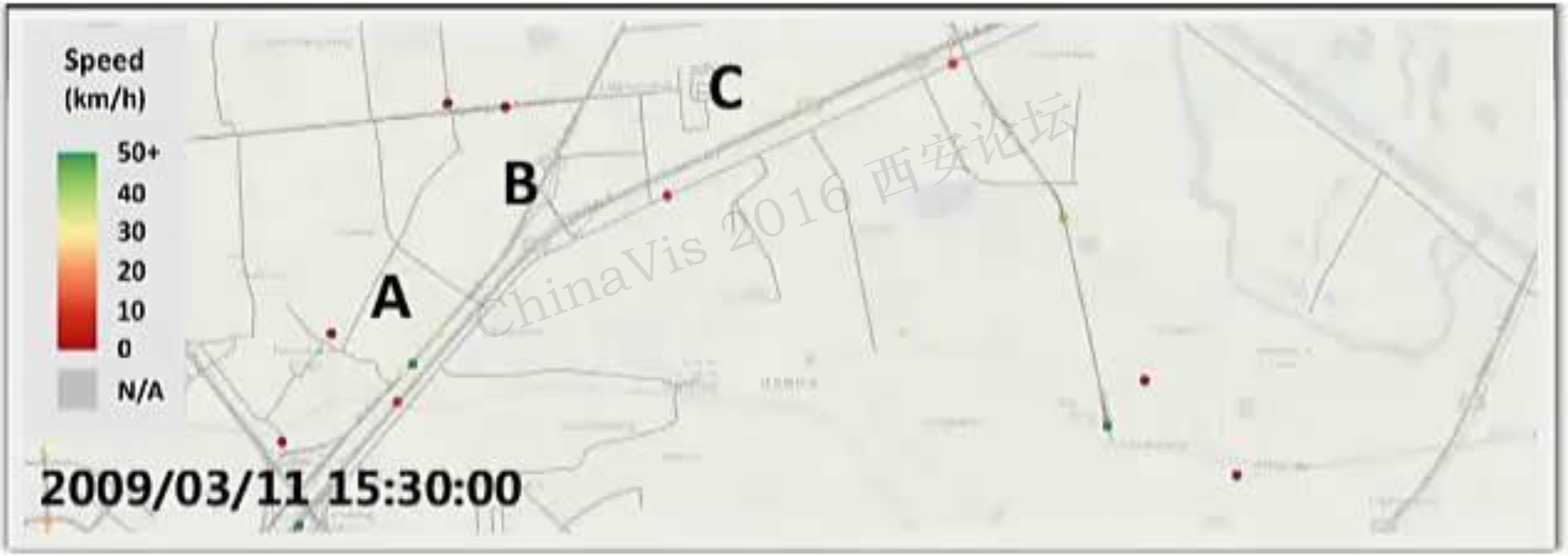




ChinaVis2016 西安

ChinaVis 2016 西安论坛

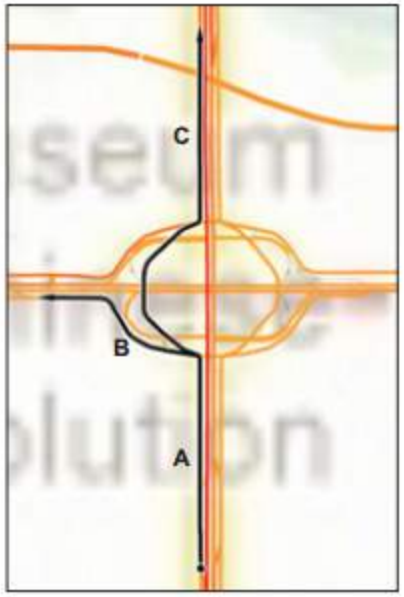
知微探秘：拥堵传播



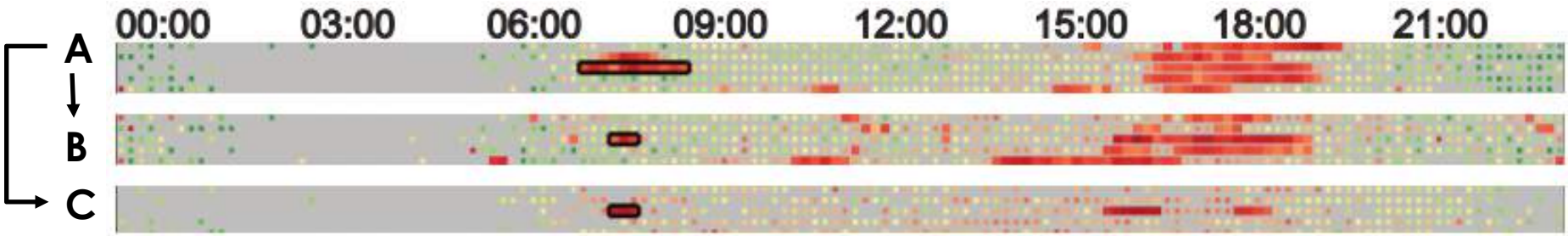
ChinaVis 2016 西安论坛

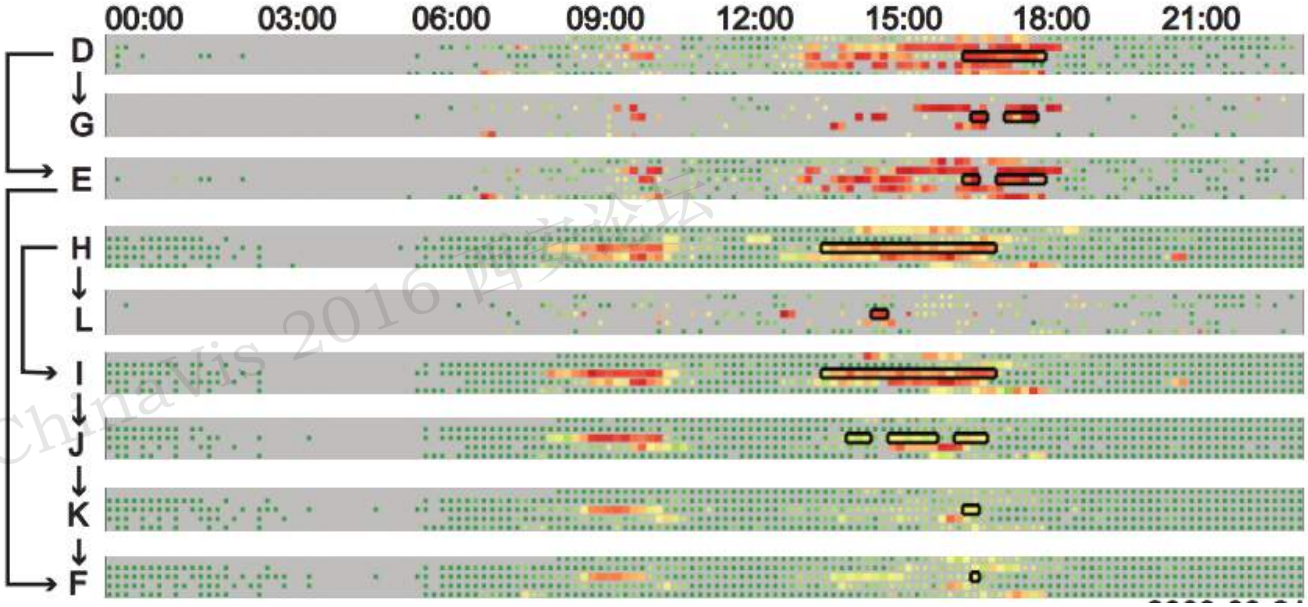
拥堵传播

- 空间相邻路径
- 路径拥堵事件
- 拥堵事件时间先后

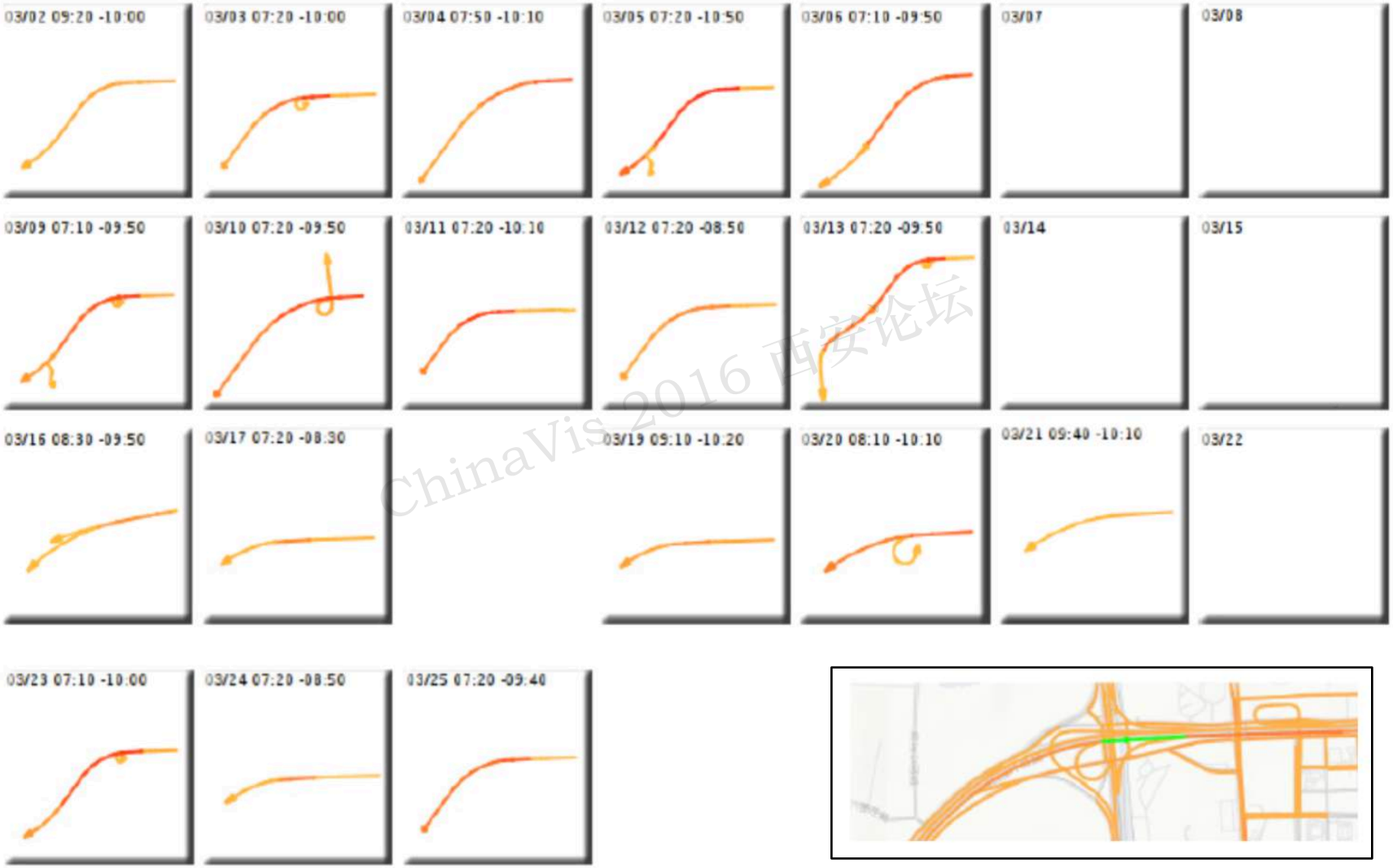


ChinaVis 2016 西安论坛





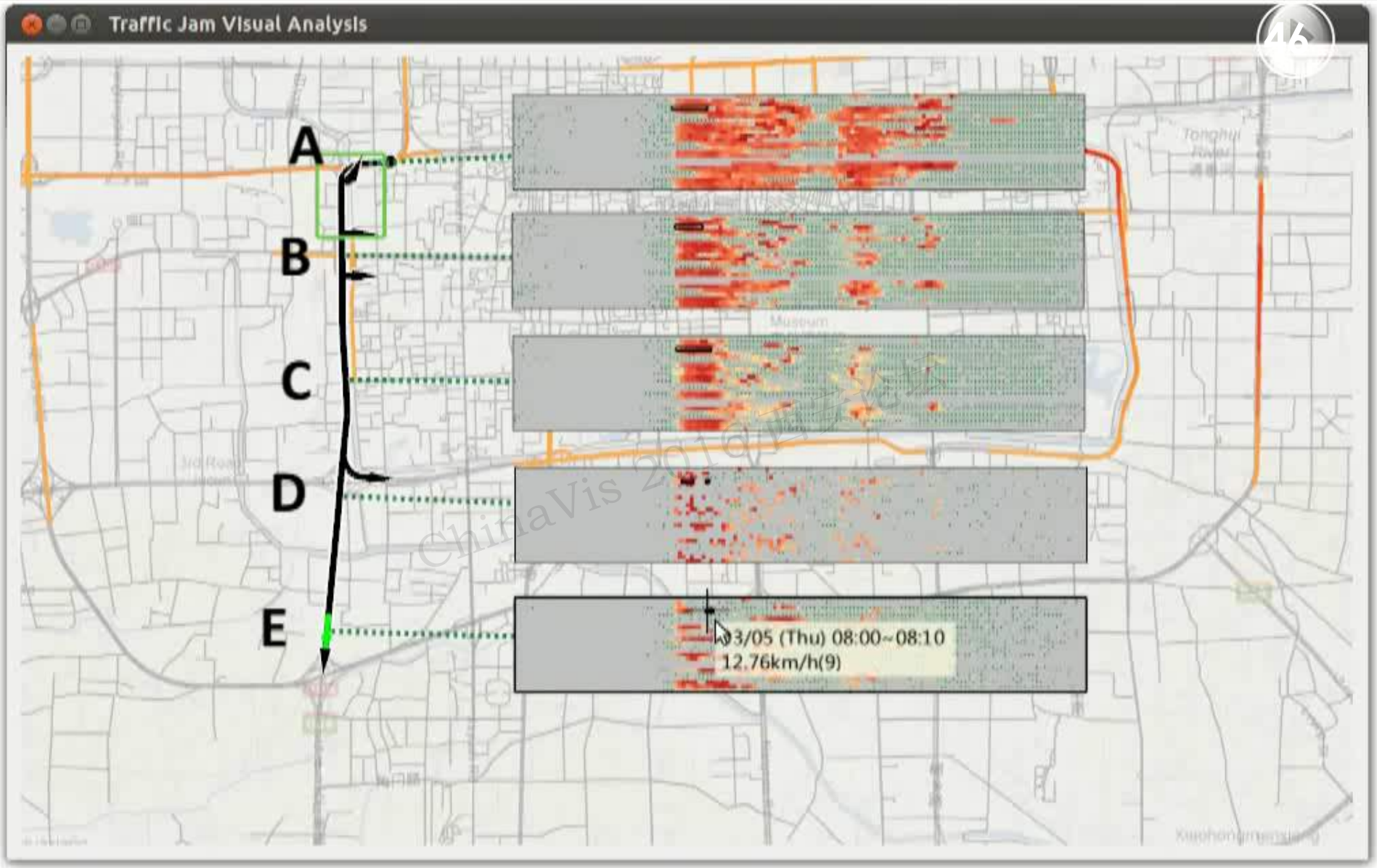
ChinaVis 2016 西安论坛



ChinaVis 2016 西安论坛

Temporal Filtering

ChinaVis 2016 西安论坛



Watching Taxi Trajectory Animation

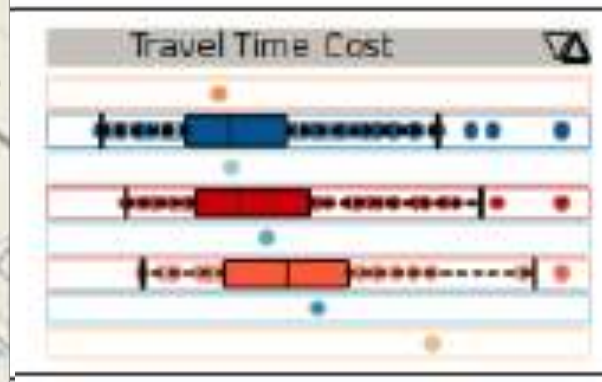
ChinaVis 2014 北京

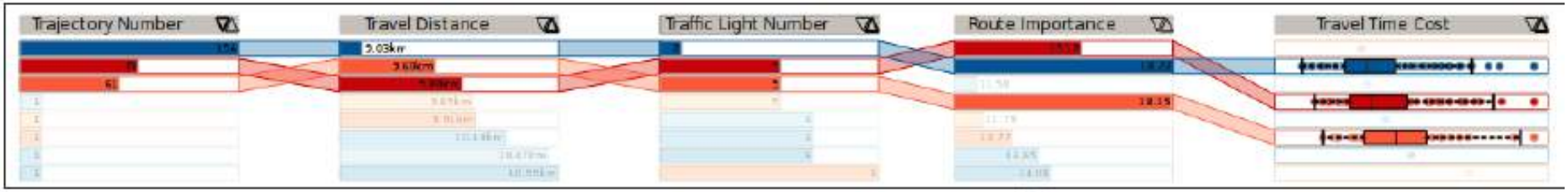
ChinaVis 2016 西安论坛

知微探秘：路径选择因果分析

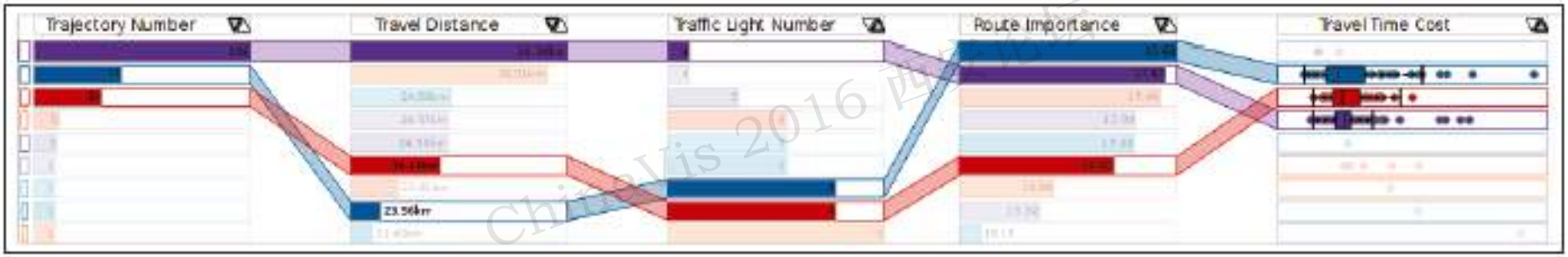
ChinaVis 2016 西安论坛

多条可选路径

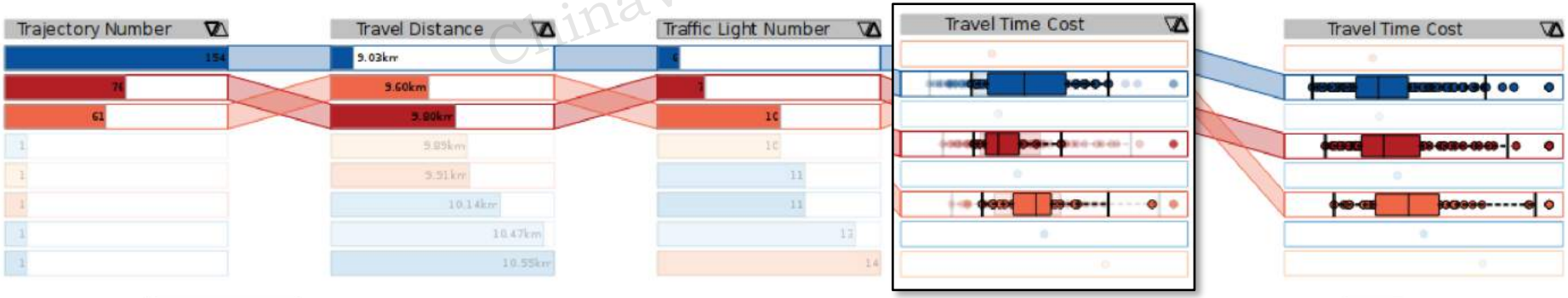


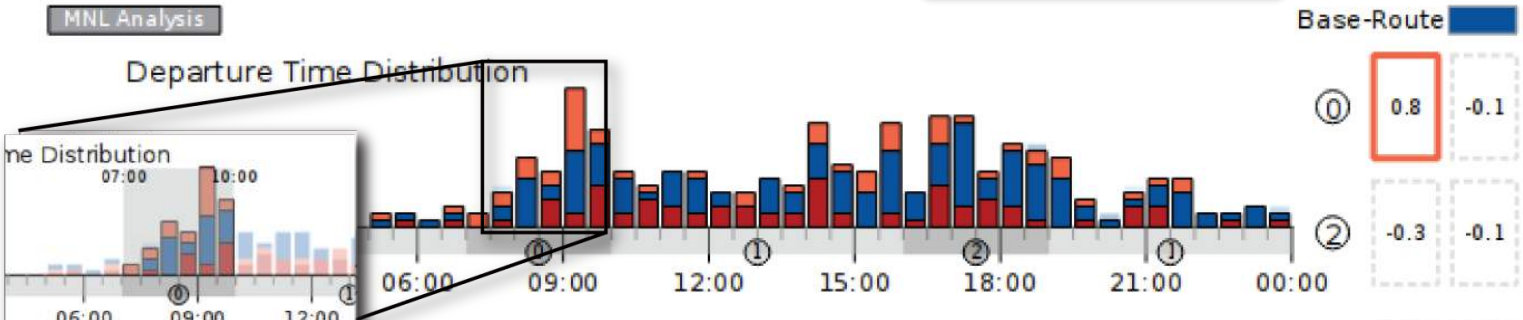


ChinaVis 2016 西安 演示



ChinaVis 2016 西安

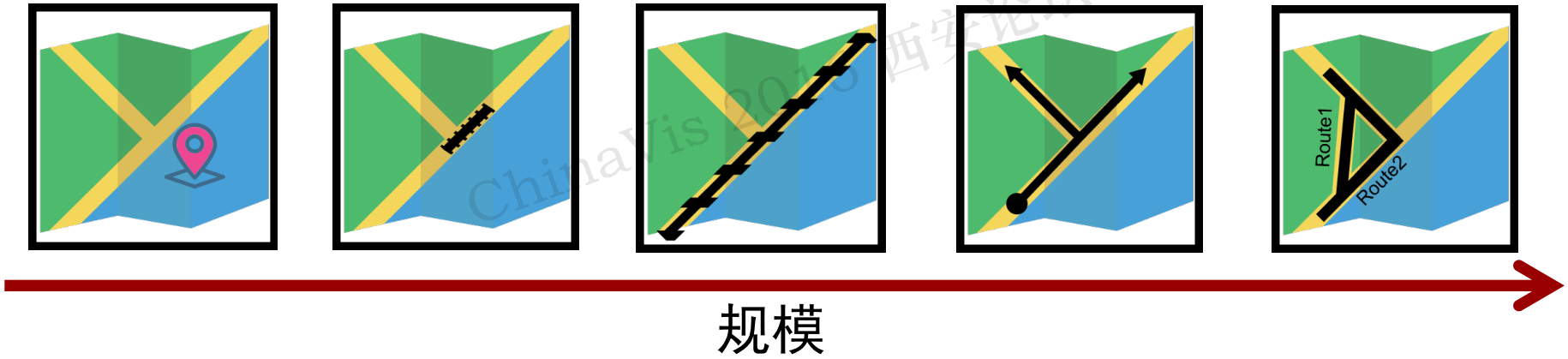




ChinaVis 2016 西安论坛

小结

- 城市移动数据知微探秘



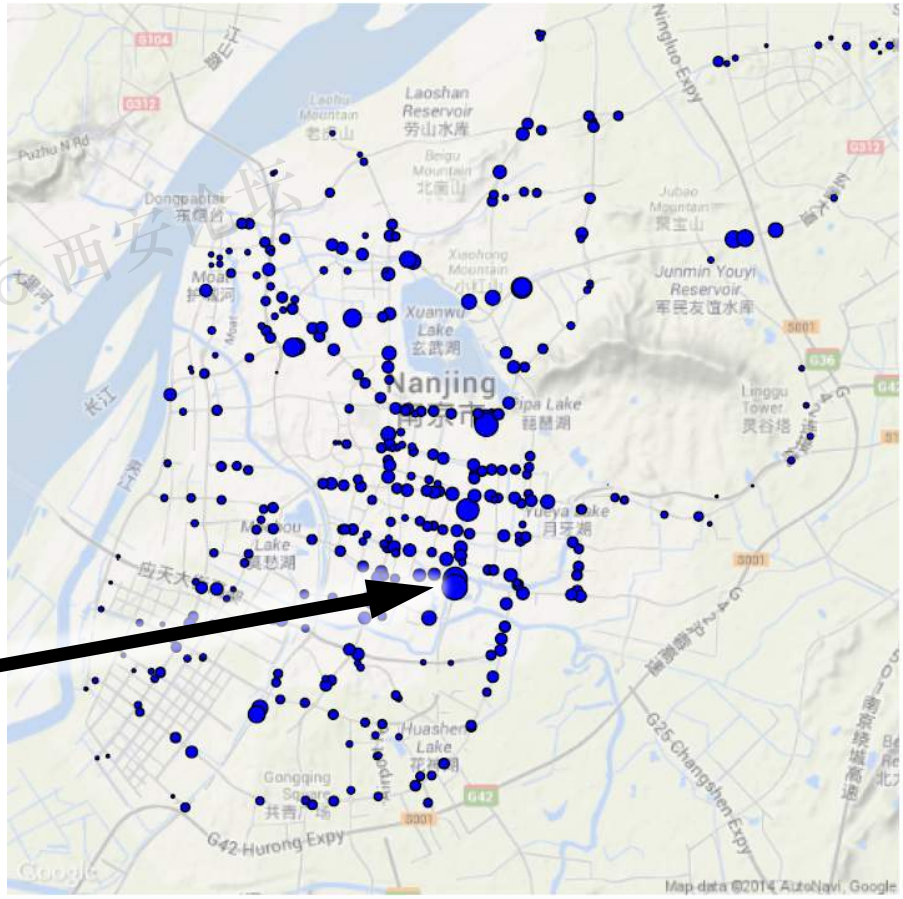
ChinaVis 2015 西安论坛

城市RFID数据

[Wang et al. IEEE TVCG 2014] (VAST)

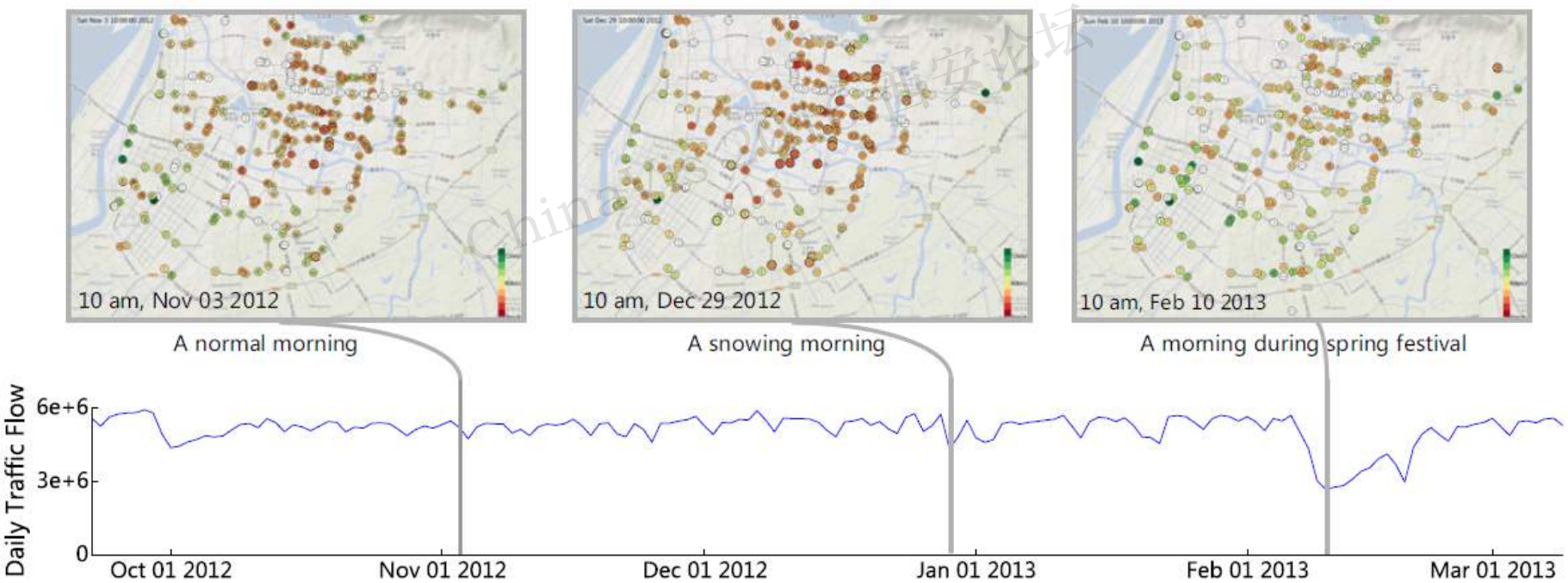
- Video detector data
  - 472 transportation cells in Nanjing
  - Over 1 million vehicles
  - 870 million vehicle passing records
  - 169 consecutive days
  - 39 GB data size

ChinaVis 2016 西安论坛



城市RFID数据

- Normal traffic　　Slower traffic (redder cells)　　Faster traffic (greener cells)



Traffic flow is rather stable during this 169 days, except for during Spring Festival

Data Characteristic

		Our cell-based sparse trajectory	Traditional GPS trajectory
Data sampling	Location	Sparse (only at cells)	Quasi-continuous (can be anywhere)
	Population	Dense (almost all vehicles)	Subset (usually one type of vehicles)
Derived properties	Flow volume	Available	Not available
	Traffic speed	Accurate	Can be biased
Analysis task	Network analysis	Accurate	Can be biased
	Vehicle tracking	Low resolution (can be simulated locally)	Accurate (require map matching)

ChinaVis 2016 西安论坛

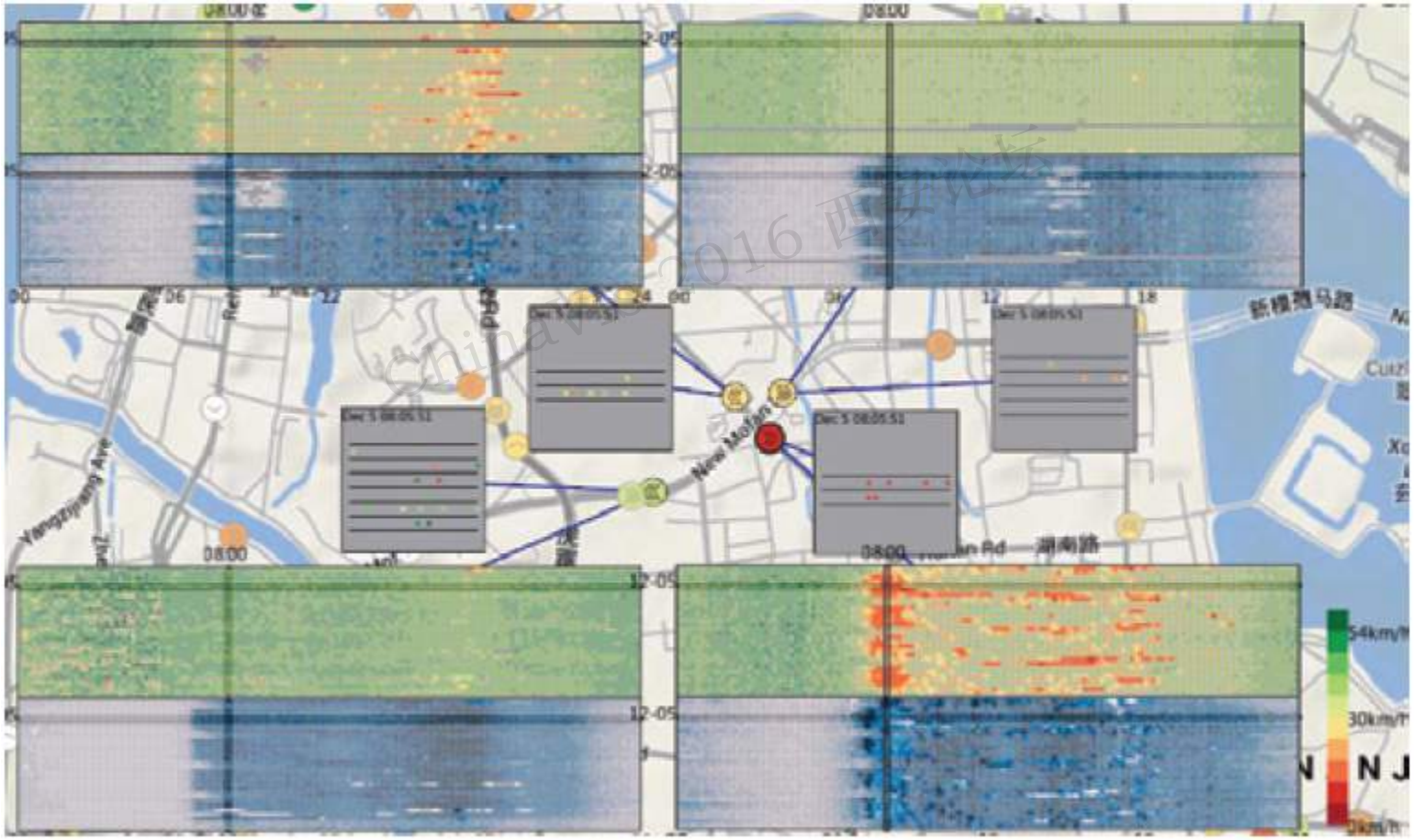
Network Exploration

A: railway station
B: inter-city express
C: commercial cente
D: High-tech zone

Based on data in Dec. 1st, 2012

ChinaVis 2016 西安论坛

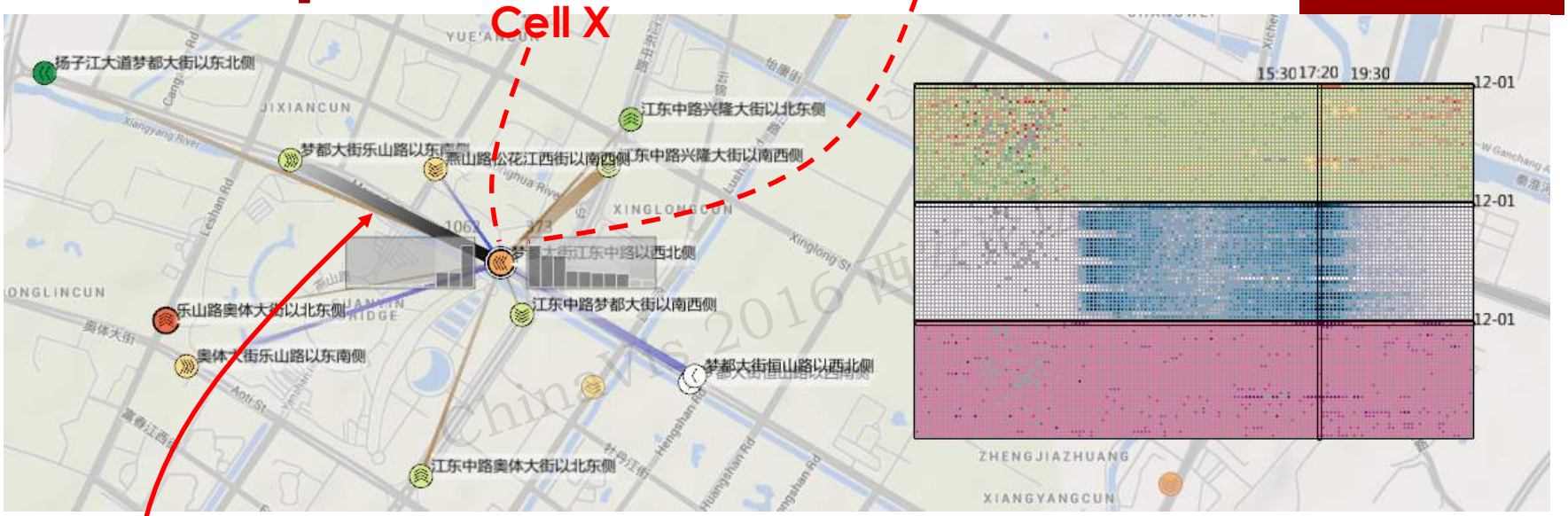
Cell Exploration



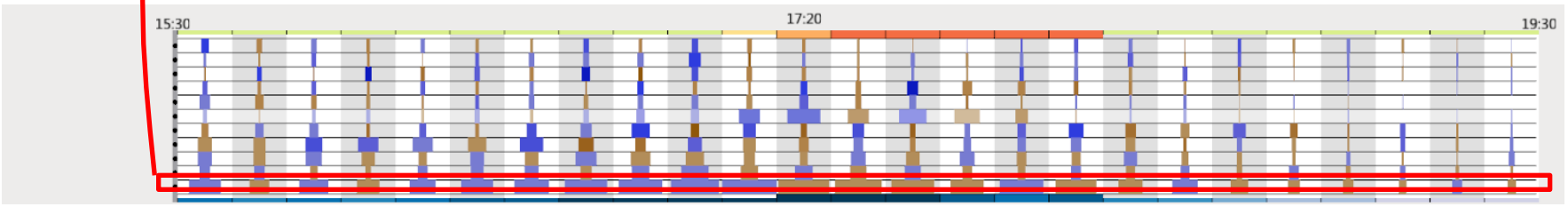
chinavis2016 西安论坛

Link Exploration

Nanjing Olympic Center
9am-5pm: Car Exhibition
6pm-?: Voice of China

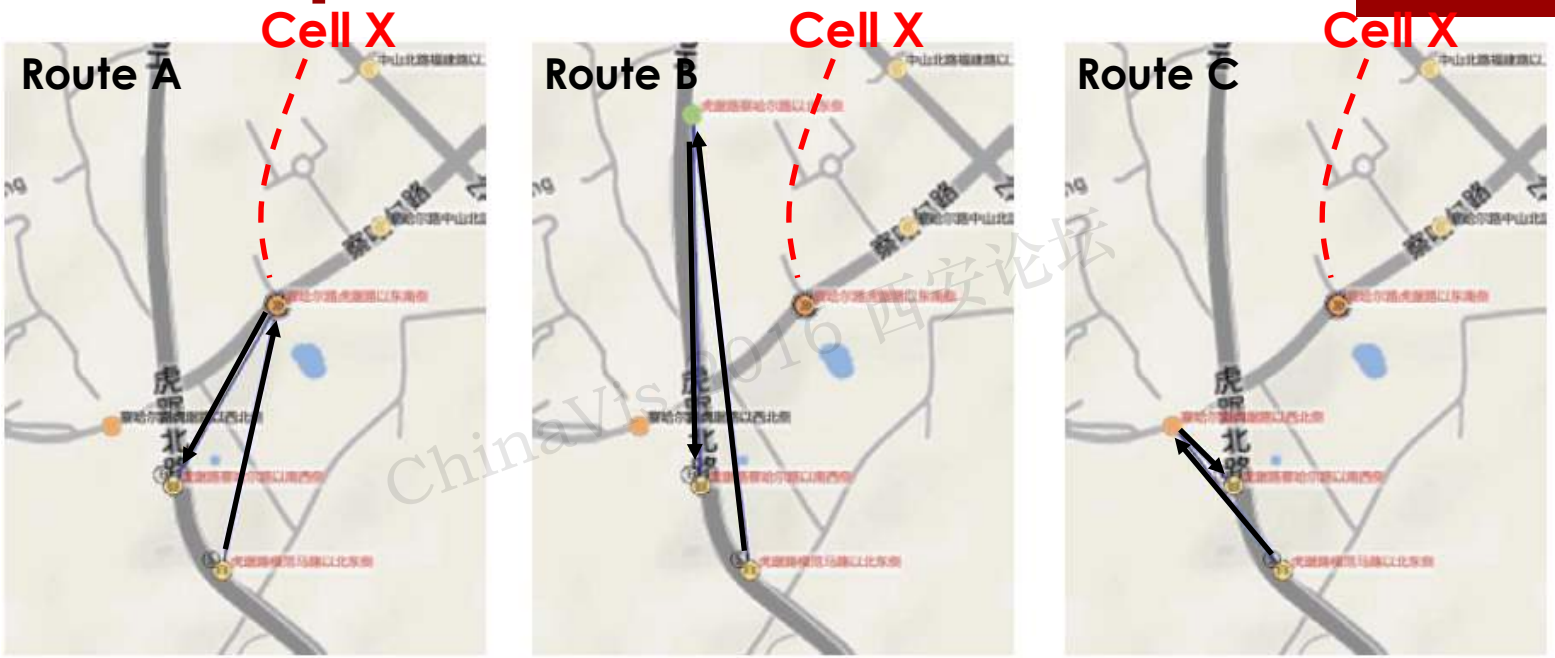


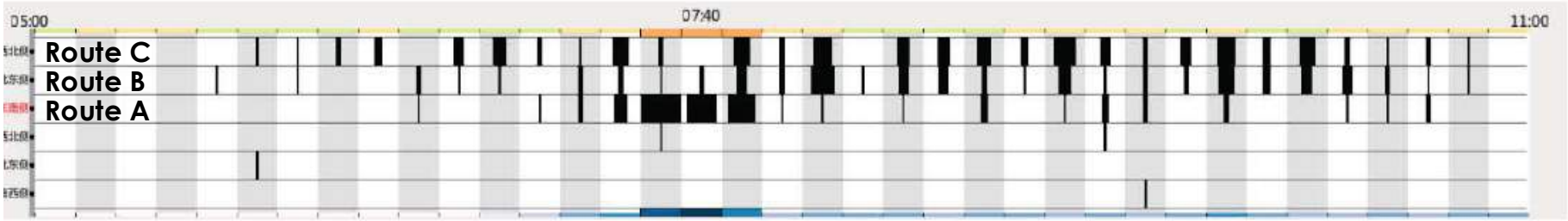
This turn-around link behaves differently during congestion on cell X



Congestion on cell X

Route Exploration





Route A behaves differently during congestion on Cell X

Congestion on Cell X

ChinaVis 2016 西安论坛

可视化—发展趋势

- 数据对象从传统的单一数据来源扩展到多来源，多维度，多尺度等广泛的数据；
- 用户从少数专家用户扩展到广泛的不特定的群体；
- 研究更为强调方法的可扩展性和开发的简捷性；
- 革命性的新方法和原有普遍适用的可视化技术和工具结合；
- 面向领域和大众的可视化工具库和协同与众包的可视分析方式将被更广泛地引入到各种综合分析工具和平台，大幅度提升各类大数据分析的易用性和有效性；
- 沉浸式可视分析将在今后3-5年涌现 （New)
- 可视化即为服务，Visualization As Service

s

ChinaVis 2016 西安论坛

致谢

-

China 2015 西安论坛

北京大学可视化与可视分析研究组

VisPKU

ChinaVis 2015 西安论坛

实验室主页 http://vis.pku.edu.cn/

实验室博客 http://vis.pku.edu.cn/blog

空间数据可视分析网站 http://vis.pku.edu.cn/trajectoryvis/en/index.h

公众号“可视分析”