北京城市实验室
Beijing City Lab

ID of the slides

24

Slides of BCL

www.beijingcitylab.com

How to cite

Author(s), Year, Title, Slides at Beijing City Lab, http://www.beijingcitylab.com

E.g. Long Y, 2014, Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest, Slides at Beijing City Lab, http://www.beijingcitylab.com

利用公共交通刷卡数据
分析北京通勤出行和城市贫困

龙　瀛，清华大学城市规划工学博士

北京市城市规划设计研究院，高级工程师

北京城市实验室，主任

![BCL logo]

BCL
Beijing City Lab

- Home
- Projects
- Members
- Working papers
- Slides
- Data released
- Blogs
- About

The Beijing City Lab (BCL) is a virtual research community, dedicated to studying, but not limited to, China's capital Beijing. The Lab focuses on employing interdisciplinary methods to quantify urban dynamics, generating new insights for urban planning and governance, and ultimately producing the science of cities required for sustainable urban development. The lab's current mix of planners, architects, geographers, economists, and policy analysts lends unique research strength.

Urban Growth Boundaries of 176 Chinese Cities

Full process hand made from planning drawing collection to digitization. Courtesy of students in Zhejiang University. longying1980@gmail.com


Create your own custom maps with CartoDB

www.beijingcitylab.org OR longy.jimdo.com

About | Privacy Policy | 🖶 Print Version | Sitemap |
✉ Recommend this site!
longying1980@gmail.com

Login

Beijing City Lab

2013年10月，龙瀛博士发起北京城市实验室（Beijing City Lab，BCL），BCL专注于运用跨学科方法量化城市发展动态，开展城市科学研究。BCL是中国第一个开放的定量城市研究网络，通过邀请学者发布其工作论文等形式阐释其对城市研究的最新见解，通过数据分享行为为科研群体提供开放的城市定量研究数据。

中国蓝色之都：厦门

Polluted months

- 0-1
- 1-2
- 2-3
- 3-4
- 4-5
- 5-6
- 6-7
- 7-8
- 8-9
- 9-10
- Water bodies
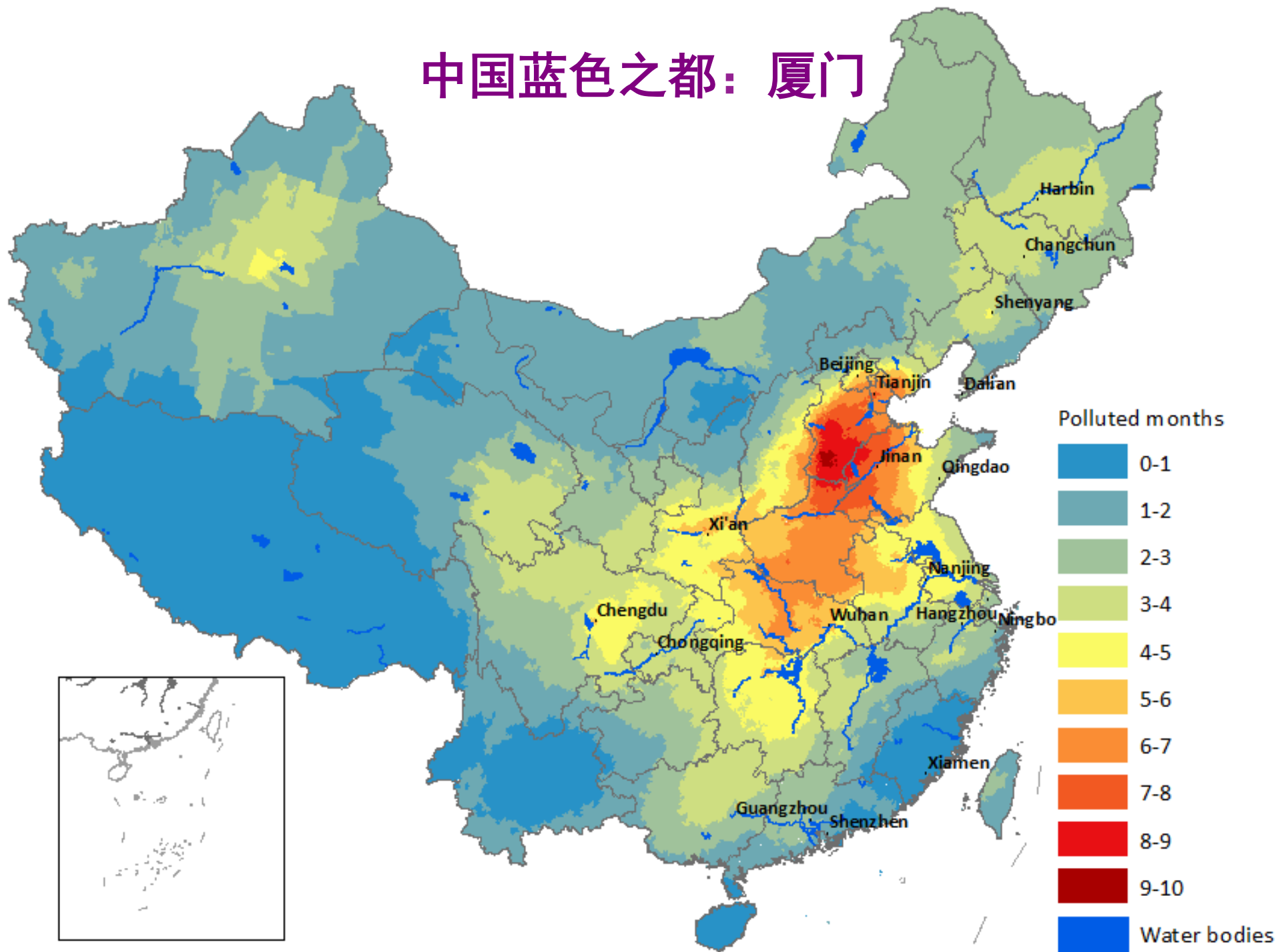
Beijing City Lab, BCL, 北京城市实验室

- 组织架构
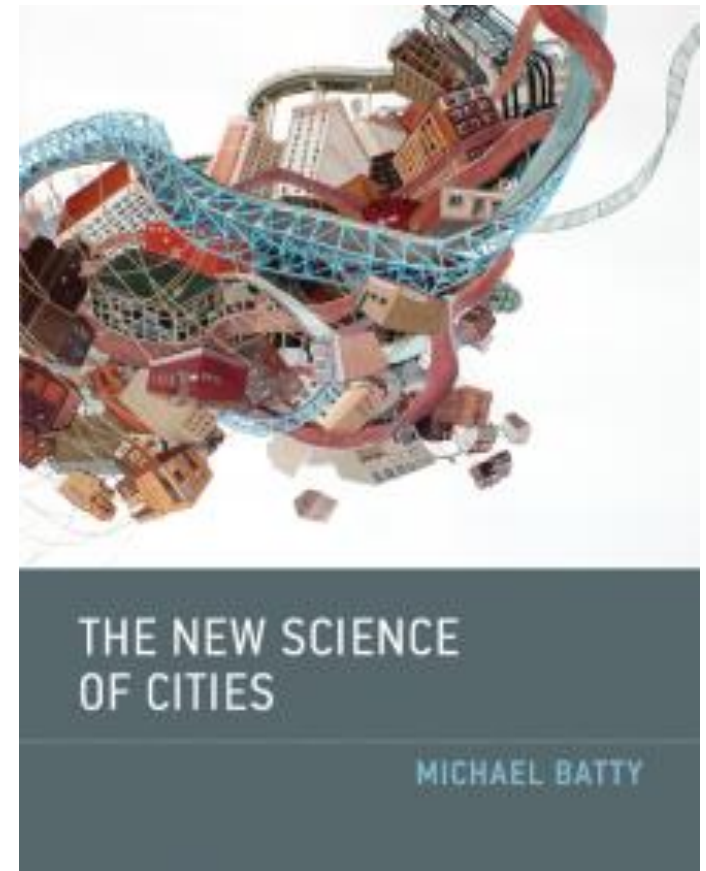 - 核心团队（×7）—两个数据工程师
  - 资深学人组成的荣誉会员（×11）
  - 作为骨干的研究员（×24）
  - 青年学生会员（×38）
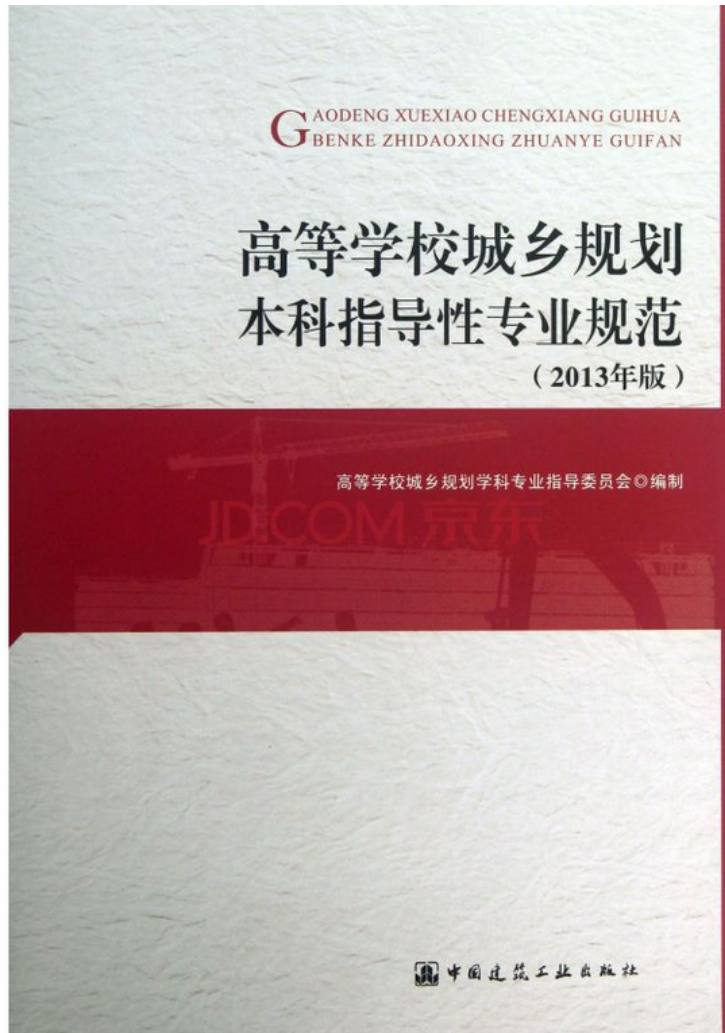 - 大量关注者(6000+)
- 定位
  - 一个定量城市研究学术网络
  - 一个开放的共享平台(40篇工作论文+24项研究数据）
  - 一次科学地理解城市的尝试
  - 一组吸引公众参与的可视化表达
  - 关注北京、放眼全国

一次科学地理解城市的尝试

- 规划信息化？ ✖

- 规划新技术？ ✖

- 定量城市研究 — 城市科学 (A Science of Cities)?✔

THE NEW SCIENCE OF CITIES

MICHAEL BATTY

未来4、6、10年？

高等学校城乡规划本科指导性专业规范（2013年版）

高等学校城乡规划学科专业指导委员会◎编制

中国建筑工业出版社

- 地理信息系统应用（**10**门核心课程之一）

- （城市与区域发展知识领域的知识单元）城乡技术与信息，**32**学时，包括城乡规划信息技术、城市发展模型、城市系统工程

- （规划技术类方向知识单元）城市系统分析方法，**32**学时，统计分析方法、城市模型建构方法、城市模拟技术

- 地理信息系统应用（城乡规划理论与方法之子领域），**32**学时

- 其他知识单元：城市用地分类及其适用性评价、城乡社会综合调查研究、城乡规划公众参与

- 规划支持靠教育，不是靠城市研究学者
- 广义的规划支持（任何城市研究）与狭义的规划支持（系统）

来自五洲四洋的访客（BCL visitors）

EMAIL: LONGYING1980@GMAIL.COM
BCL网址: WWW.BEIJINGCITYLAB.ORG OR
LONGY.JIMDO.COM

城市模型学术报告会
暨2014 北京城市实验室年会
Urban Modeling Workshop & 2014 BCL Conference

BCL Beijing City Lab

时间PM	人物	事件/报告题目
2:00-2:10	毛其智	致欢迎辞
2:10-2:20	龙 瀛	北京城市实验室（BCL）介绍
2:20-2:45	龙 瀛	大模型及其中国应用案例
2:45-3:10	李 栋	位置微博与城市规划
3:10-3:35	刘行健	多尺度城市网络分析
3:35-4:00	茅明睿	规划人：基于刷卡和微博数据的规划行业人脉分析"
4:00-4:10		间歇
4:10-4:35	王江浩	志愿者地理大数据分析与可视化：中国案例研究
4:35-5:00	吴 康	中国PM2.5的人口暴露评估：基于精细单元的空间分析
5:00-5:25	周江评	中国城市之城市形态、职住平衡与通勤效率
5:25-6:00	大家一起来	嘉宾点评和自由讨论

2014年6月11日（周三）下午2时至6时 清华大学建筑学院王泽生报告厅（建筑馆一层）

清华大学 新闻网

首页 ＞ 综合新闻 ＞ 内容

最新新闻列表

城市模型学术报告会暨2014年北京城市实验室年会在清华大学举办

清华新闻网6月12日电（通讯员 **袁晓辉**）6月11日，"城市模型学术报告会暨2014年北京城市实验室年会"在清华大学建筑学院举办。

图为报告会现场。

　　本次报告会由清华大学人居环境实验室（TSHSI Lab）和北京城市实验室（Beijing City Lab，BCL）共同组织， 来自中国科学院、清华大学、北京大学、北京交通大学、北京建筑大学、北京联合大学、北京航空航天大学、首都经贸大学、北京测绘设计研究院、清华同衡城市规划设计研究院、北京城市规划设计研究院、中国城市规划设计研究院、武汉大学、河海大学、中山大学、美国爱荷华州立大学、美国北卡大学、美国佐治亚大学等30多个单位的150多人参加了此次学术报告会。

BCL积累了大量的开放数据和大数据

- ## 物理空间
  - 全国所有城市的行政边界、路网、现状地块、城镇用地范围、规划许可（部分城市）、土地利用图（多年）、数字高程模型、不透水面、夜光遥感反演的城镇用地（多年）、自然地理要素等

- ## 人口分布
  - 全国2000和2010乡镇街道尺度人口密度、地块尺度人口空间化和属性合成

- ## 居民生活质量评价
  - 全国所有城市环境信息（包括PM2.5）、兴趣点POI、公共服务设施、房价、公交线路和站点、餐厅

BCL积累了大量的开放数据和大数据

- ## 人类活动与移动
  - 人类活动热力图（部分城市）、签到、空间微博、城市间交通联系、北京的公交卡（多年）、居民出行调查（多年）和出租车轨迹

- ## 对未来的判断
  - 规划图（200+城市）、近期城市扩张情景

BCL开展的一系列定量城市研究项目
(传统数据、大数据、开放数据、大的开放数据)

Projects

1 BUDEM

2 Urban Growth Boundaries

3 Bus Landscapes

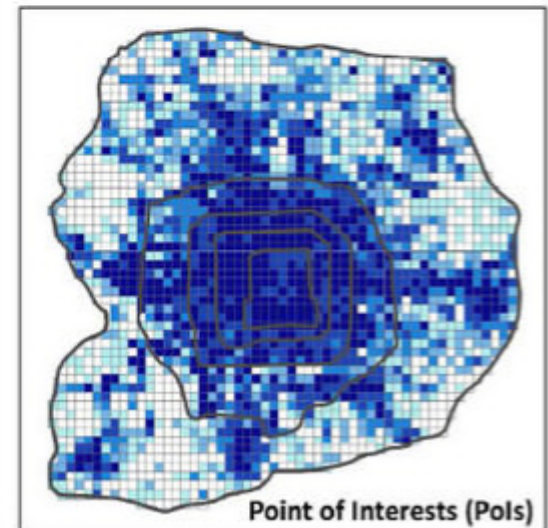4 Population China

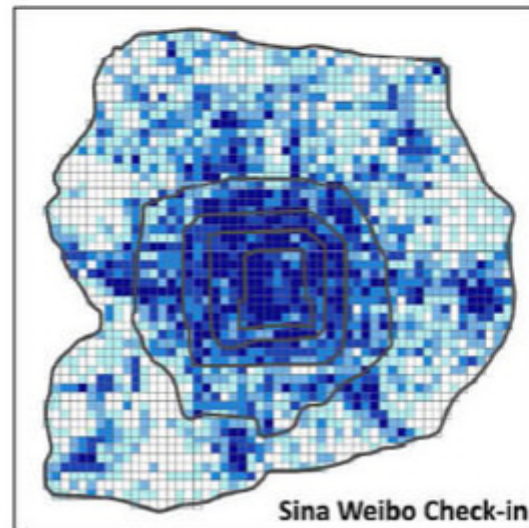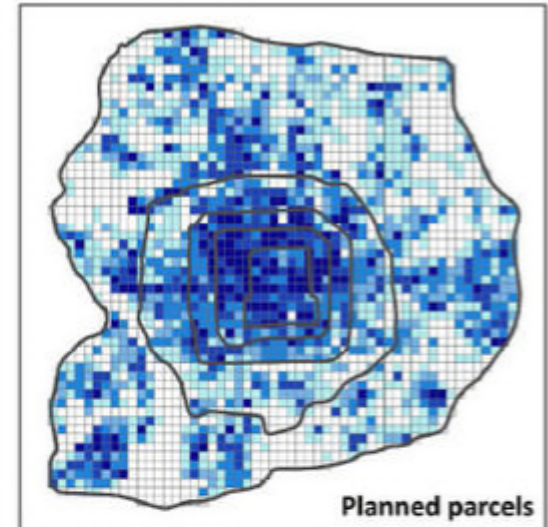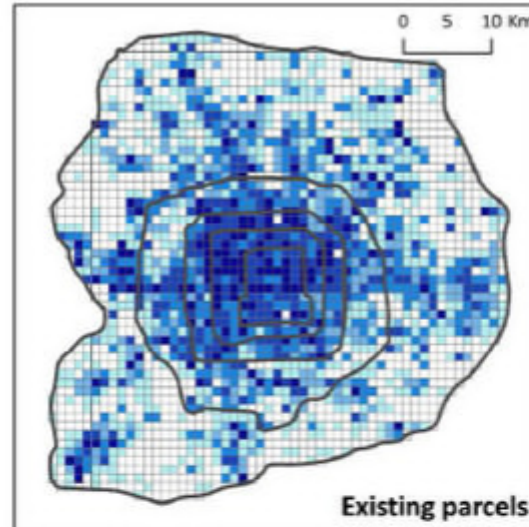5 Planning Support Systems

5 Urban Form

6 Population Synthesis

7 Social Network Mining

8 Big Model

9 Beijing Parking

10 Urban Network Analysis

BCL排行（RANKING）敬请期待

利用公交和地铁刷卡数据
分析北京通勤交通和城市贫困问题

1 INTRODUCTION

Smart card data from public transportation
智能交通卡、一卡通、八达通…

- Smart card data collected by automated fare collection systems
  - Bus
  - Subway
- Since 1990's the use of smart card has become significant (Blythe, 2004)
- Overwhelmingly adopted by Chinese cities
  - Over 100 cities in 2007
- With precise bus trip information (both boarding and getting off)
  - A spatial resolution of bus stop
  - A temporal resolution of second

Existing researches using SCD

- Most researches advocate providing decision making for planning and design of public transportation system (see Pelletier et al (2011) for a review).
  - This is also the focus of Chinese papers
- In South Korea, Joh and Hwang (2010) analysed the trip trajectories using four million individuals' trips from the bus SCD, and correlated them with land use characteristics in the Seoul Metropolitan Area.
- In the U.K., Roth et al (2011 coauthored with Michael Batty@CASA) used real-time "Oyster" card database of individual person movements in the London subway to reveal the polycentric urban structure.
- Less attention was paid on analyzing jobs-housing relationships as well as commuting pattern of a metropolitan city.

This talk is regarded with

- Using Beijing bus SCD for identifying cardholders' housing and job places

- Identifying commuting trips of cardholders

- Analyzing and mapping bus commuting pattern of Beijing

2 DATA

Bus lines
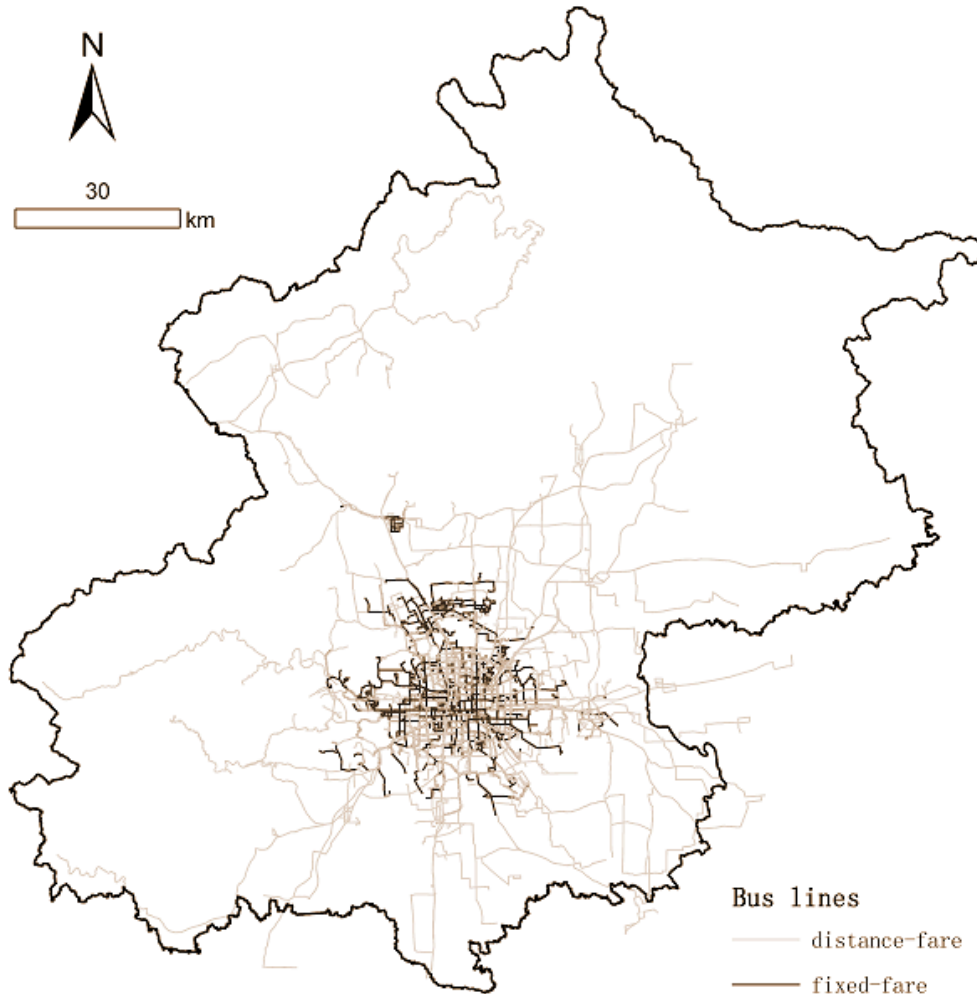
- The Beijing Metropolitan Area
  - 16410 sqkm
- 1287 bus lines
- Two types
  - Distance-fare (long dist, inner city to suburb)
  - Fixed-fare (short dist, within inner city)

Bus lines
distance-fare
fixed-fare
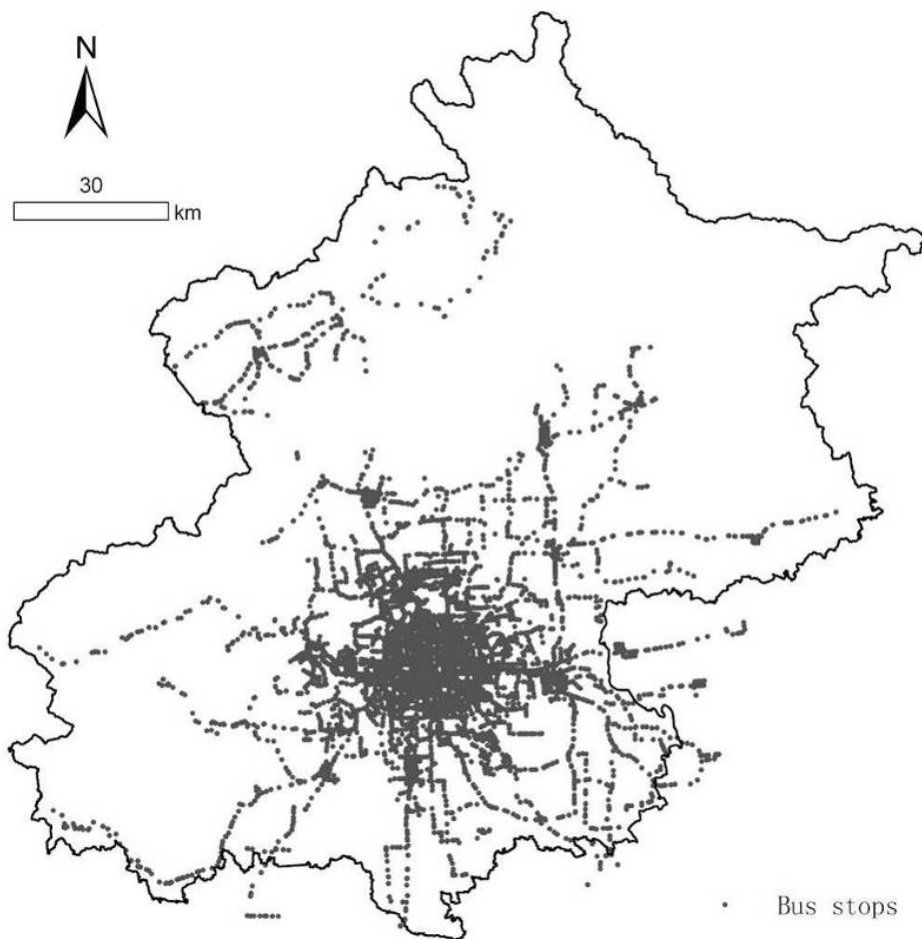
Bus stops

- 8691 bus stops
- The spatial resolution of identification
- The average distance between two stops
  - 231 m

Traffic analysis zones
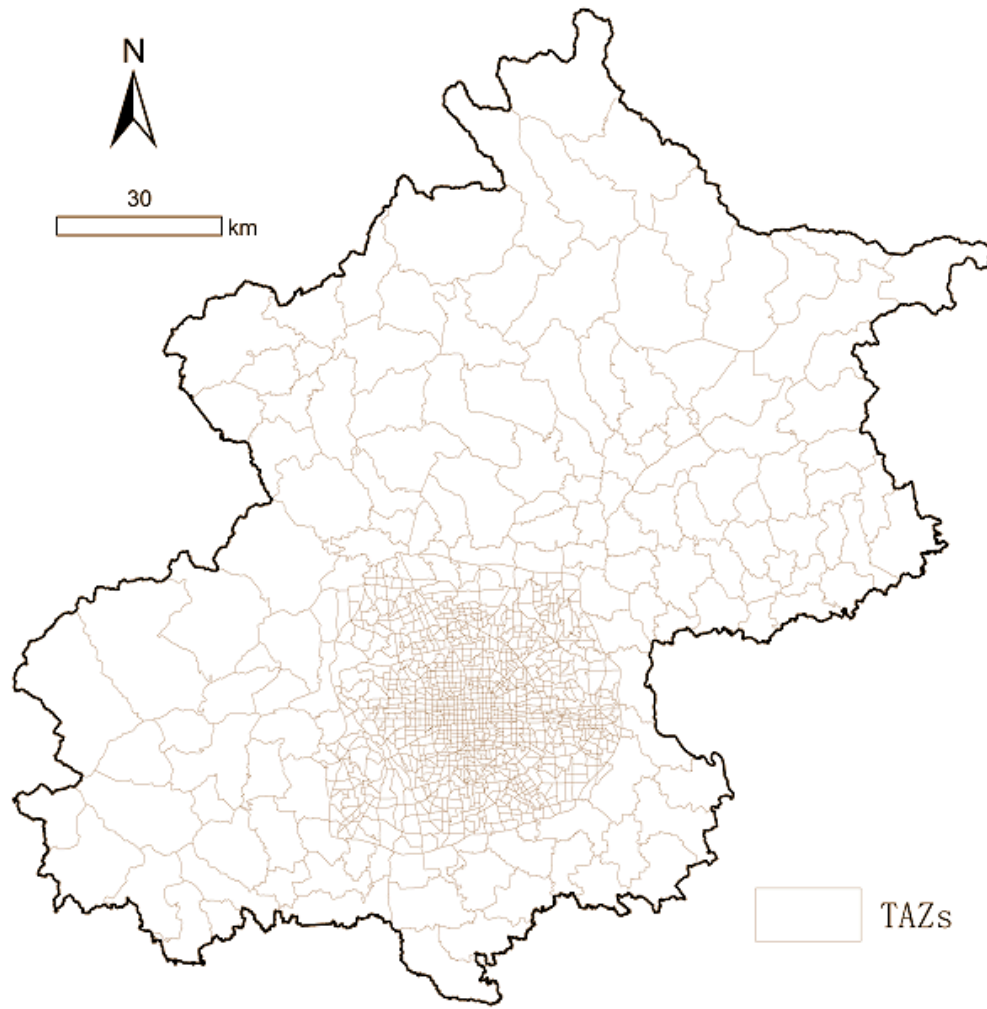
- 1118 TAZs
- For aggregating identification results
  - Originally in the bus stop level

Land use pattern
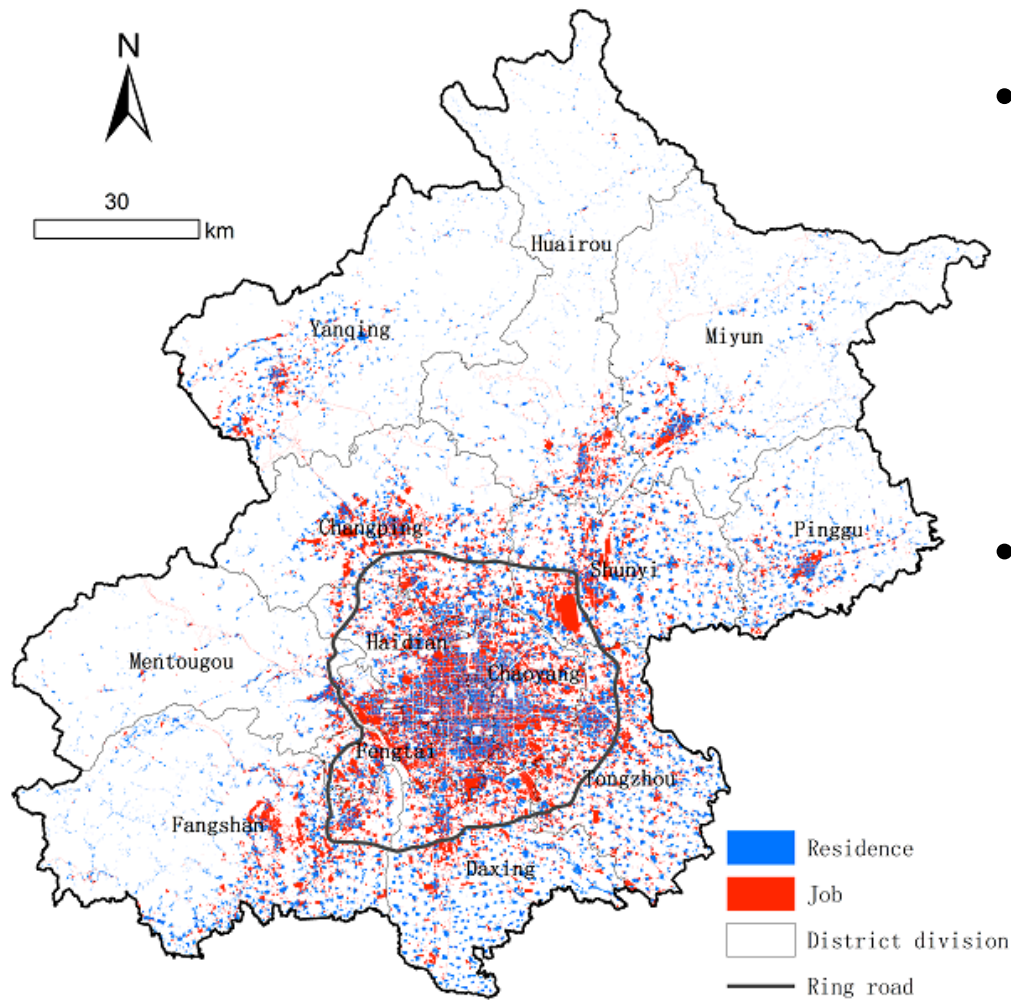
- 133503 parcels in the BMA
  - 29112 residential parcels
  - 57285 parcels with job positions
    - Job parcel

- For calculating residential or job potential of each stop
  - In case of periodic identification results using one-week data

r23

Bus SCD in Beijing

- Anonymous card
- Over 90% bus riders use smart cards till April 2007

- One week in April 2008 from
  - Monday to Sunday
  - Subway not included
- 80 million records for 8.5 million cards/persons
  - One record for a bus riding of a cardholder
  - 1.3 daily bus ridings per person

SCD data structure

Variable	Exemplified Values
Card ID	"10007510038259911", "10007510150830716"
Card Type	1, 2, 3, 4
Line ID	602, 40, 102
Line Type	0, 1
Driver ID	11032, 332
Vehicle ID	111223, 89763
Departure Data	2008-04-08
Departure Time	"06-22-30", "11-12-09"
Departure Stop	11, 5, 14
Arrival Time	"09-52-05", "19-07-20"
Arrival Stop	3, 14, 9
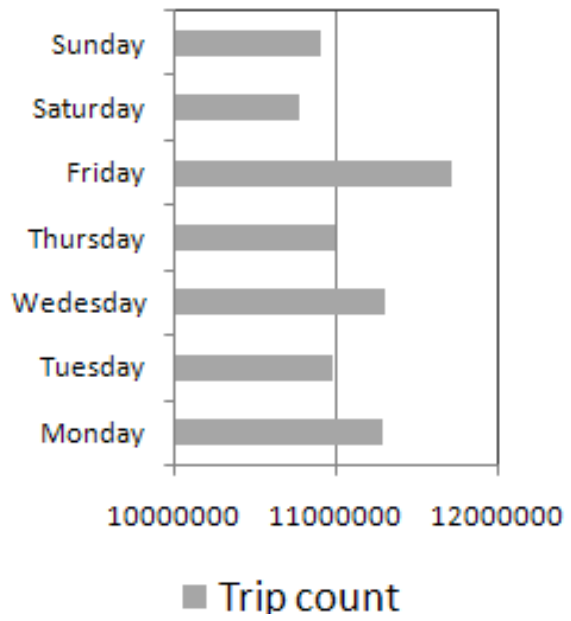
Trip count (bus riding) for each day

Trip count in each hour on Tue and Sat

Trip density in the TAZ level in the whole week

26

Count of bus ridings for each TAZ

0:00-1:00

N

▨ (yellow)	1 – 146
▨ (light orange)	147 – 383
▨ (orange)	384 – 716
▨ (brown)	717 – 1288
▨ (dark red)	1289 – 3411
▢	Ring road

10 km

27
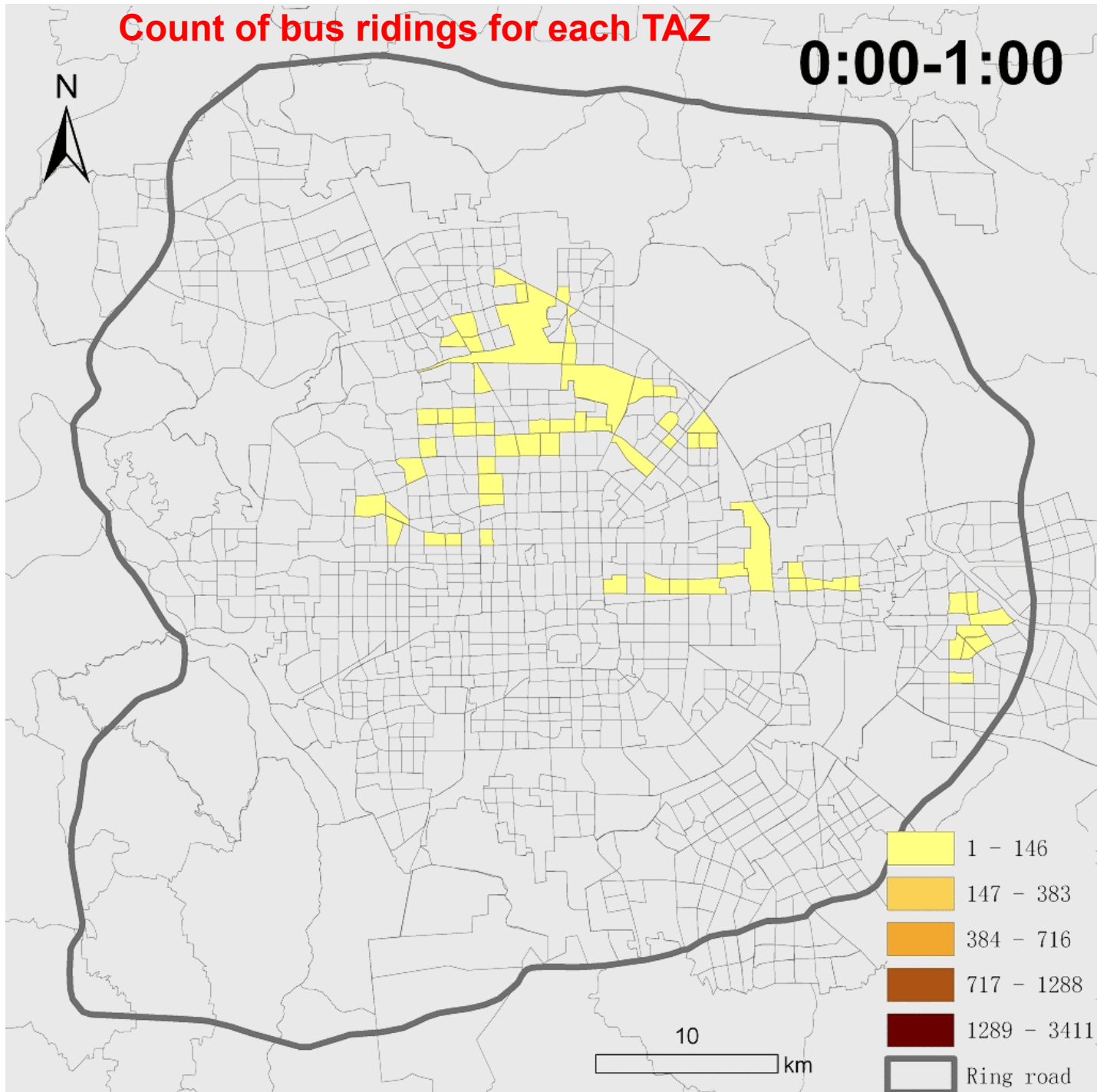
Special attention paid to fixed-fare trips

- ## Distance-fare
  - With full spatiotemporal information
- ## Fixed-fare
  - Only arrival time and stop ID
  - No departure time and stop ID

The Beijing Household Travel Survey

(The 2005 survey)

- Conducted in 2005 for the whole BMA
- 81,760 households/208,290 persons, with 1.36% sampling rate
  - 800 thousand trips
- Information included:
  - Socio-economic attributes of household and persons
  - One-day travel diary of each person
    - Mode, purpose (e.g. commute), OD, departure and arrival time, etc.
- For setting rules for identifying jobs-housing places and commuting trips

3 APPROACH

Data preprocessing

- Data cleaning
  - Abnormal records due to inappropriate operation on fare machines
- Merge adjacent trips for each cardholder
  - Converting segments into real trips
- Geocode the SCD
  - by linking the bus stop ID in the SCD with the bus stop layer.
- Combine trips of each cardholder
  - to retrieve cardholder's full bus travel diary
- Two data models proposed for identifying commuting pattern
  - PTD: Position-start Time-Duration
    - For identifying housing-jobs places
    - {H0, 18:00 (-1), 13 h} and {J0, 8:00, 9h}
  - TRIP: departure / arrival time and space
    - For identifying commuting trips
    - {H0, 7:00, J0, 8:00} and {J0, 17:00, H0, 18:00}

Identification of housing places using one-day data

- The departure bus stop of the first trip (TRIP1) will be the housing place of a cardholder.

- Grounded on:

  - In the 2005 survey, 99.5% person's first trips start from home

- Note that a person is assumed to live around the identified stop with a maximum distance of 750 m

  - Retrieved from walking trips in the 2005 survey

  - This could be the spatial resolution of our identification results

Identification of job places using one-day data

- An activity of a non-student (except at home) with a duration longer than 6 hours is assumed to be working.
  - *If:*
    - *Condition 1*: The card type is not the student card
    - *Condition 2*: $D_k > 360$
    - *Condition 3*: $k <> 1$
      - he/she is sleeping at home when $k=1$
  - *Then:*
    - the k_{th} place P_k will be regarded as the job place of this cardholder
- Grounded on:
  - In the 2005 survey, 96% persons work over 6 hours per day.

Combining each day's result
to get more solid housing-jobs places

- One day result might not be stable, e.g.
  - A shopping, hospital, or creation activity longer than 6 h
  - One night sleeping at a friend's home
- Periodic pattern of cardholders with identified housing-jobs results like
  - M T W T F S S
  - a a a b a a a (Person 1)
  - **a a a b b b c (Person 2)**
  - a a b b c c d (Person 3)

- A decision tree for combining seven daily result
  - A complex procedure
- For aaabbbc pattern:
  - Housing or job potential of each stop
  - "a" with a higher housing potential was selected as the final housing place of Person 2

Start

$n \leq 1$ — Y → No housing place identified (P_H = "")

N

$N=1$ — Y → P_H = most frequent places   (if two, the one with the highest p selected)

N

the maximum stop count of all clusters = 1? — Y → No housing place identified (P_H = "")

N

there is only one cluster with the maximum stop count? — Y → P_H = most frequent stop in this cluster   (if two, the one with the highest p selected)

N

most frequent place count is not the same in maximum clusters? — Y → P_H = the most frequent place in the maximum cluster with most frequent place count

N

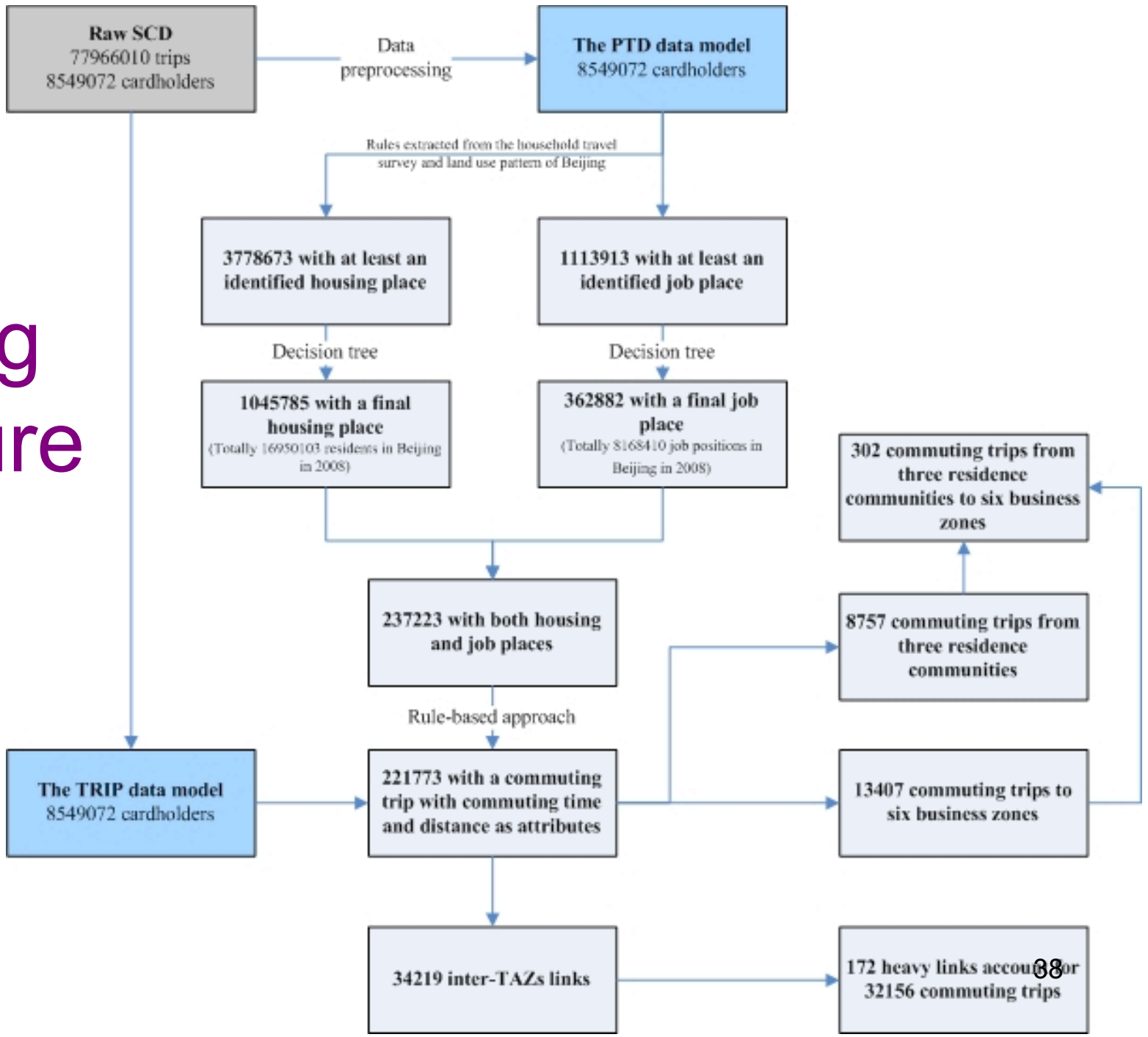P_H = the most frequent place in maximum clusters with the highest p

Identification of commuting trips

- Limited to those cardholders with both identified housing and job places.

- Rules used:
  - The boarding bus stop of the first trip in a day is the identified housing place.
  - The job place appears in trips of a day.
  - Both the housing and job places are identified in the same day.

- Commuting time and distance are calculated for each identified commuting trip.

4 RESULTS

A big picture

Raw SCD
77966010 trips
8549072 cardholders

Data preprocessing

The PTD data model
8549072 cardholders

Rules extracted from the household travel survey and land use pattern of Beijing

3778673 with at least an identified housing place

1113913 with at least an identified job place

Decision tree

Decision tree

1045785 with a final housing place
(Totally 16950103 residents in Beijing in 2008)

362882 with a final job place
(Totally 8168410 job positions in Beijing in 2008)

302 commuting trips from three residence communities to six business zones

237223 with both housing and job places

8757 commuting trips from three residence communities

Rule-based approach

The TRIP data model
8549072 cardholders

221773 with a commuting trip with commuting time and distance as attributes

13407 commuting trips to six business zones

34219 inter-TAZs links

172 heavy links account for 32156 commuting trips

38

1,045,785 cardholders (12.2% of all 8549072 cardholders) are identified with housing places, and 362,882 cardholders (4.2% of all) are identified with job places.
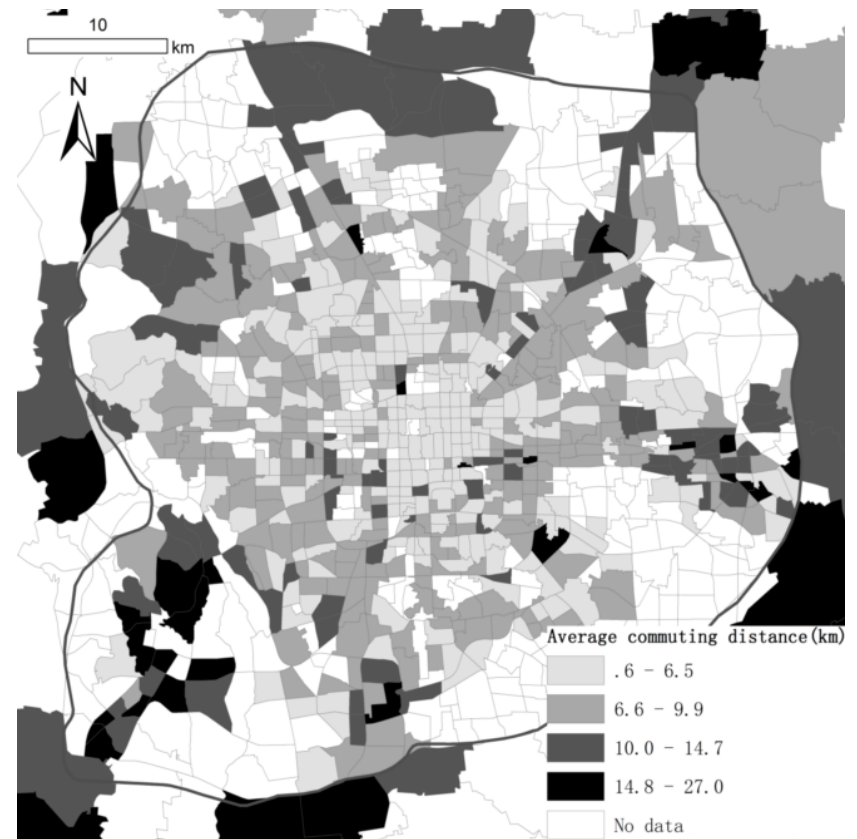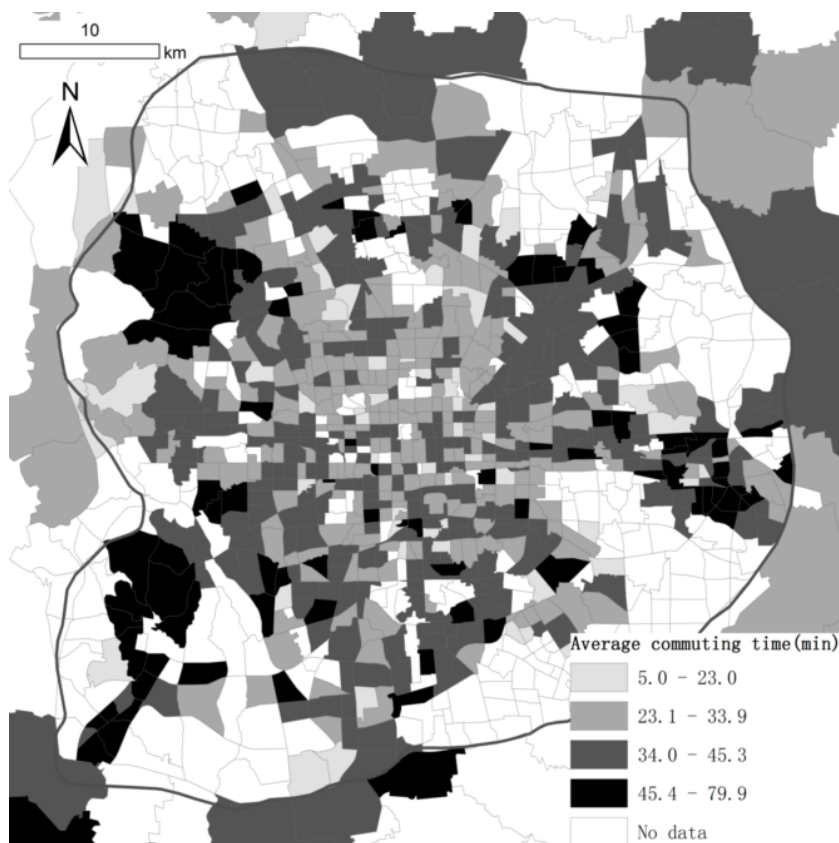

Cardholder density with an identified housing place

Cardholder density with an identified a job place

Commuting trips identification

221,773 cardholders identified with commuting trips

Average commuting **time** and **distance** for each TAZ   40

Comparing with existing researches

| Name	Travelling modes and year	Sample size	Average commuting time (min)	Average commuting distance (km)
Our commuting trips	**Bus, 2008**	**221773**	**36.0 (24.2)**	**8.2 (7.0)**
The 2005 survey	Bus, 2005	6651	40.5 (23.1)	8.4 (8.3)
Liu and Wang, 2011	Bus, 2007	307	46.3 (N/A)	N/A
Wang and Chai, 2009	Bus, 2001	227	55.1 (30.4)	N/A
Zhao et al, 2011	Bus and metro, 2001	220	52.4 (26.6)	N/A

Note that the number in the bracket is the standard deviation of the average commuting time and distance. Except our commuting trips, bus samples in other researches are extracted from the whole survey with all travelling modes.
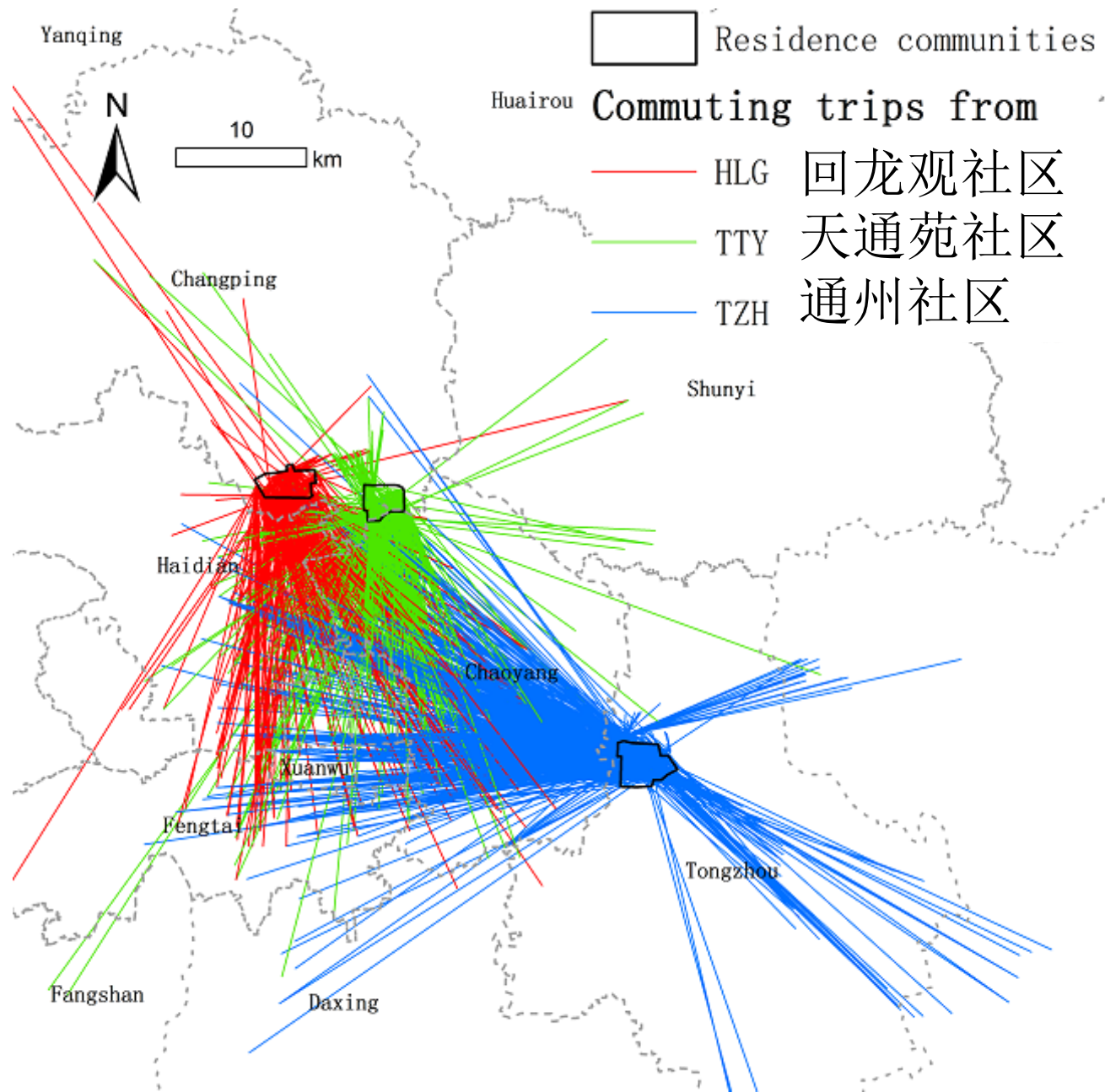
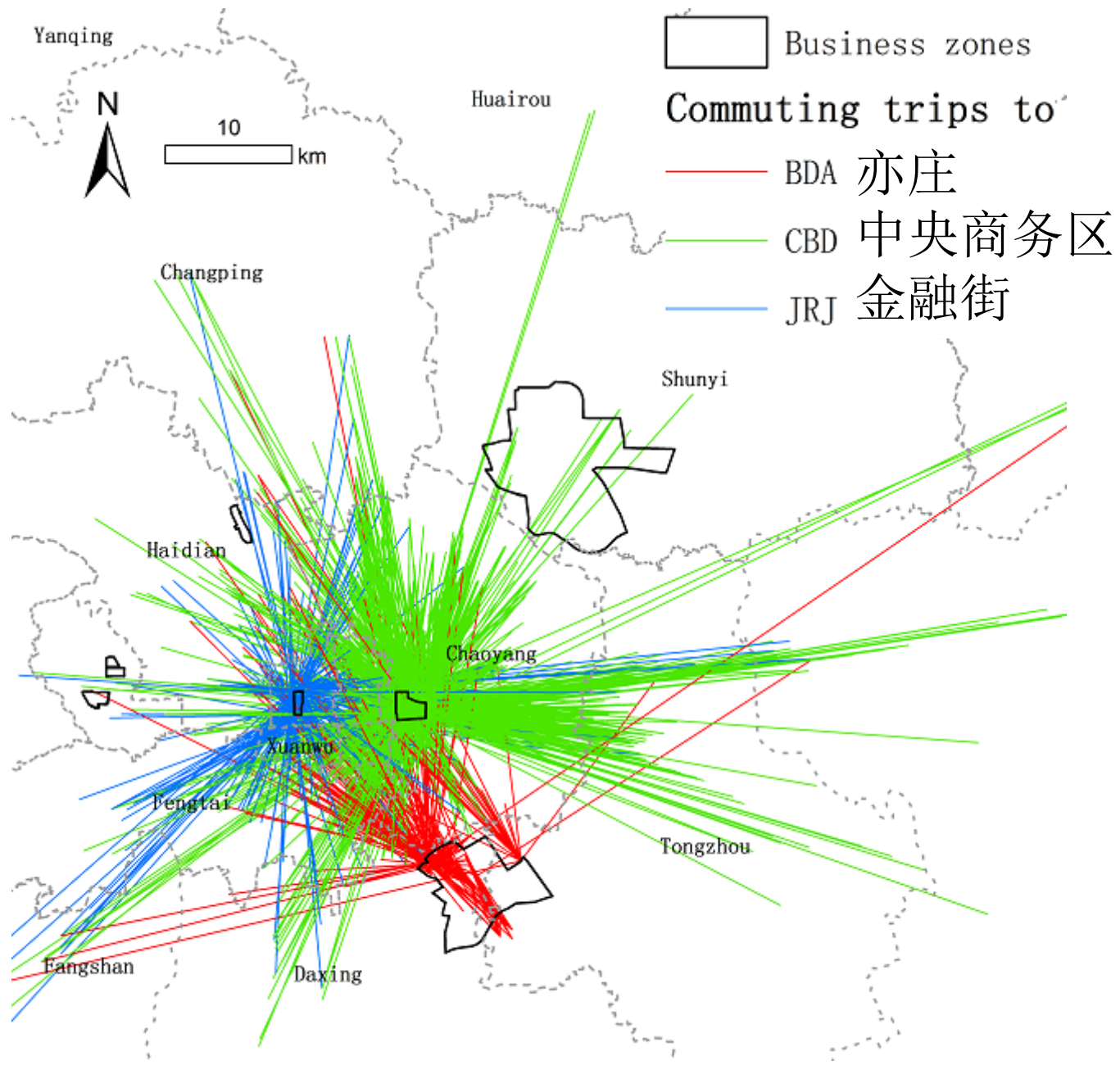# Bus commuting pattern

- **Where to where**
- **Extreme commuting**

42
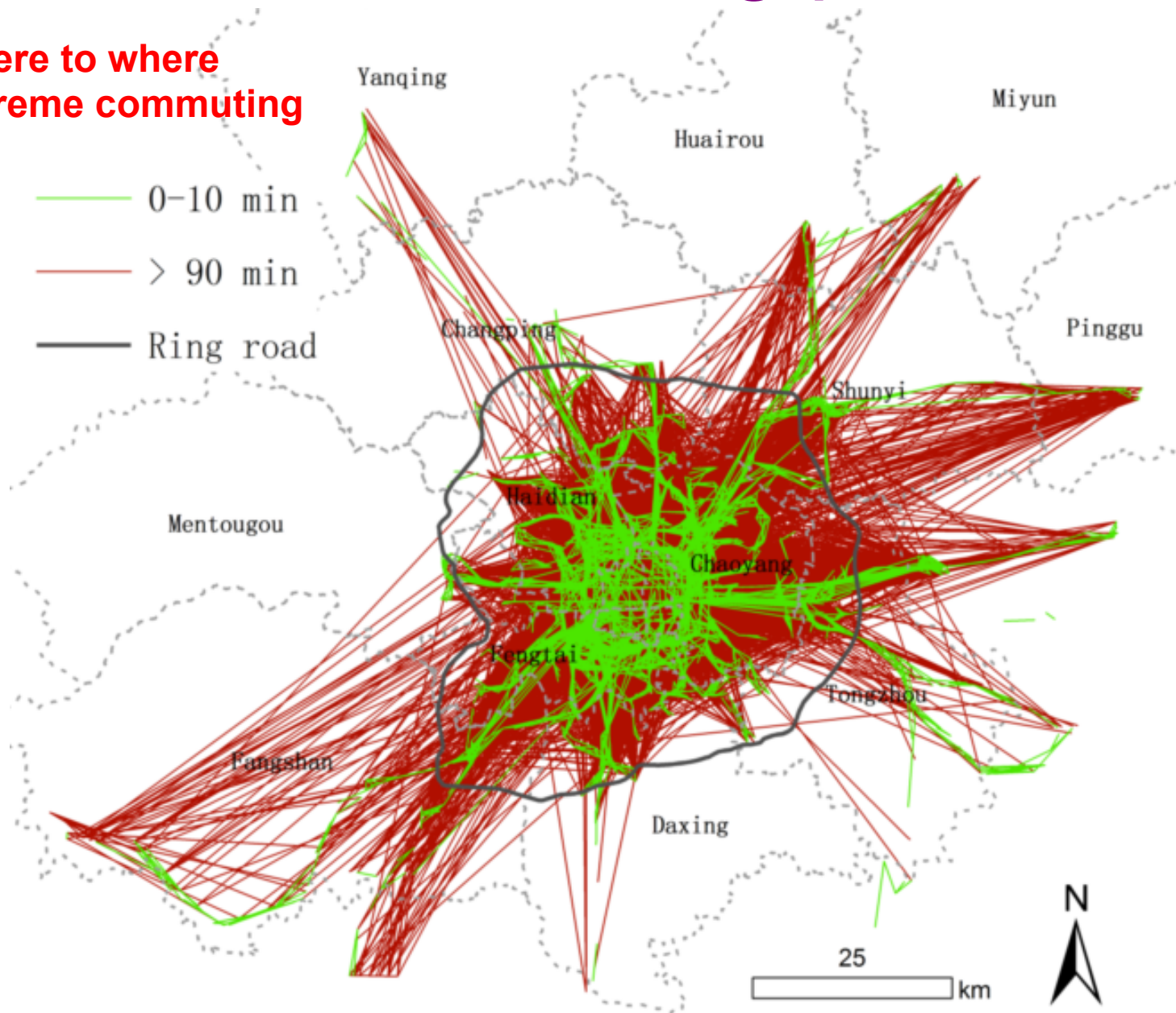
- **Aggregated in the TAZ level**
- **The head/tail division rule used**
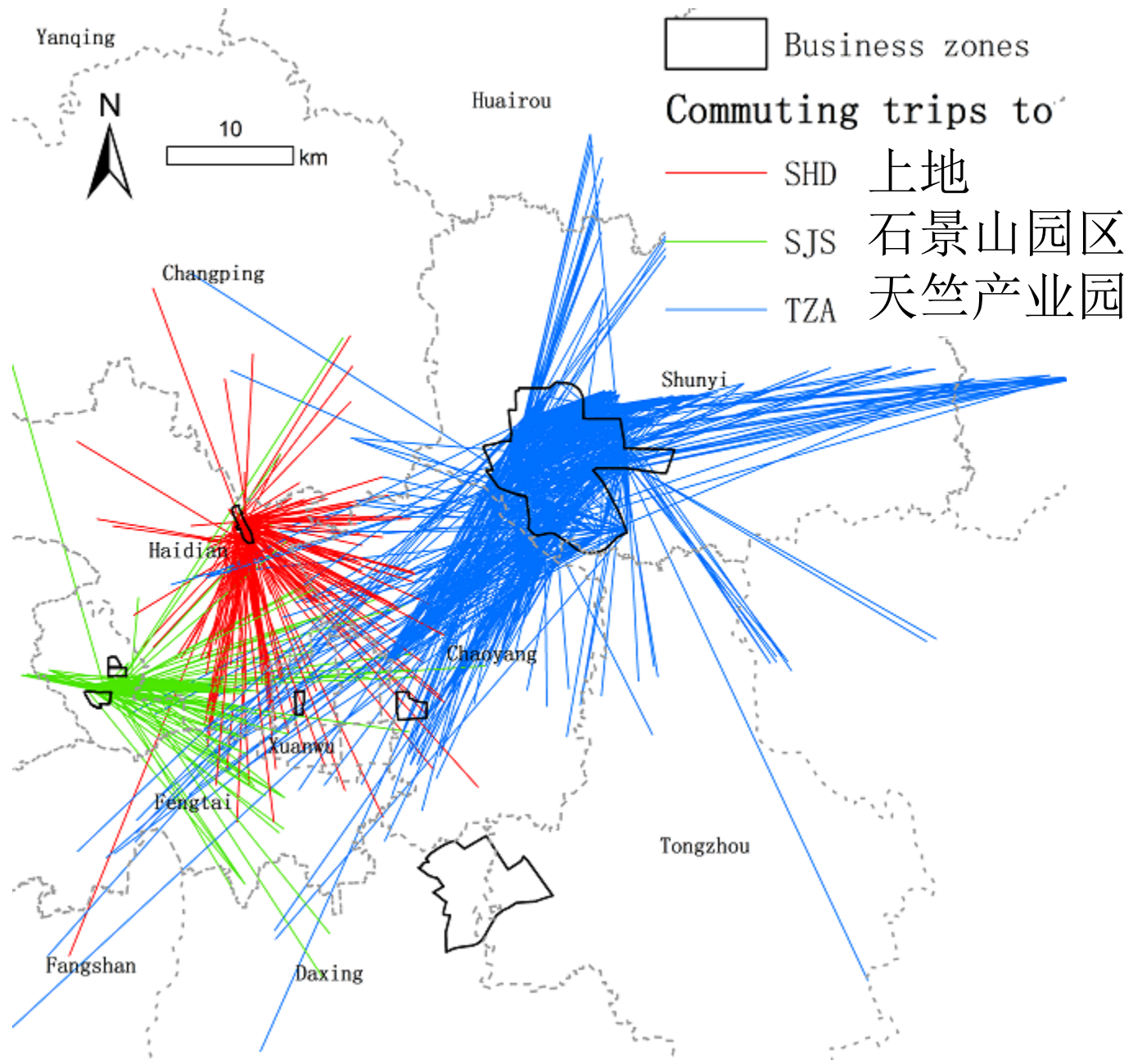- **Policy suggestion on BRTs**
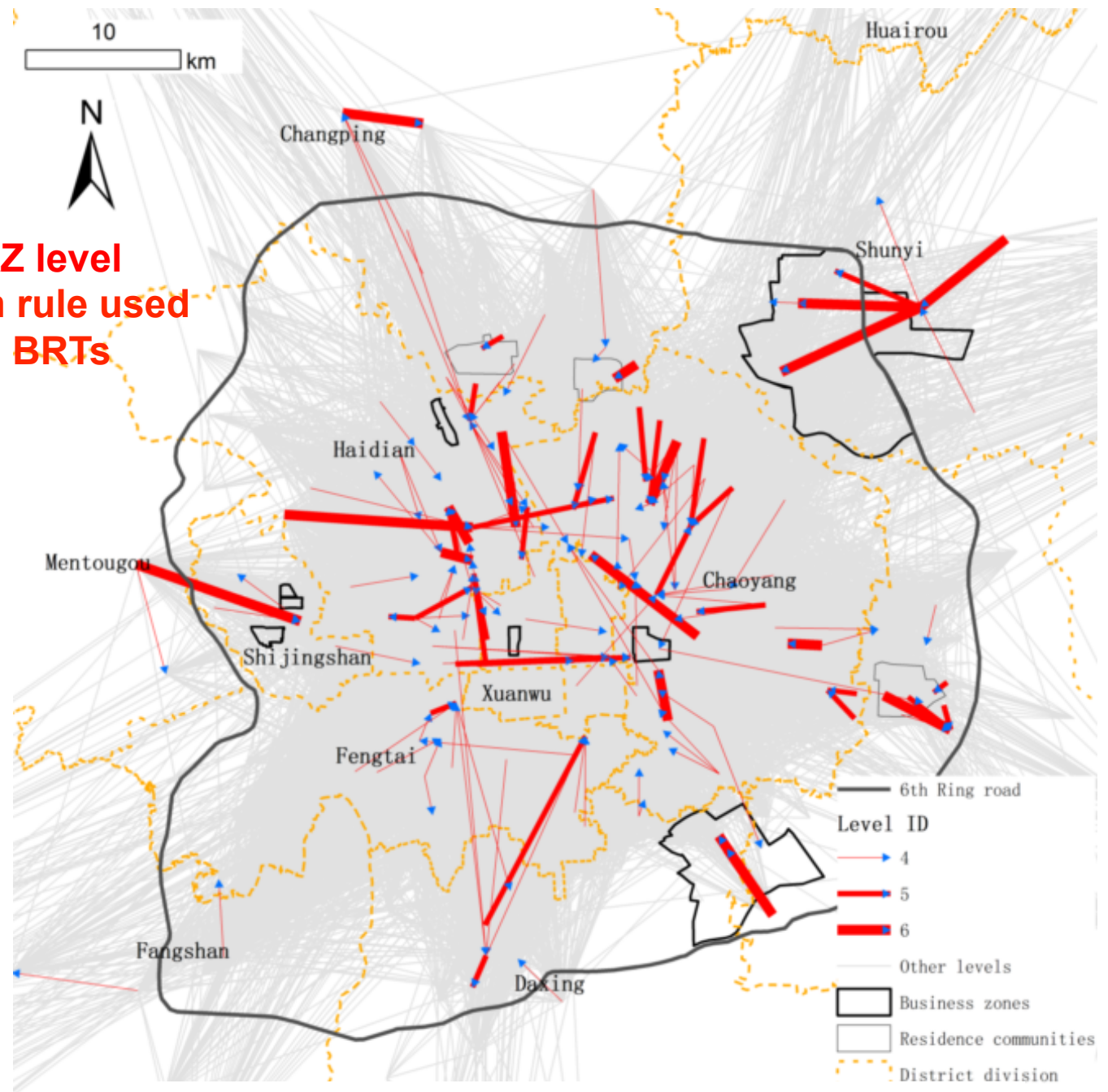
Inter-TAZ commuting mainstream

Commutes from main residence communities

Commutes to main business zones (1)

Commutes to main business zones (2)

Commuting time and distance of various residence communities and business zones of Beijing

Zone name	Commuting time(min)	Commuting distance(km)	% of all identified commuting trips
Trips from residence communities			3.9
TZH	45.1	10.0	1.4
HLG	39.4	7.0	1.0
TTY	36.2	6.1	1.5
Trips to business zones			6.0
CBD	41.4	9.4	2.7
SHD	40.4	6.7	0.3
JRJ	34.9	7.1	0.5
TZA	31.6	10.0	1.3
SJS	28.4	6.9	0.3
BDA	26.6	6.4	0.8

? Only 302 commuting trips (0.14% of all identified trips) are from 3 residence communities to 6 business zones in Beijing.

5 CONCLUSIONS AND DISCUSSION

Concluding remarks

1. Propose an algorithm for identifying housing-jobs places as well as commuting trips using rules extracted from
   – the 2005 household travel survey
   – land use pattern of Beijing
2. Identification results are acceptable via comparing with the 2005 survey and existing researches
3. Analyse and map Beijing bus commuting pattern
   – Extreme commutes
   – Mainstream of bus commutes
   – Commutes from typical residential communities and to business zones
4. SCD could be a substitute of conventional travel surveys, at least a complement.

Our contribution: A promising solution for analyzing urban dynamics

- Mining LBS data using conventional conventional surveys and urban GIS layers with sound validation results

- A decision tree for determining the final one-week result using periodic information and spatial distribution of one-day result

- Retrieve Beijing commuting pattern with more accurate spatial info and more samples in contrast to existing researches using surveys

  – Although limited to bus riders

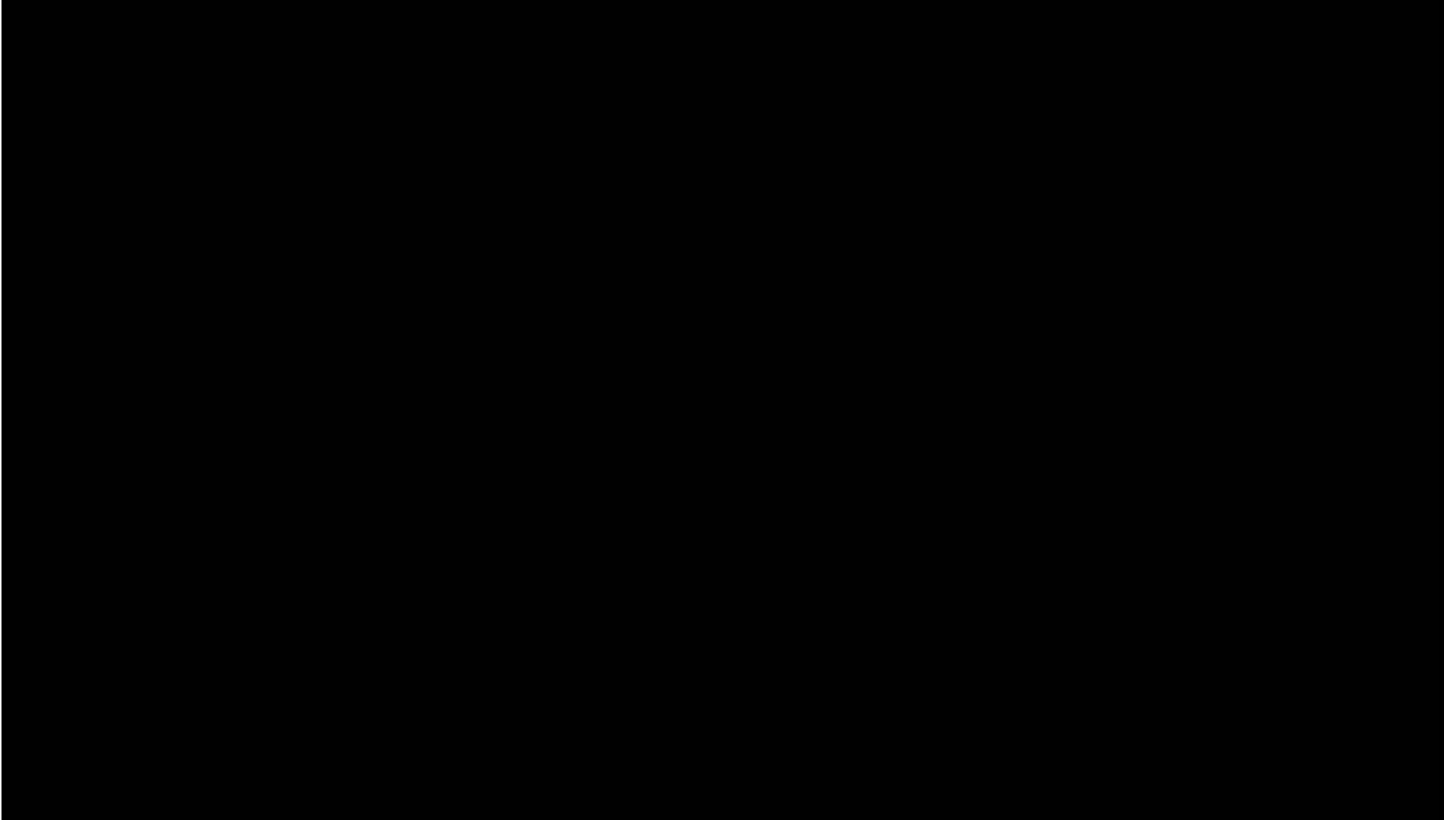# 基于2010年公交和地铁SCD的热点通勤线路
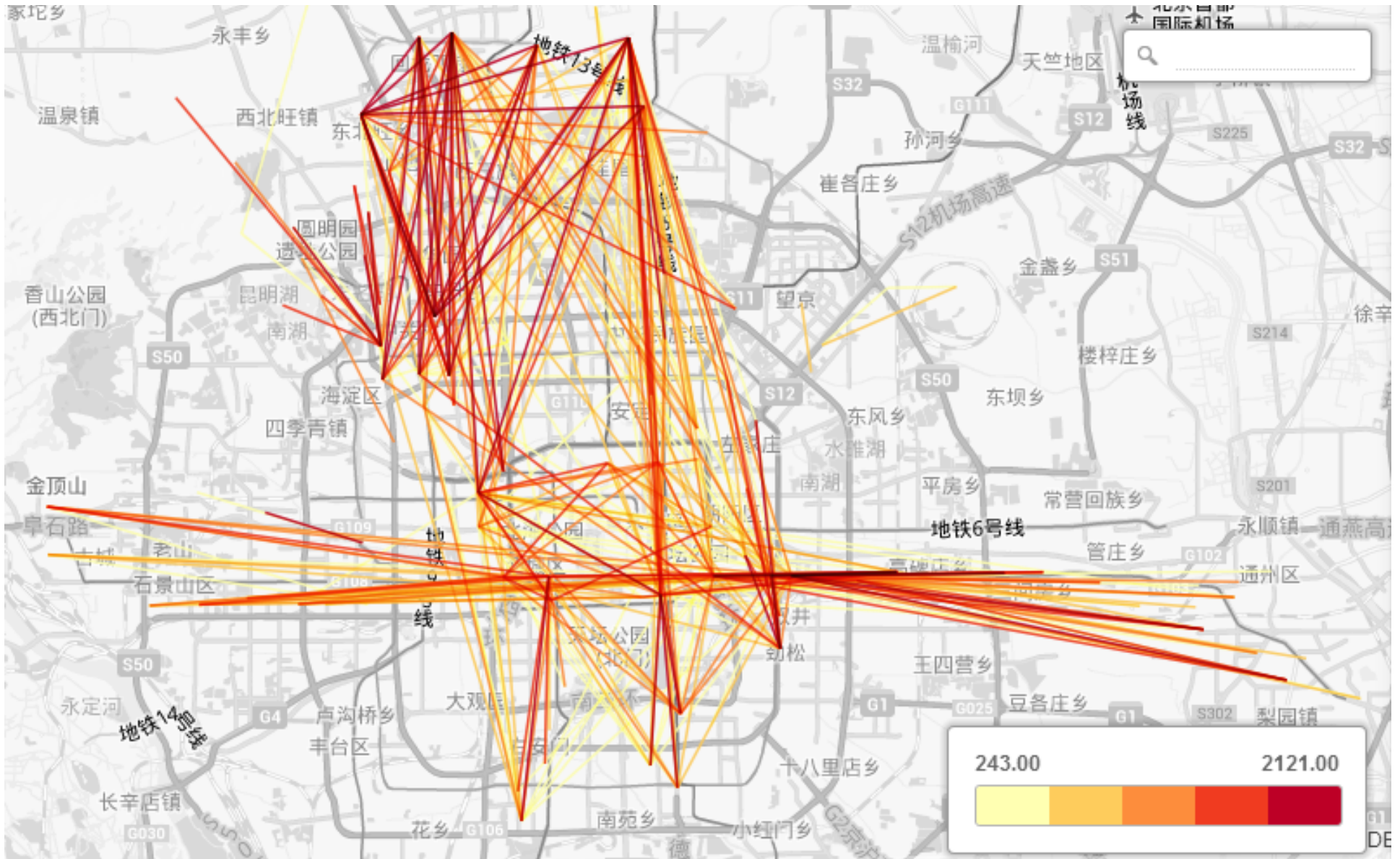
2010通勤者的居住地密度（70余万人）
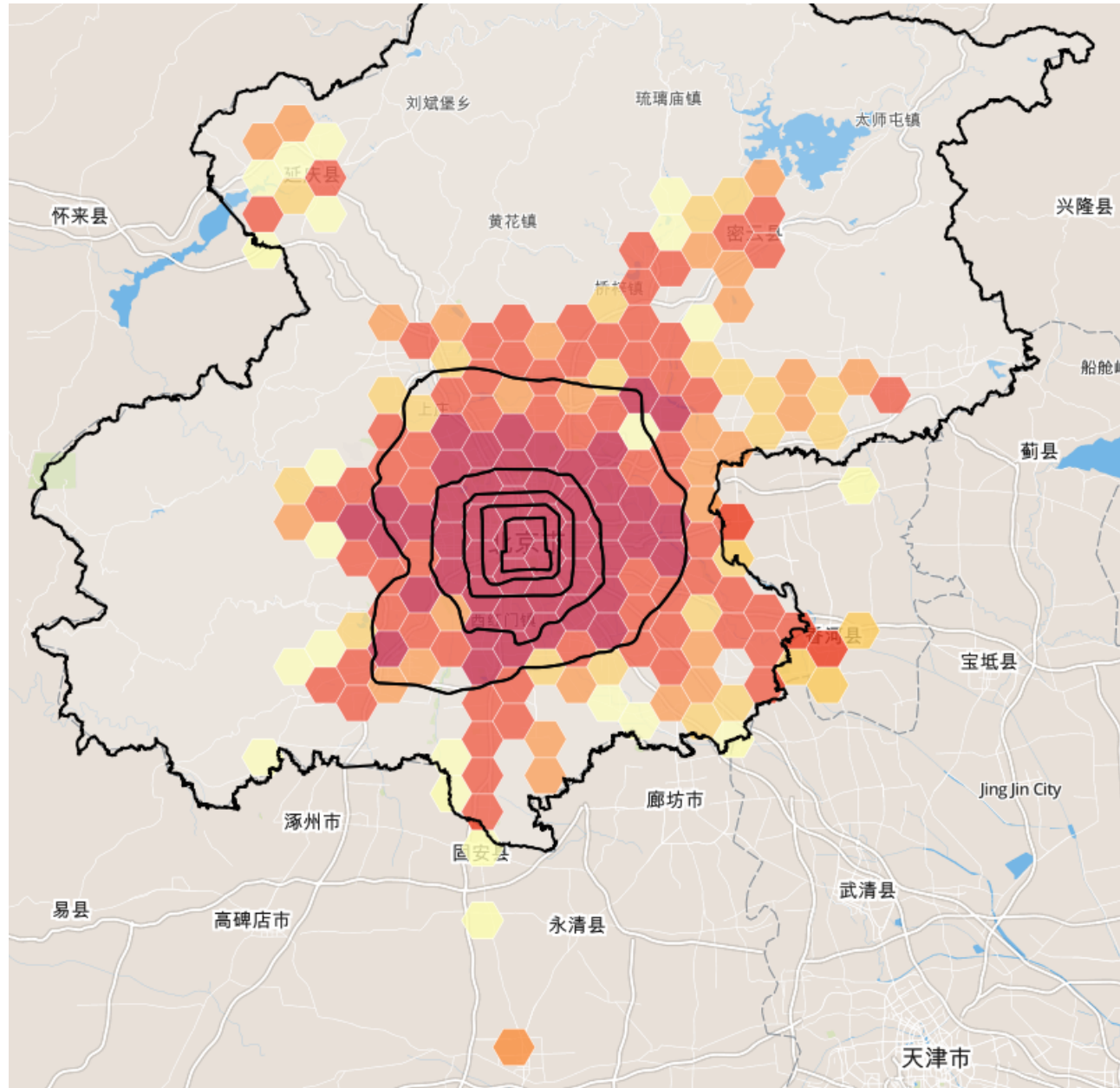
2013年某日地铁进站可视化

频繁公交出行者多为经济上的准（底层）

- **2010交通出行调查**
  - For income of 618 households, 417 are in the level 1, 166 are in the level 2, 191 are in the level 3, and 8 are in level 4.

- 柴彦威小组调查
  - Among 125 identified frequent riders, 80.8% residents are with the month income less than 6000 CNY, and most of them (50% of all) only in the 2001-4000 CNY level.

- 当地居民访谈
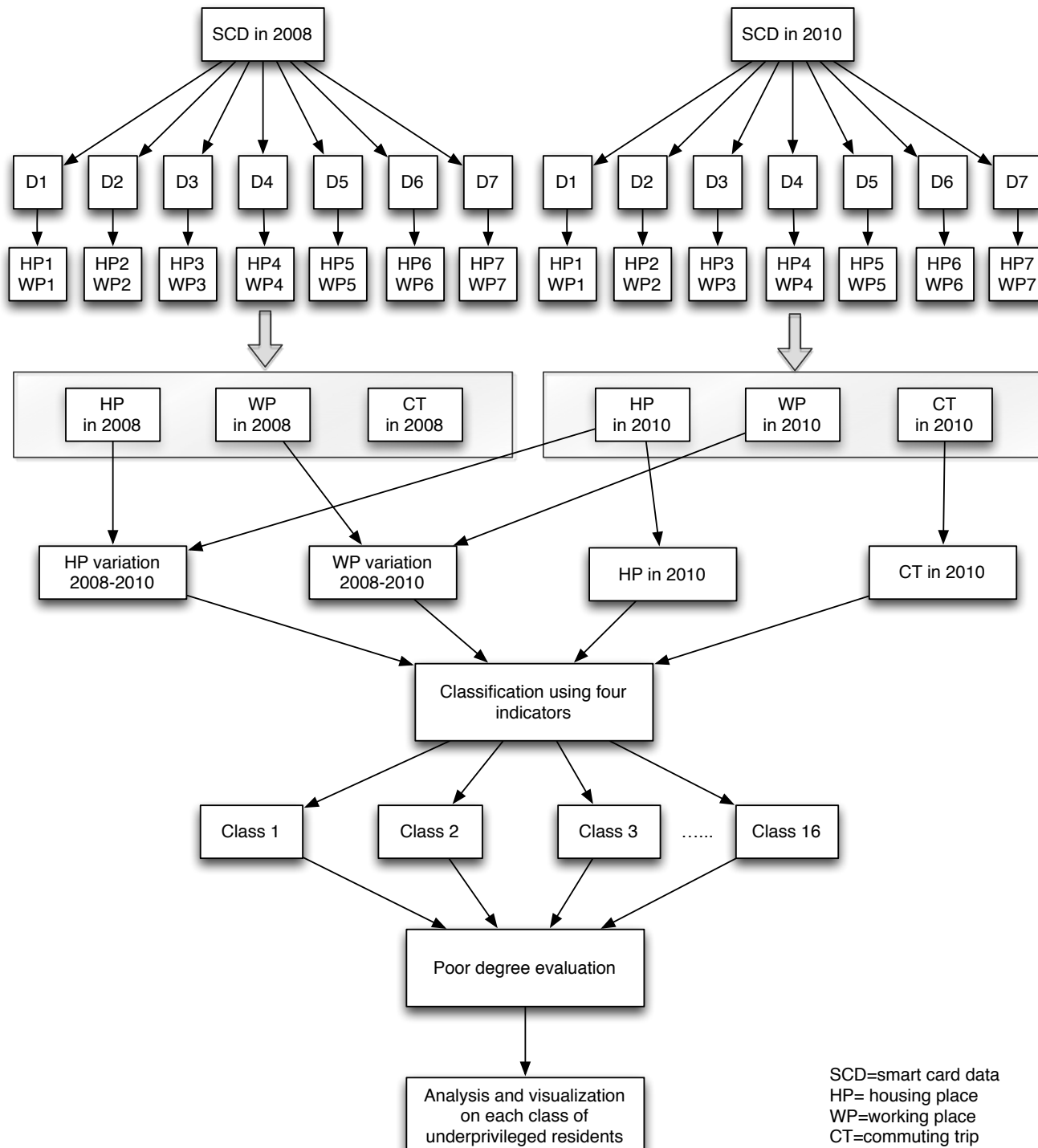
SCD=smart card data
HP= housing place
WP=working place
CT=commuting trip

Table 6 Housing place dynamics of FRs during 2008-2010

Housing place			# cardholders	Ratio (%)
Not changed			25,492	22.6
Changed			87,082	77.4
	Inward (km)		42,013	37.3
		2-5	9,211	8.2
		5-10	9,651	8.6
		10-20	13,150	11.7
		>=20	10,001	8.9
	Outward (km)		45,069	40.1
		2-5	7,990	7.1
		5-10	10,139	9.0
		10-20	16,400	14.6
		>=20	10,540	9.4
Sum			112,574	100.0

Table 7 Workplace dynamics of FRs during 2008-2010

Workplace				# cardholders	Ratio (%)
Working				14916	13.3
	Not changed			2203	2.0
	Changed			12713	11.3
		Inward (km)		6142	5.5
			2-5	1444	1.3
			5-10	1893	1.7
			10-20	2071	1.8
			>=20	734	0.7
		Outward (km)		6571	5.8
			2-5	1371	1.2
			5-10	2018	1.8
			10-20	2156	1.9
			>=20	1026	0.9
Losing job				15909	14.1
Finding a job				26919	23.9
Jobless				54830	48.7
Sum				112,574	100.00

Table 8 Commuting distance variation of FRs (with commuting trips both in 2008 and 2010)

Commuting distance in 2010 − that in 2008 (km)	# cardholders
>=20	436
10-20	1,885
5-10	2,266
2-5	2,419
0-2	2,647
-2-0	1,984
-5-(-2)	1,416
-10-(-5)	1,069
-20-(-10)	622
<=-20	172
Sum	14,916

畅想十年之后

来自五洲四洋的访客（BCL visitors）

EMAIL: LONGYING1980@GMAIL.COM
BCL网址: WWW.BEIJINGCITYLAB.ORG OR
LONGY.JIMDO.COM