

城市大数据挖掘：城市网络分析

Urban Network Analysis using Big Data

龙瀛，博士
清华大学建筑学院
2017年12月1日

Approaching the Human City: Beijing Studio
September 11 - 23, 2016
COURSE SYLLABUS

*A joint workshop between the Human Cities Initiative at Stanford University
& Tsinghua University Academy of Art and Design and the School of Engineering*

We invite you to participate in this experiment while we are holding class. Allow yourself to be fully present in the room, so you can listen to your classmates and what they have to say and share with you. Experience the freedom of not having to have your attention diverted or your mood instantly altered by whatever email or text message should come your way. You may take a phone call if it seems particularly urgent. But for most circumstances, we urge that you give yourself permission to be in control of your own time and energy— to actively choose where you want to direct your attention, as opposed to a portable device making that decision for you.

Please turn your mobile devices to “silent” or “do not disturb” mode, and do not take them out for the duration of the class. We strongly encourage you to take notes using pencil and paper— as research shows that this helps with memory retention— but If you must have a laptop to take notes, **please do not check your e-mail or browse the internet at any time.**

There will be plenty of opportunities to plug in once you leave the classroom. Let’s treat our classroom as a sacred space to enjoy the moment.

《大数据与城市规划》教学大纲

1. 大数据与城市规划概论（9月22日） W1
2. 变化中的中国城市与未来城市（9月29日） W2
3. 城市大数据类型与典型数据介绍（10月13日） W4（分组完成）
4. 城市大数据的获取（10月20日） W5（分发第一版本的北京老城数据、第一次课外沙龙）
5. 城市大数据的统计与分析（10月27日） W6（介绍中期汇报要求）
6. 城市大数据的可视化（11月3日） W7
7. 学生作业思路汇报与点评（11月10日） W8+W9（预计占用第二大节部分时间）

8. 基于图片大数据的城市空间研究（11月24日） W10
9. 城市大数据挖掘：空间句法（11月30日） W10+（亦为第二次课外沙龙）
10. **城市大数据挖掘：城市网络分析（12月1日） W11**
11. 数据增强设计（12月8日） W12
12. 战略、总体及详细规划中的大数据应用（12月15日） W13
13. 城市设计中的大数据应用（12月22日） W14
14. 大模型：跨越城市内与城市间尺度的大数据应用（12月29日） W15
15. 学生作业终期汇报与点评（1月5日） W16（一次或分两次）

上一堂课的回顾

- 空间句法概论
- 轴线分析与线段分析
- 空间句法软件Depthmap的实际操作

本讲大纲

- 城市网络分析概述
- 企业网络联系
- 交通网络联系
- 社交网络联系
- 城市规划领域的知识产出、消费与网络
- 基于兴趣点位置和名称的中国城市网络分析

城市空间新数据可以主要归纳为两类

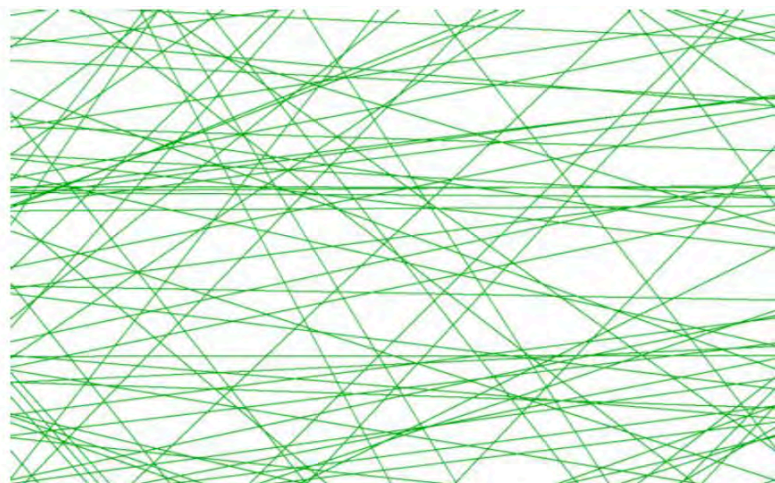
- **Density based**
- **Flow based**
 - 一个记录表示一次流动/联系 →
 - 一个记录表示多次流动/联系

3. Commute2016alpha 通勤OD(直线距离)2016

图层类型File Type : Polyline

数据量Amount : 220,084

图层重点Key Word : #起迄点 O/D
#出行次数统计



图层意义Meanings : 北京二环内之通勤路径(直线距离)

核心字段Key Elements :

1. OX, OY : 起点(Orientation)位置 · X坐标(经度)与Y坐标(纬度)
2. DX, DY : 迄点(Destination)位置 · X坐标(经度)与Y坐标(纬度)
3. people : 人数统计

基于联通数据的公里网格之间的通勤人流数量
OX/OY对应居住地的坐标
DX/Dy对应工作地的坐标

id	OX	OY	DX	DY	people	Shape_Leng	CommutelD
2	116.519	39.944	116.082	39.991	5	0.438979369	2

30. Weibos2016 微博发文地点2016

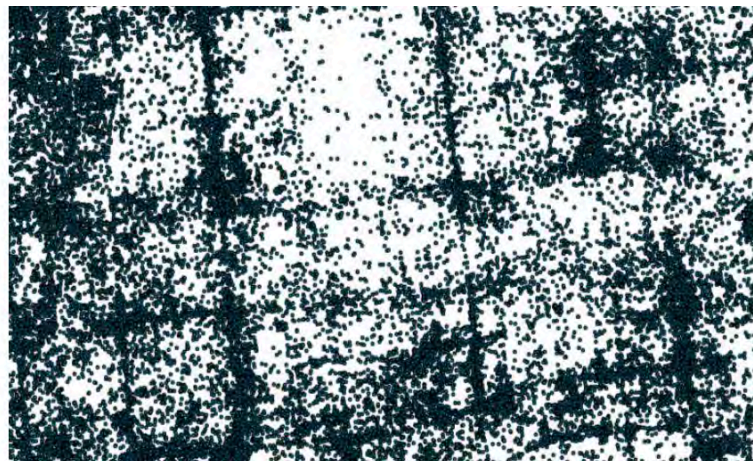
图层类型File Type : Point
数据量Amount : 947,020

图层重点Key Word : #位置 #发文内容 #按讚数量
#客户使用装置 #图片数量

图层意义Meanings : 北京二环内之微博发文地点

核心字段Key Elements :

1. text : 发文文本
2. SOURCE : 装置来源(iphone、小米、华为等)
3. FAVORITED : 点赞数量
4. USER_ID : 用户ID



可以根据每个人的不同时间的微博地点记录推导对应的轨迹（但是并没有考虑两次签到之间的时间关系）

CREATED_AT	TEXT	SOURCE	FAVORITED	PIC_COUNT	USER_ID	CLIENT_MBL
2015-04-24 01:27:00+08	飞机晚点35分钟·机场高速夜间维修保养堵车18分钟·天宁寺桥凌晨还查酒驾堵车15分钟...眼瞅着一辆河北牌司机被请下车带走了。 ..	iPhone 6	0	0	1341199832	iPhone-975F6618-8B8D-44B3-B38E-78FF35711172

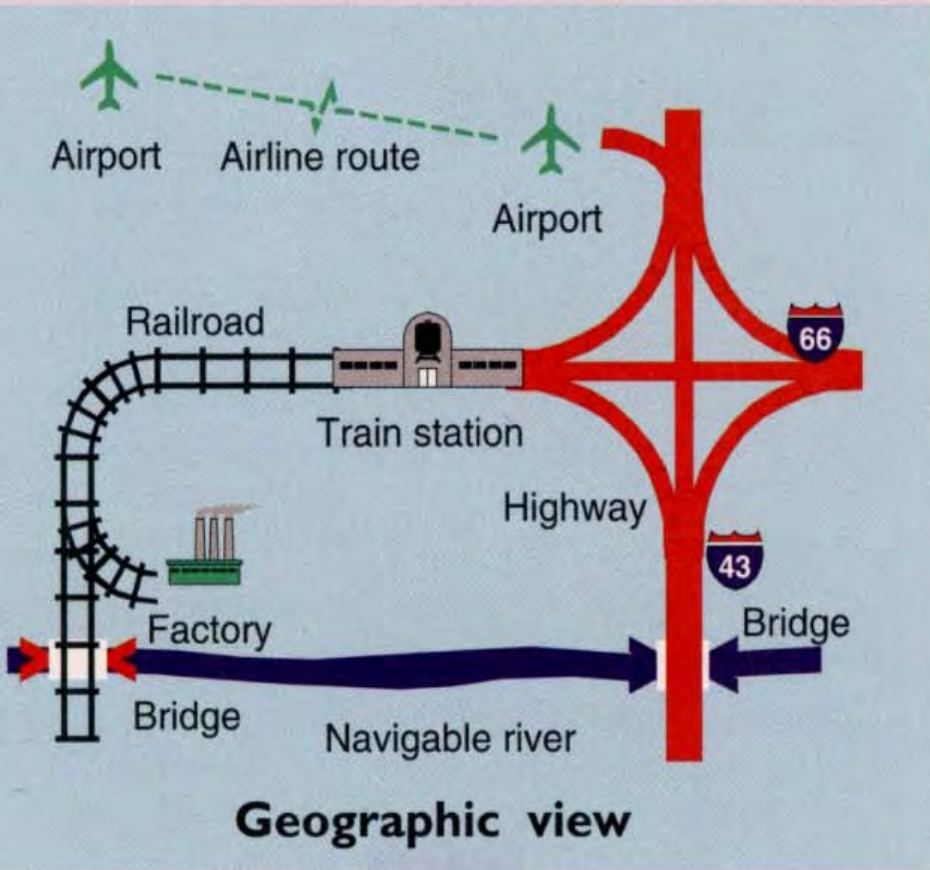
共享数据中的flow based的相关数据

- 原始记录数据量较大，建议在更大的空间尺度（如交通分析小区、街区等）汇总后，进行分析、统计和可视化
 - 需要构建点与更大的空间尺度的单元之间的关系（spatial join?）
- 可以用于表达研究范围内不同空间单元之间的联系

一、城市网络分析概述

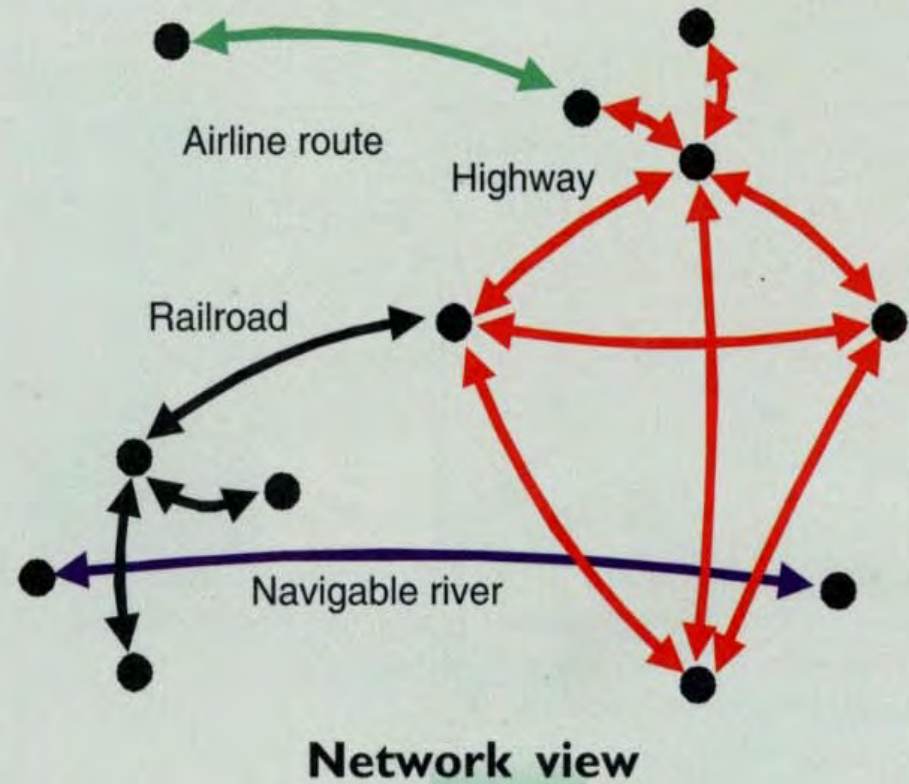
Two views of a network

You can view a network as a collection of geographic objects such as rails, roads, stations, and bridges and also as a pure network of edges and junctions.



A geometric network is the representation of geographic features that comprise a network.

Geometric network

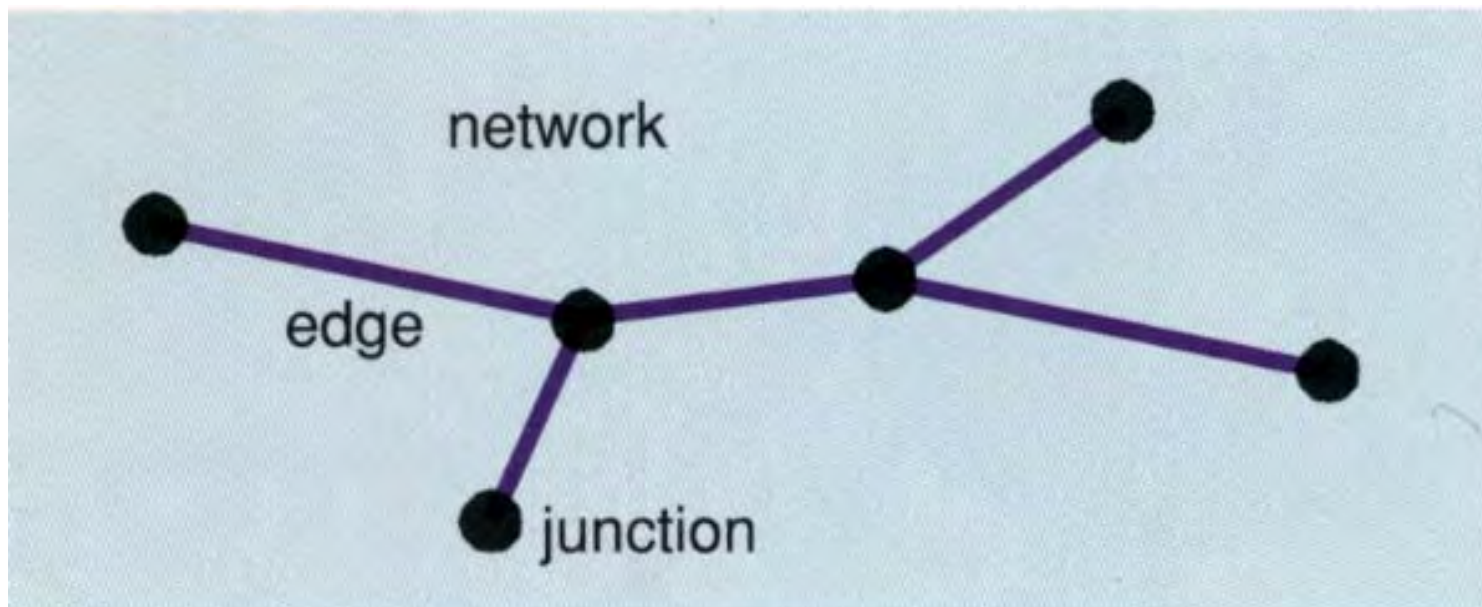


A logical network is a pure graph of junction elements and edge elements.

Logical network

网络图论

- $G=(V,E)$ ，其中 V 是点集， E 是边集。
- 网络是由节点与边组成的。



基本的统计特征

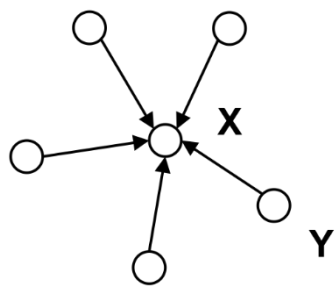
• 度分布(degree distribution)

- 节点 v_i 的度 k_i 定义为与该节点相连接的边的数目。直观上看，一个节点的度越大就意味着这个节点在某种意义上越“重要”。网络中所有节点 v_i 的度 k_i 的平均值称为网络的(节点)平均度，记为 $\langle k \rangle$ 。
- 网络中节点的度的分布情况可用分布函数 $p(k)$ 来描述。度分布函数反映了网络系统的宏观统计特征。 $p(k)$ 表示的是一个随机选定的节点的度恰好为 k 的概率分布。
- **出度和入度**的相关研究解释：
 - 在社会网络应用中，出度是扩张性 (expansiveness)的测度，而入度则是接收性 (receptivity)或受欢迎程度 (popularity)的测度。如果我们考虑友谊关系的社会计量关系，一个有较大出度的行动者是喜欢结交朋友的人，一个有较小出度的行动者结交的朋友较少。一个有较大入度的行动者是很多人都喜欢与之交往的人，而一个有较小入度的行动者被较少的人选作朋友。
 - 在城市网络应用中，点出度、点入度分别代表了各城市的吸引力、控制力。点出度为从该点出发与其直接联系的点数，点入度为到达该点并与该点直接联系的点数。

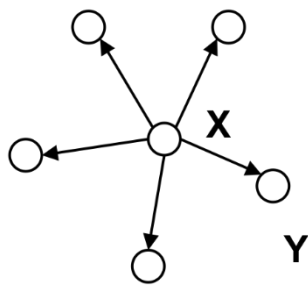
基本的统计特征

- 介数 (betweenness)

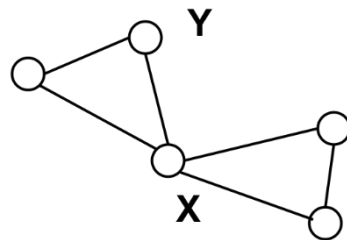
- 介数反映了节点或边的作用和影响力，如果一对节点间共有 B 条不同的路径，其中有 b 条经过节点 i ，那么节点 i 对这对节点的介数的贡献为 b/B 。把节点 i 对所有节点对的贡献累加起来，再除以节点对总数，就可得到节点 i 的介数。
- 类似的，边的介数定义为所有节点对的最短路径中经过该边的数量比例。
- 介数越大 \rightarrow 经过该节点/边的最短路径越多 \rightarrow 经过该节点/边的信息量越大



入度



出度



介数

常用的中心性方法

- 介数中心性方法(betweenness centrality C^B):通过假设信息仅仅沿着最短路径来传播.如果 n_{jk} 是连接节点 j 和 k 之间最短路径数目, $n_{jk}(i)$ 是连接节点 j 和 k 之间包含着节点的最短路径数目, 结果节点 i 的介数中心性定义为

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j \in G, j \neq i, k \in G} \sum_{k \neq i, k \neq j} n_{ik(i)} / n_{jk}$$

- 邻近中心性方法 (Closeness centrality C^C): 一个节点 i 的邻近中心性是基于最小距离或最短路径 d_{ij} 来定义的, 即从节点 i 到节点 j 穿过边的最小数目, 并被定义为

$$C_i^C = (L_i)^{-1} = \frac{N-1}{\sum_{i \in G} d_{ij}}$$

从 i 到所有其他节点的平均距离

- Bonacich's 特征向量中心性方法(Bonacich's eigenvector centrality C^λ):基于一个节点的中心性值是由邻接顶点的值决定这样一种思想, 这种方法不仅考虑到它的邻居节点中心性值, 而且考虑到一个节点在图中的位置.节点 i 的特征向量中心性定义为

$$C_i^\lambda = \sum_{j=1}^N A_{ij} C_j^\lambda$$

- 这就直接引导出了众所周知的特征向量计算问题 $\lambda S = AS$, 并且最大特征值对应的特征向量就是这个向量中心性($C^\lambda = S$).

常见的网络分析数据源

- 传统数据（多为城市间尺度，城市-城市）
 - 人口普查公报中的人口迁移矩阵
 - 居民交通出行调查（城市内尺度）
 - 城际交通（航班、火车、长途客车等）
 - 公司的总部与分支关系
 - 文献统计（如作者所在城市与其研究的城市）
- 新数据（城市间与城市内皆有）
 - 出租车轨迹
 - 公共交通刷卡记录
 - 共享交通工具轨迹（如滴滴与摩拜）
 - 手机信令（位置移动与通讯联系）
 - 社交媒体（不同时刻的签到位置）
 - 兴趣点（所在城市与和名称所体现的城市）

城市网络分析的主要步骤



- 依托大数据的城市网络分析主要分为数据获取、数据处理、数据分析及结论应用四个步骤。

常用的网络分析工具

- ArcGIS / Python (本次课堂中)
- Convert a feature into point (Feature to Point)
- Calculate the coordinates of points (Calculate Geometry in ArcMap)
- Convert tabulate flows into Lines (XY to Line)
- Analyze and visualize the Lines in ArcGIS

Welcome to GaWC - the leading academic thinktank on cities in globalization

[The World According to GaWC](#) [GaWC Books](#)

Tweets by @GaWC

GaWC @GaWC
How not to fry the grandchildren? A response to Peter Taylor's 'cities in climate change' [dx.doi.org/10.1080/122659...](https://doi.org/10.1080/122659...)

23h

GaWC Retweeted
Urban Studies Jnl @USJ_online
Free access till March 2017: World City Typologies and National City System Deterritorialisation: USA, China, Japan ow.ly/gCur304PUoB

24 Oct

GaWC Retweeted
Sergio Montero @sergemont
The new special issue of @urbgeog: An attempt to de-center "gentrification" from the experience of Latin America: tandfonline.com/toc/rurb20/37/8

Embed

[View on Twitter](#)

About GaWC

[What's New](#)

[Projects](#)

[Research Bulletins](#)

[Data](#)

[Media](#)

Gateways into GaWC

[New Visitors](#)

[Researchers](#)

[Practitioners](#)

[Teaching Resource](#)

[Visualisation](#)

[Search GaWC](#)

We are proud to be part of the global intellectual commons - please explore!

You can navigate through GaWC in two ways. Either use the links above to go straight to the content pages that interest you; or use the [Gateways](#) to see a selection of resources appropriate to your particular needs.

Honorary Founders

Manuel Castells [University of Southern California](#)
John Friedmann [University of British Columbia](#)
Sir Peter Hall [University College London](#)
Saskia Sassen [Columbia University](#)
Sir Nigel Thrift [University of Warwick](#)

Director

[Peter Taylor](#) (Northumbria)

Associate Directors

[David Bassens](#) (Brussels)

[Jon Beaverstock](#) (Bristol)

[Martijn Burger](#) (Rotterdam)

[Ben Derudder](#) (Ghent)

[James Faulconbridge](#) (Lancaster)

[John Harrison](#) (Loughborough)

[Michael Hoyler](#) (Loughborough)

[Xingjian Liu](#) (Hong Kong)

[Zachary Neal](#) (Michigan State)

[Kathy Pain](#) (Reading)

[Christof Pamreiter](#) (Hamburg)

[Allan Watson](#) (Loughborough)

[Frank Willox](#) (Ghent)

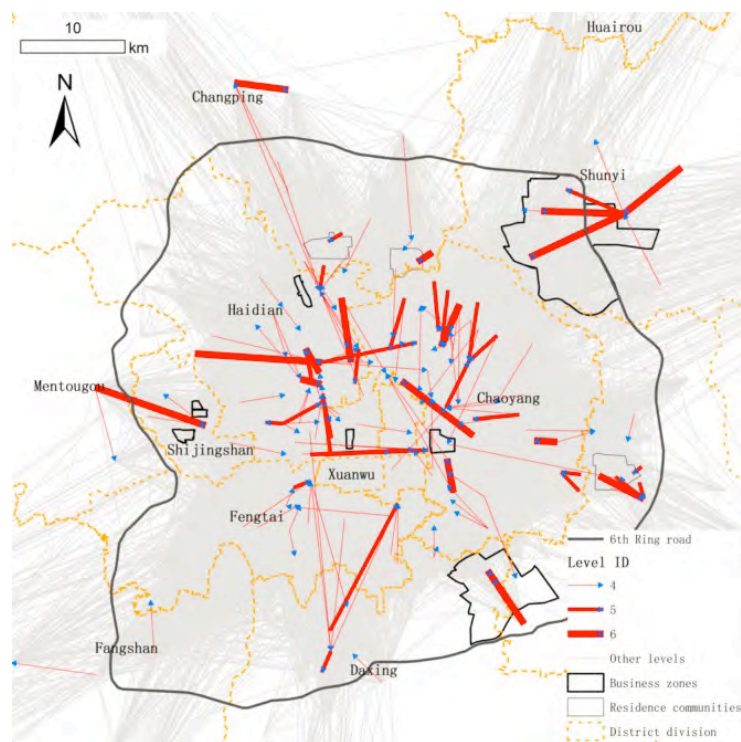
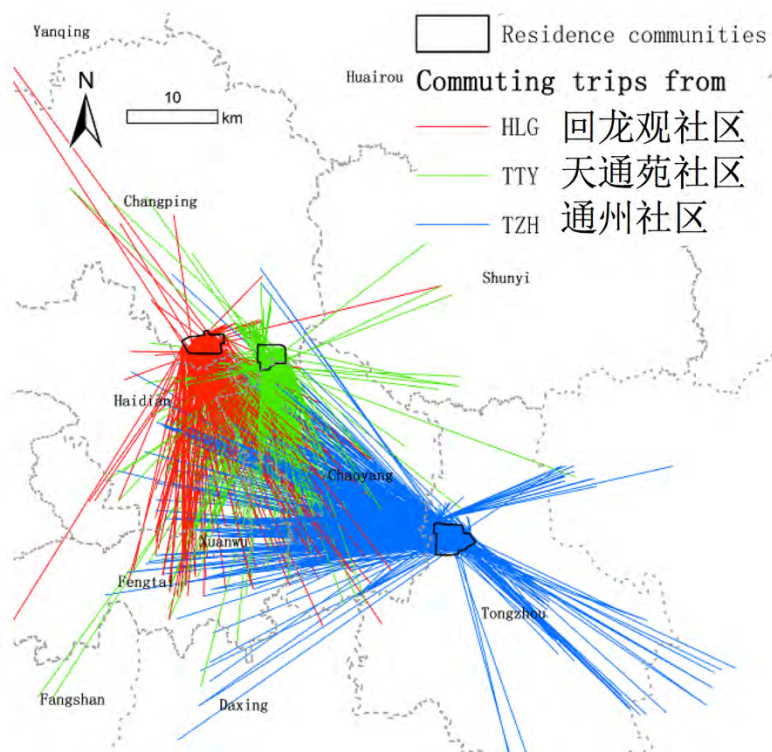
[Email GaWC](#)



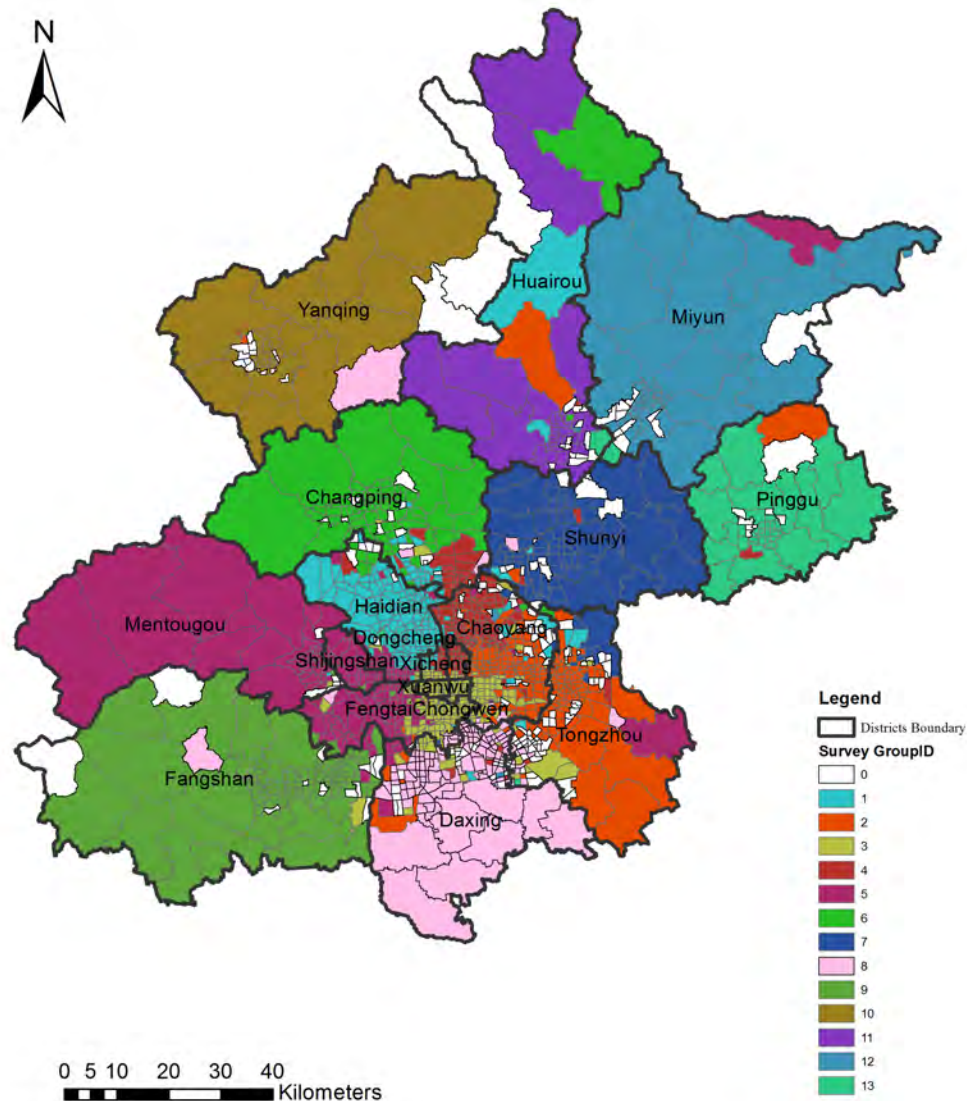
- Globalization and World Cities Research Network
- <http://www.lboro.ac.uk/gawc/>
- GaWC的中国拥护者们

城市内部网络

- 城市交通流数据是主要的数据源，多为有向流数据（有向网络）
 - 评价所关注区域与周边的联系（左图）
 - 评价主导的联系方向/边（右图）
 - 评价主要的联系节点



- 社区识别 (Community detection) 是有效的基于流数据对空间进行分割的方法
- In the study of complex networks, a community is defined as a subset (group) of the whole network and the nodes in the same community are densely connected internally and grouped together. The identification of such densely connected nodes in networks is called community detection.
- Popular community detection methods can be classified into two groups: graph partitioning and hierarchical clustering. Graph partitioning divides a network graph into a set of non-overlapping groups, while hierarchical clustering seeks to build a hierarchy of clusters of nodes, such that for each cluster there are more internal than external connections.



The spatial distribution of community detection results of one-day household survey data

二、企业网络联系

基于企业普查数据的城市产业联系分析

感谢华南理工大学赵渺希教授对本部分课件的支持

• 企业总分支机构

湖北省有色金属企业联营公司武汉经营部	武汉市	江汉区	湖北省有色金属企业联营公司	武汉市	青山区	6373
湖北省科学器材公司附属科技开发经营部	武汉市	江汉区	湖北省科学器材公司	荆门市	钟祥市	6571
中国电子进出口湖北分公司	武汉市	武昌区	中国电子进出口湖北分公司	荆门市	钟祥市	6399
湖北九隆文化科技有限公司广告分公司	武汉市	洪山区	湖北九隆文化科技有限公司	荆门市	钟祥市	7440
湖北天王电器有限责任公司中北路分公司	武汉市	武昌区	湖北天王电器有限责任公司	荆门市	钟祥市	6571
长江轮船海外旅游总公司(武汉)分公司	武汉市	江汉区	长江轮船海外旅游总公司	武汉市	江汉区	6610
武汉市华中衡器厂经营部	武汉市	江汉区	武汉市华中衡器厂	武汉市	江汉区	6373
武汉市联碱化工公司经营部	武汉市	江汉区	武汉市联碱化工公司	武汉市	武昌区	6369
武汉市普爱综合经营部	武汉市	硚口区	武汉市普爱综合公司	武汉市	武昌区	6353
武汉市武船华泰经营部	武汉市	武昌区	武汉市武船华泰公司	武汉市	武昌区	6599
武汉中力物流有限公司武昌南分公司	武汉市	武昌区	武汉中力物流有限公司	武汉市	武昌区	5720
武汉市单洞物回经营部	武汉市	江汉区	武汉市单洞物回公司	武汉市	武昌区	6539
武汉市商医综合经营部	武汉市	江汉区	武汉市商医综合公司	武汉市	武昌区	6529
江汉石油管理局驻武汉办事处招待所	武汉市	汉阳区	江汉石油管理局驻武汉公司招待所	武汉市	江汉区	6690
湖北印刷厂分厂	武汉市	武昌区	湖北印刷厂	荆门市	钟祥市	2320

分公司

总公司

行业代码

- 筛选出带有分公司、分厂等字眼的企业
- 字段处理生成企业总部名称
- 地址解析，抓取总公司与分公司所在的区县
- 形成表

• 武汉城市圈企业矩阵表

企业总部

企业分支

	安陆市	赤壁市	崇阳县	大悟县	大冶市	鄂州市	汉川市	红安县	黄梅县	黄石市	黄州区	嘉鱼县	罗田县	麻城市	蕲春县	潜江市	天门市	通城县	通山县	团风县	武汉市	
安陆市	140	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
赤壁市	0	227	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
崇阳县	0	12	25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
大悟县	1	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
大冶市	0	1	0	0	302	0	0	0	0	32	0	0	0	0	0	0	0	0	0	0	0	0
鄂州市	0	0	0	0	1	269	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	7
汉川市	0	0	0	0	0	0	141	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
红安县	0	0	0	0	0	0	0	21	0	0	0	0	0	3	0	0	0	0	0	0	0	0
黄梅县	0	0	0	0	0	0	0	0	36	0	1	0	0	0	0	0	0	0	0	0	0	0
黄石市	0	3	0	0	21	2	0	0	0	395	0	0	0	0	1	1	0	0	0	0	0	9
黄冈市	0	0	0	0	0	4	0	0	0	0	151	0	6	0	1	0	0	0	0	0	0	1
嘉鱼县	0	3	0	0	0	0	0	0	0	0	0	74	0	0	0	0	0	0	0	0	0	4
罗田县	0	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0
麻城市	0	0	0	1	0	0	0	0	0	0	0	0	0	124	0	0	0	0	0	0	0	0
蕲春县	0	0	0	0	0	0	0	0	0	0	1	0	0	0	39	0	0	0	0	0	0	0
潜江市	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	237	3	0	0	0	0	23
天门市	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	7	284	0	0	0	0	1
通城县	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0
通山县	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0
团风县	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	8	0	1
武汉市	3	3	3	2	3	3	10	2	0	1	3	5	0	1	0	1	2	1	1	0	0	2378
武穴市	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
浠水县	0	0	0	0	1	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0
仙桃市	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
咸宁市	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3
孝昌县	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
孝感市	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
阳新县	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
英山县	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
应城市	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
云梦县	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

数据分析

Network Analyst

OD 成本矩阵

- 起始点 (O)
- 目的地 (D)
- 线 (O)
- 点障碍 (O)
- 禁止型 (O)
- 成本增加型 (O)
- 线障碍 (O)
- 禁止型 (O)
- 成本按比例增加型 (O)
- 面障碍 (O)
- 禁止型 (O)
- 成本按比例增加型 (O)

加载位置

加载自: 城市点

仅显示点图层

仅加载选定行(S)

排序字段(O): FID

位置分析属性

属性	字段	默认值
Name		
Target/DestinationCount		
CurbApproach		车辆的任意一侧
Cutoff_长度		

位置定位

使用几何(G)

搜索容差(T): 5000 米

使用网络位置字段(N)

属性	字段
SourceID	
SourceOID	
PosAlong	
SideOfEdge	

高级(A)... 关于加载位置 确定 取消

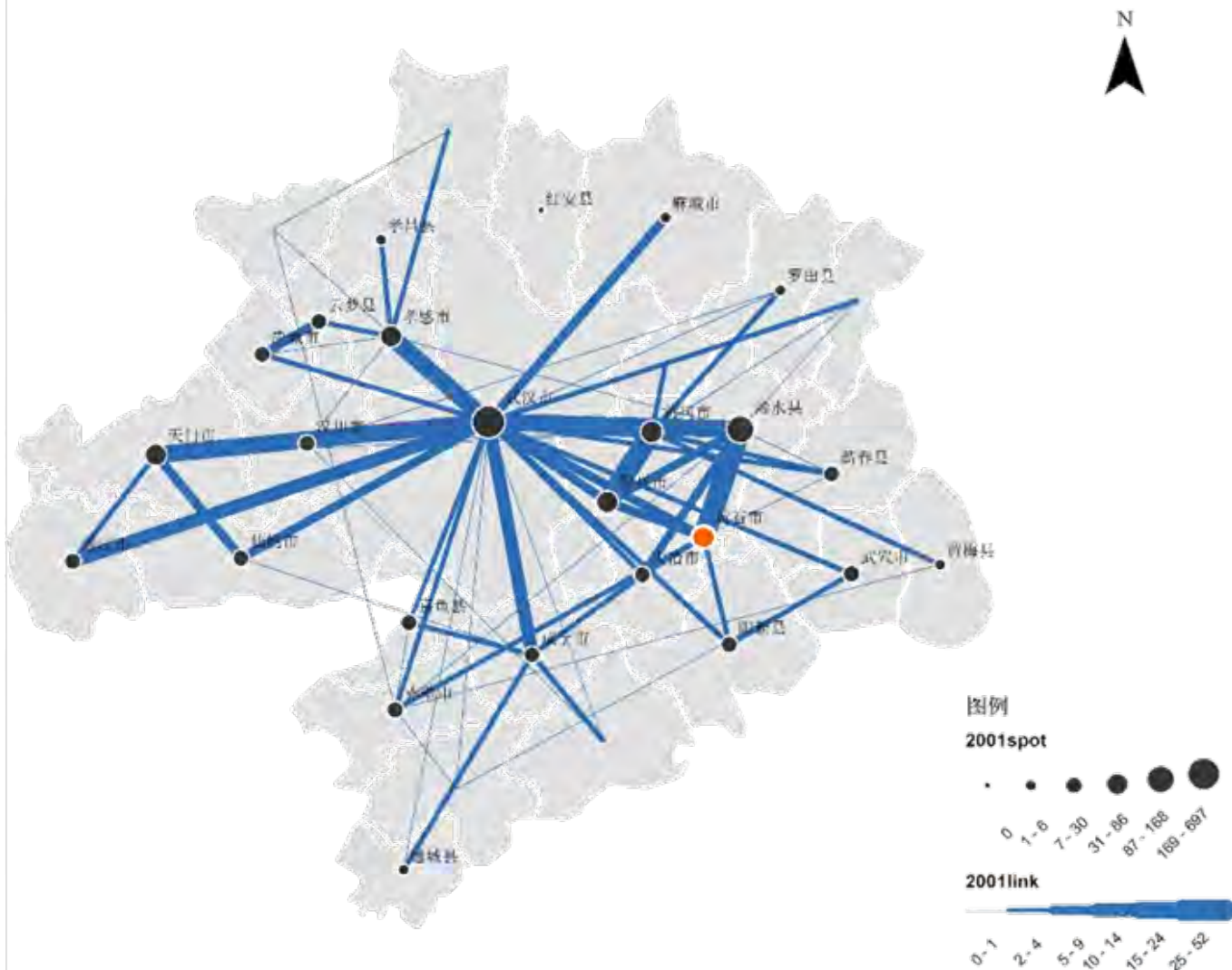
1024928.595 4038958.161 未知单位

数据分析

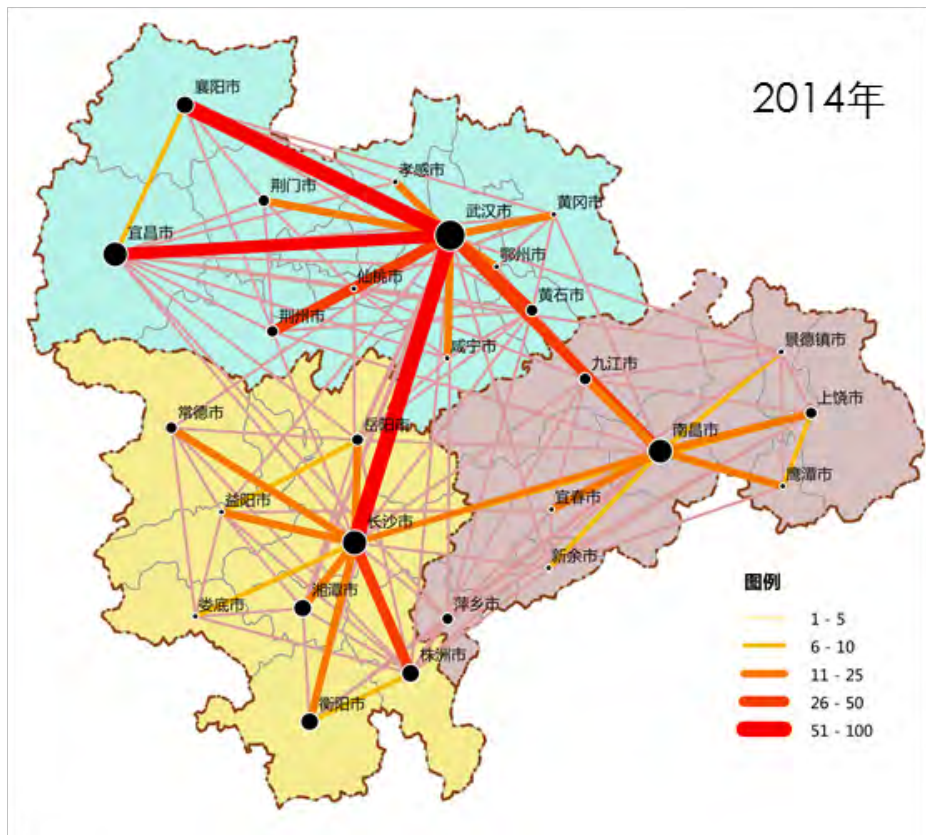
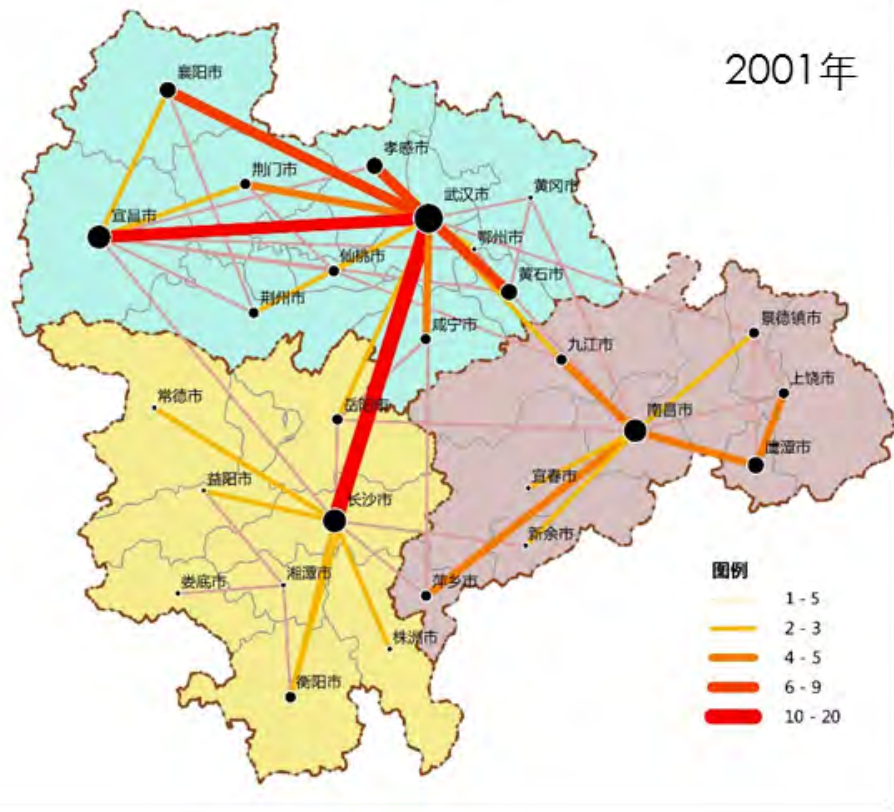
	位置 21	位置 22	位置 23	位置 24	位置 28	位置 31	位置 25	位置 2	位置 3	位置 30	位置 18	位置 4	位置 5	位置 6
位置 21 安陆市	140													
位置 22 赤壁市		227	1											
位置 23 崇阳县		12	25											
位置 24 大悟县	1			21										
位置 28 大冶市		1			302					32				
位置 31 鄂州市					1	269				1	2			
位置 25 汉川市							141							
位置 2 红安县								21						3
位置 3 黄梅县									36			1		
位置 30 黄石市		3			21	2				395				
位置 18 黄冈市						4					151		6	
位置 4 嘉鱼县		3										74		
位置 5 罗田县													40	
位置 6 麻城市				1										124
位置 7 蕲春县											1			
位置 8 潜江市														
位置 9 天门市							2							
位置 1 通城县														
位置 10 通山县														
位置 17 团风县											2			
位置 27 武汉市	3	3	3	2	3	3	10	2		1	3	5		1
位置 19 武穴市			1											
位置 11 浠水县					1					1	2			
位置 12 仙桃市														
位置 13 咸宁市						2						2		
位置 26 孝昌县														
位置 20 孝感市	1													
位置 29 阳新县								1						
位置 15 英山县											1			
位置 14 应城市												1		
位置 16 云梦县													1	

位置	联系度
位置 21 - 位置 21	140
位置 21 - 位置 22	0
位置 21 - 位置 23	0
位置 21 - 位置 24	0
位置 21 - 位置 28	0
位置 21 - 位置 31	0
位置 21 - 位置 25	0
位置 21 - 位置 2	0
位置 21 - 位置 3	0
位置 21 - 位置 30	0
位置 21 - 位置 18	0
位置 21 - 位置 4	0
位置 21 - 位置 5	0
位置 21 - 位置 6	0
位置 21 - 位置 7	0
位置 21 - 位置 8	0
位置 21 - 位置 9	0
位置 21 - 位置 1	0
位置 21 - 位置 10	0
位置 21 - 位置 17	0
位置 21 - 位置 27	1
位置 21 - 位置 19	0
位置 21 - 位置 11	0
位置 21 - 位置 12	0
位置 21 - 位置 13	0
位置 21 - 位置 26	0
位置 21 - 位置 20	0
位置 21 - 位置 29	0
位置 21 - 位置 15	0





武汉城市圈2001年企业分支机构图



2001年—2014年长江中游城市群企业网络演化图

- 长江中游城市群的企业关联网
 - “一核独大”向“三足鼎立”
 - 黄石

三、交通网络联系

基于航空班次数据的城市对外航空网络联系

机票预订 > 机票查询 > 航班查询

航班时刻查询

航线

出发城市

换

航班号

到达城市

出发城市

查询

到达城市

热门航班推荐

广州 → 长沙

起抵时间: 22:35 - 23:50

南方航空 CZ3644 ¥ 260起

广州 → 梅州

起抵时间: 20:20 - 21:20

南方航空 CZ3337 ¥ 210起

广州 → 北海

起抵时间: 10:45 - 12:00

南方航空 CZ3353 ¥ 210起

广州 → 南宁

起抵时间: 22:45 - 00:15 ±1

南方航空 CZ3298 ¥ 160起

广州 → 湛江

起抵时间: 06:40 - 07:50

南方航空 CZ3323 ¥ 350起

广州 → 晋江

起抵时间: 06:50 - 08:15

昆明航空 KY9659 ¥ 411起

国内城市出发航班时刻

国际航班时刻表 >>

按拼音字母查询 A B C D E F G H J K L M N P Q R S T U W X Y Z

A 阿尔山航班 阿克苏航班 阿拉善右旗航班 阿拉善左旗航班 阿勒泰航班 阿里航班 安康航班 安庆航班 鞍山航班 安顺航班



B 百色航班 保山航班 包头航班 巴彦淖尔航班 北海航班 北京航班 毕节航班 博乐航班

C 长白山航班 长春航班 常德航班 昌都航班 长沙航班 长治航班 常州航班 朝阳航班 成都航班 赤峰航班 池州航班 重庆航班

D 大理市航班 大连航班 丹东航班 稻城航班 大庆航班 大同航班 达县航班 大庸航班 达州航班 德令哈航班 迪庆航班 东营航班 敦煌航班

E 鄂尔多斯航班 额济纳旗航班 二连浩特航班 恩施航班

F 佛山航班 阜阳航班 抚远航班 富蕴航班 福州航班

- 数据透视构造航班间城市的联系矩阵
 - 横列  目的城市
 - 纵列  始发城市

- 原理
 - 通过航班的始发站与终点站分别所在的城市来量化城市间的航空交通联系度。
 - 对各行各列的数据进行汇总，作为该区县与其他区县的联系总量。

数据分析

Network Analyst

OD 成本矩阵

- 起始点 (O)
- 目的地点 (D)
- 线 (O)
- 点障碍 (O)
 - 禁止型 (O)
 - 成本增加型 (O)
- 线障碍 (O)
 - 禁止型 (O)
 - 成本按比例增加型 (O)
- 面障碍 (O)
 - 禁止型 (O)
 - 成本按比例增加型 (O)

加载位置

加载自: 城市点

仅显示点图层

仅加载选定行(S)

排序字段(O): FID

位置分析属性

属性	字段	默认值
Name		
TargetDestinationCount		
CurbApproach		车辆的任意一侧
Cutoff_长度		

位置定位

使用几何(G)

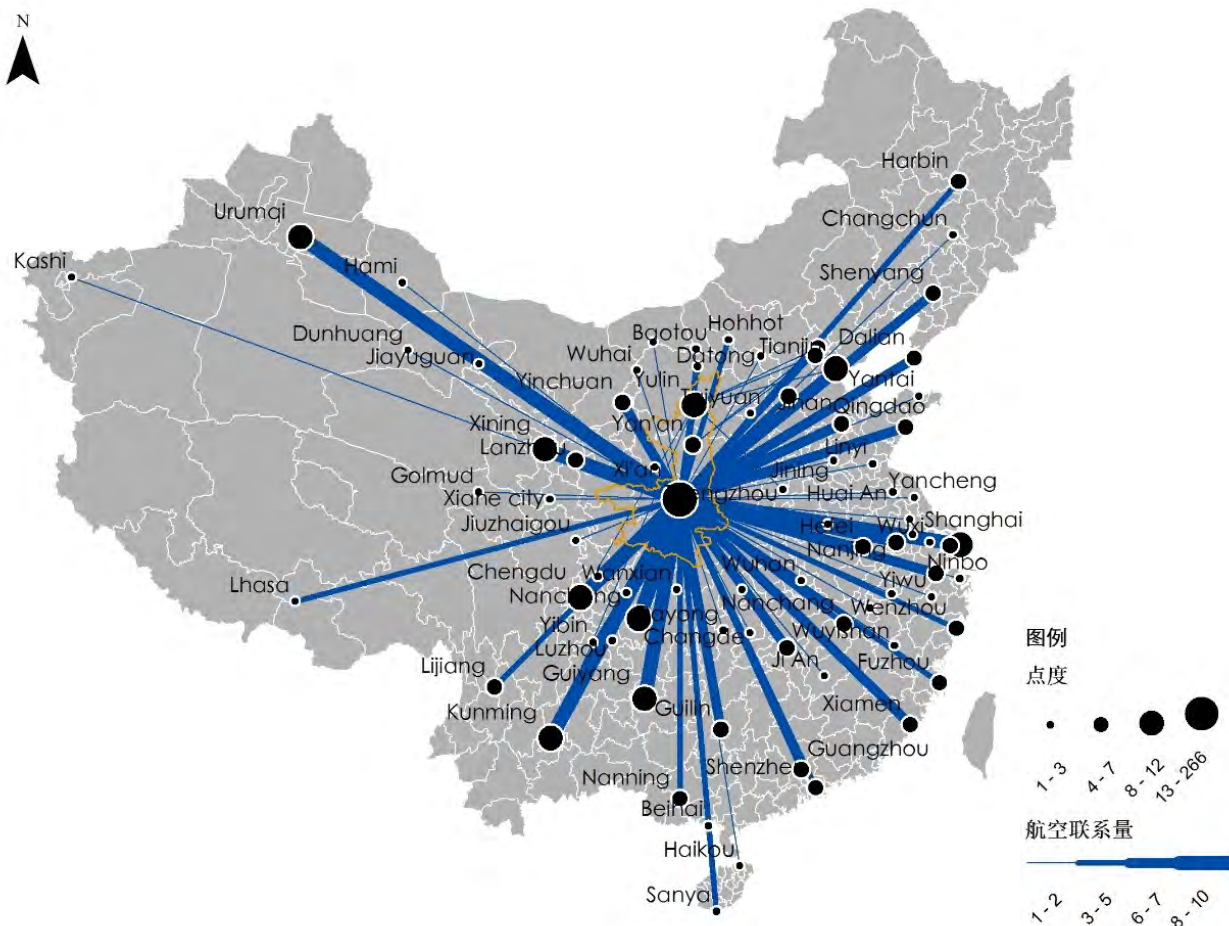
搜索容差(T): 5000 千米

使用网络位置字段(N)

属性	字段
SourceID	
SourceOID	
PosAlong	
SideOfRoad	

高级(A)... 关于加载位置 确定 取消

1024928.595 4038958.161 未知单位



关天城市群对全国的航空网络图

- 关中-天水城市群

- 航空辐射主要以**京津冀、长三角以及成渝地区**城市为主；
- 与**乌鲁木齐**的联系较为密切，与**东南沿岸**城市关系较弱。

四、社交网络联系

基于新浪微博位置数据的区域城市联系

获取Access Token

请求URL: https://api.weibo.com/2/place/nearby/pois.json

位置服务普通读取接口

place/nearby/pois

GET POST

点击获取文档

lat: 39.98437

lon: 116.30987

请求: 折叠请求

请求方式: get

请求URL: https://api.weibo.com/2/place/nearby/pois.json

请求参数:

lat=39.98437&long=116.30987&access_token=2.009ac2ICdLUHVE0cc11d8f33RgPJOD

返回的内容:

FAQ | 常见错误代

```
{
  "pois": [
    {
      "poiid": "B2094757D06FA3FD4499",
      "title": "南京大牌档(中关村店)",
      "address": "海淀区中关村大街5号中关村广场购物中心D区津乐汇2楼(近鼎好大厦)",
      "lon": "116.3134320120655",
      "lat": "39.98291860614086",
      "category": "75",
      "city": "0010",
      "province": null,
      "country": "80086",
    }
  ]
}
```

注册账号 → 接口进入 → 设置参数 → 抓取数据



所在地



微博发布者



来自地

原理：一条微博信息产生一对城市间的联系

数据分析

Network Analyst

OD 成本矩阵

- 起始点 (O)
- 目的地点 (D)
- 线 (O)
- 点障碍 (O)
- 线障碍 (O)
- 面障碍 (O)

加载位置

加载自: 城市点

排序字段(O): FID

属性	字段	默认值
Name		
TargetDestinationCount		
CurbApproach		车辆的任意一侧
Cutoff_长度		

位置定位

使用几何(G)

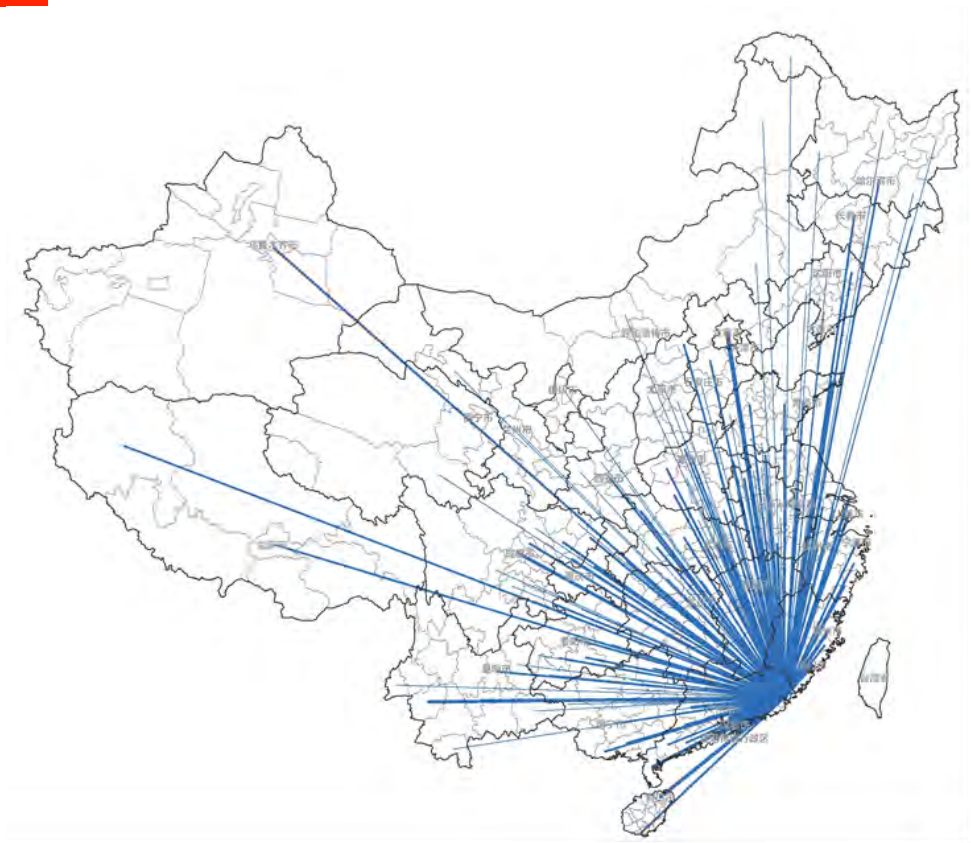
搜索容差(T): 5000 米

使用网络位置字段(N)

属性	字段
SourceID	
SourceOID	
PosAlong	
SideOfEdge	

高级(A)... 关于加载位置 确定 取消

1024928.595 4038958.161 未知单位



序号	城市	微博数量	序号	城市	微博数量
1	广州市	3236	16	中山市	73
2	深圳市	1786	17	肇庆市	71
3	北京市	1004	18	武汉市	68
4	香港	334	19	湛江市	62
5	东莞市	206	20	珠海市	56
6	惠州市	203	21	成都市	55
7	汕尾市	183	22	河源市	53
8	梅州市	142	23	临沧市	50
9	佛山市	121	24	海口市	47
10	厦门市	109	25	泉州市	43
11	重庆市	103	26	福州市	41
12	上海市	100	27	江门市	37
13	上海市	100	28	澳门	36
14	清远市	92	29	茂名市	33
15	赣州市	73	30	长沙市	30

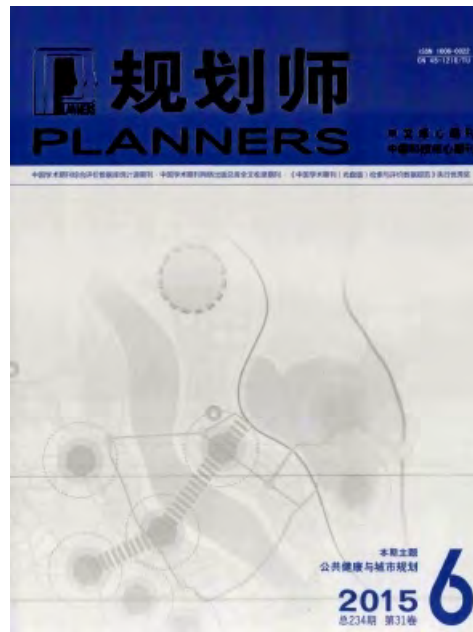
潮汕地区外来人口流动分析

- 潮汕地区的外来人口构成情况来看
 - 来自广州的人口数量最多
 - 其次是深圳和北京
 - 此外香港、东莞、惠州、汕尾等地也有一定的数量

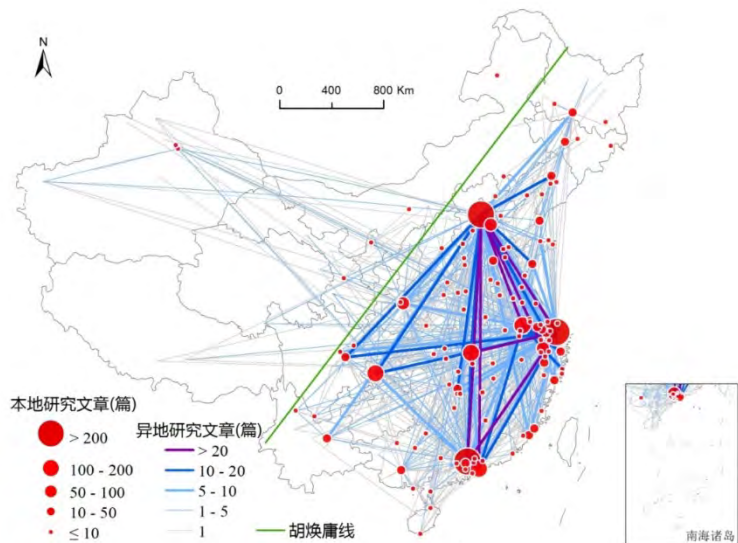
五、中国城市规划领域的知识产出、消费与网络

基于2000-2015城市规划四大期刊的分析

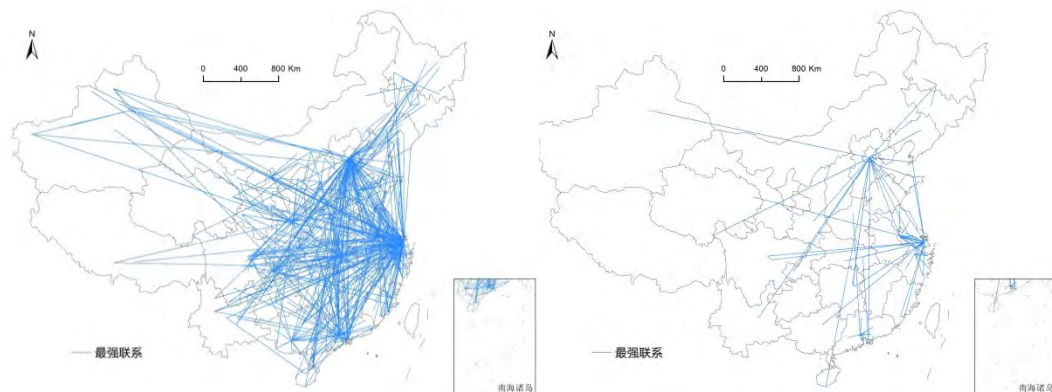
- **知识的产出和消费**是后工业社会的中心问题
 - **知识的消费地区**：大部分城市规划研究以实际项目为依托，有指定的研究区域。将被研究的地区视为城市规划知识的消费地区；
 - **知识的产出者**：与研究项目相关的高校科研工作者、规划设计院/所、规划管理部门、咨询公司的研究人员。
- 国内，城市规划领域已有文献计量研究不足
 - **没能充分利用**互联网丰富的信息资源；
 - **不足以清晰的刻画**出城市规划研究的区域规律；
 - 与其他领域的计量分析**较为相似**。
- 国外，Matthiessen等人利用计量文献分析的方法，对全球知识网络展开了一系列的研究
 - 首先分析了欧洲的研究中心，探索了世界研究中心的变化规律；
 - 然后通过合作关系和论文的引用关系，分析了全球知识网络格局。
- **总体上**，国内学者对于能刻画城市规划知识产出、消费的研究者与被研究城市的空间规律的探究，还有待深入。



城市网络规律

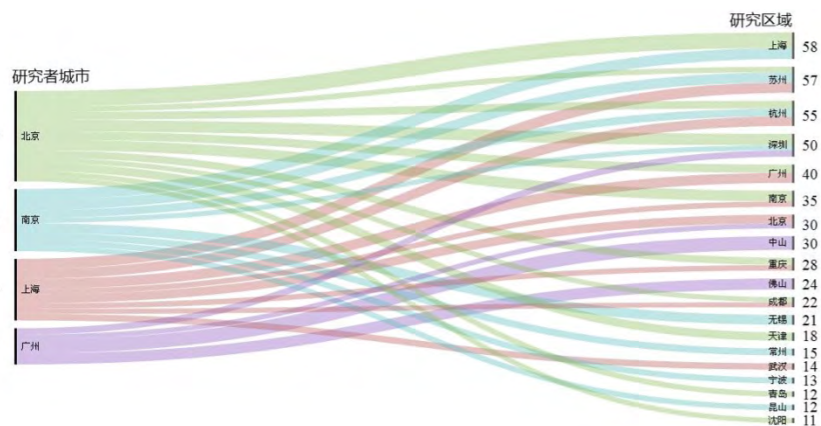


知识产出—消费网络格局（无向连接）

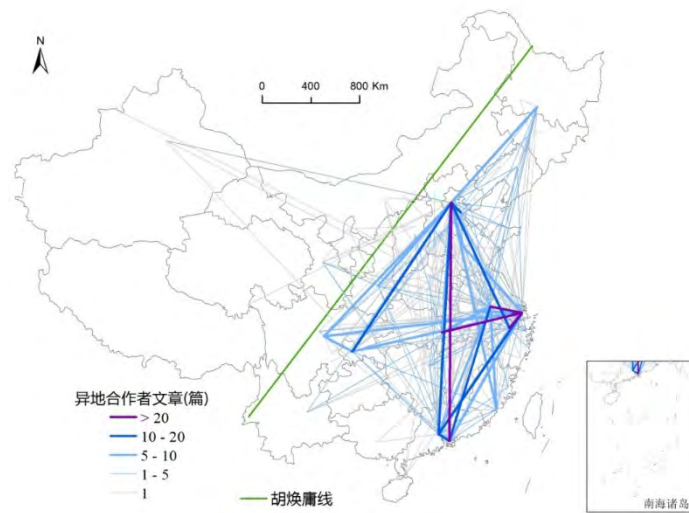


a. 所有最强联系网络 b. 最强联系大于3的网络

知识产出-消费最强联系网络



研究者—被研究城市有向连接图

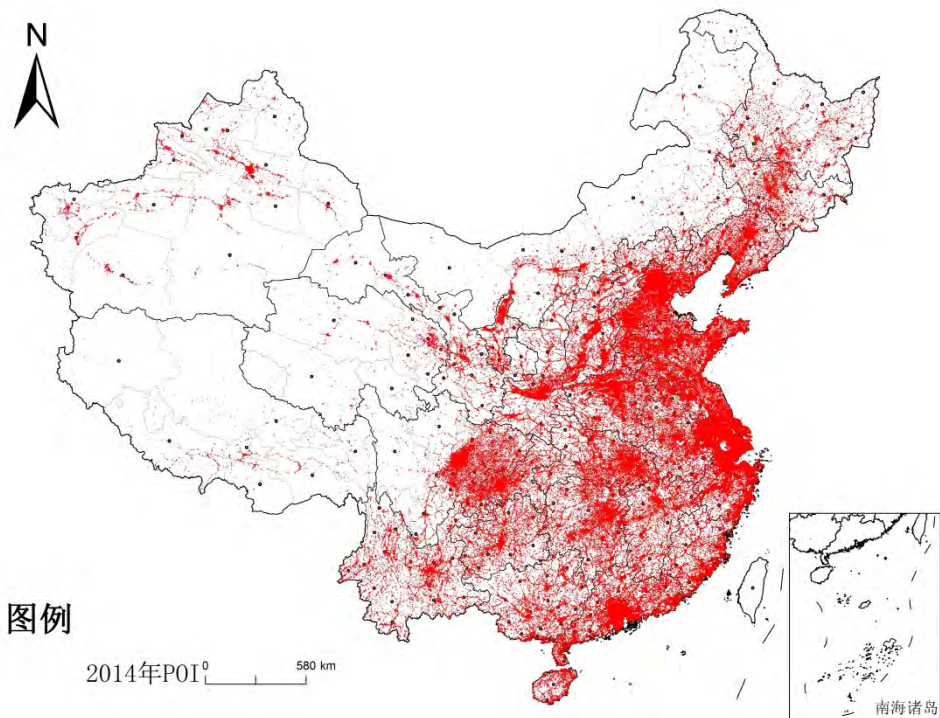
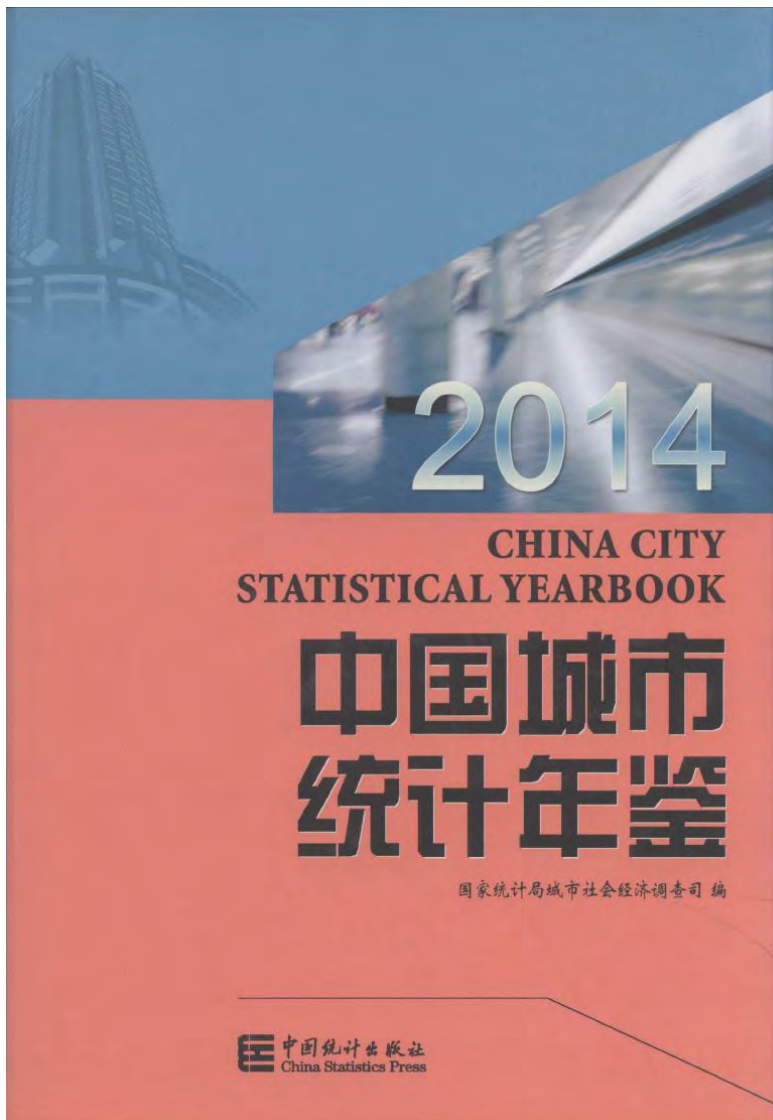


合作者城市联系强度图

六、基于兴趣点位置和名称的中国城市网络分析

课外阅读材料（辅以共享的论文）

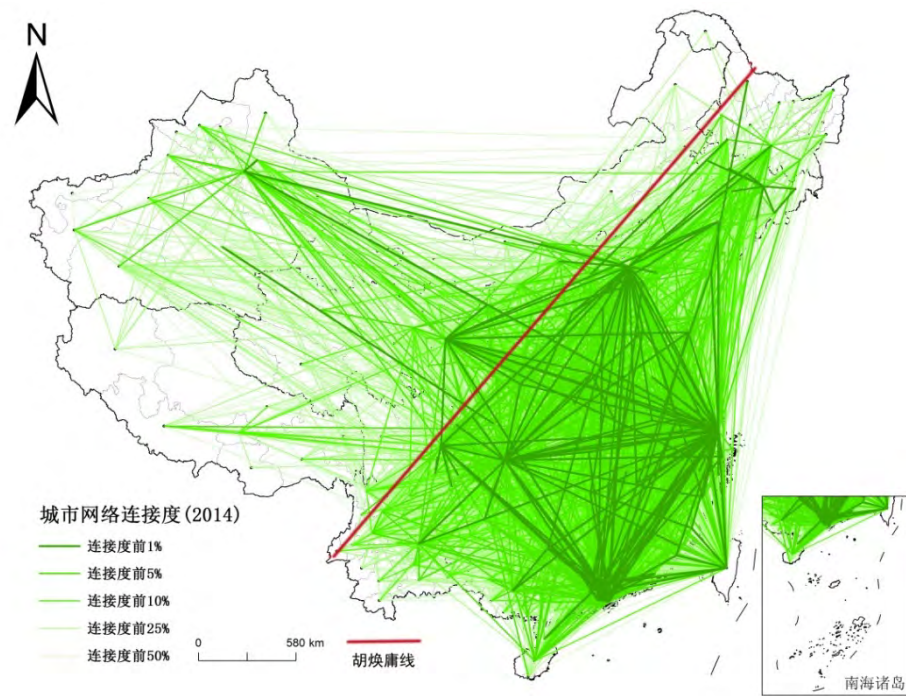
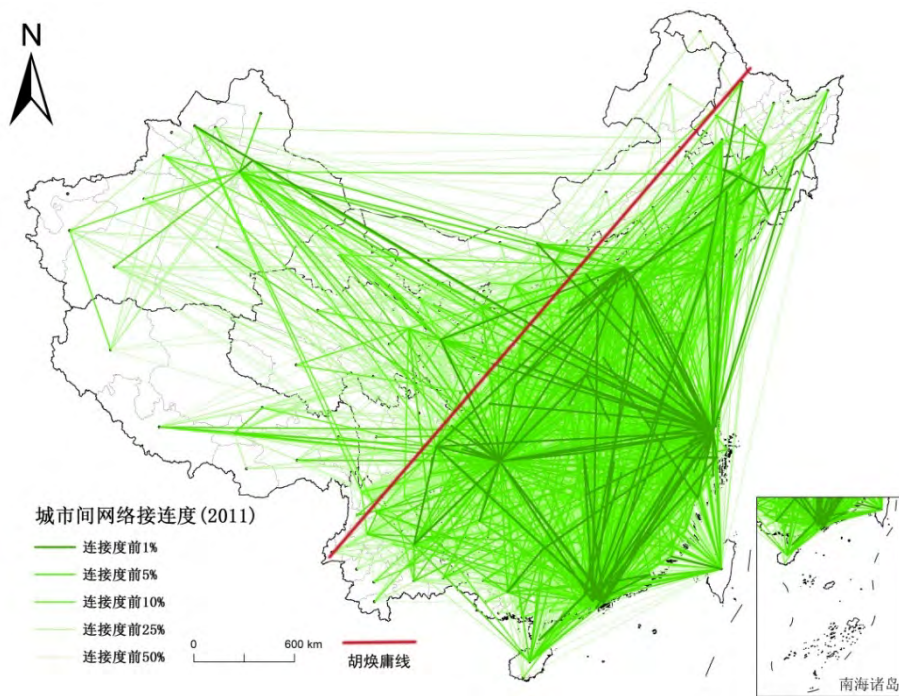
- 城市网络研究起源于对“流的空间”（Space of Flow）的理解。
- 国内外相关学者多利用航班及货运量、港口吞吐量、公路车流量、铁路流量等交通流，以及高级生产性服务业数据，这些多以母子公司、企业总部分支机构等为研究对象。但这些当中，对人文要素的关注较少，而这些数据也不易获取，这也是相关研究的难以开展的主要原因。
- 伴随着互联网信息大数据时代到来，通过计算机编程数据运算等各种手段，获得了一些新类型的数据类型，如微博、公交卡、百度指数、手机信令、POI网点等数据。
- 本研究利用数据采集技术，获取到中国各城市的百度POI网点数据，从人们经常参与其中的生产生活服务类网点入手，根据不同城市的人在各个地区开设的网点信息而产生的社会空间联系，尝试着构建相应的网络模型，分析中国社会城市网点的等级联系等特征。



- **研究思路：**根据POI的名称字段的特征进行模型构建。如果一个城市的POI中含有其他城市的名称字段，则代表这两个城市之间产生了联系，POI个数越多，城市间的联系强度越大，形成的网络等级越高。
- 根据以上思路，结合**地理信息空间处理方法**对城市相关数据进行分析。主要基于各个城市POI网点的位置及名称类型，筛选出有城市字段的POI网点，建立OD模型空间关联的数据库，并通过创建全国的网络数据集，进一步绘制城市之间的网络连接图并作相应分析，具体基于Python+ArcGIS实现。
- 其中，一些城市字段名称有些特殊，对最终的结果分析造成了一些影响，因此在程序编写代码时对其进行相应的List处理，消除影响。
- **主要测度指标：**

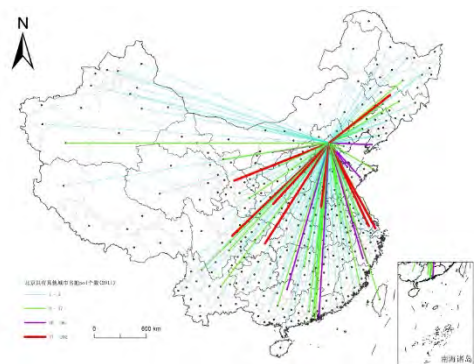
城市间联系流	出度中心性	入度中心性	总中心性
城市的多元性	距离衰减系数	变异系数	

中国城市网络空间格局特征

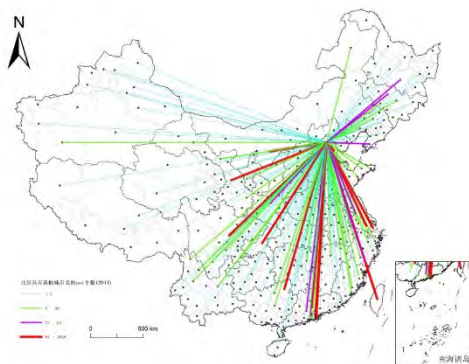


- 总体上，中国城市网络的空间格局差异较大，发展不平衡。

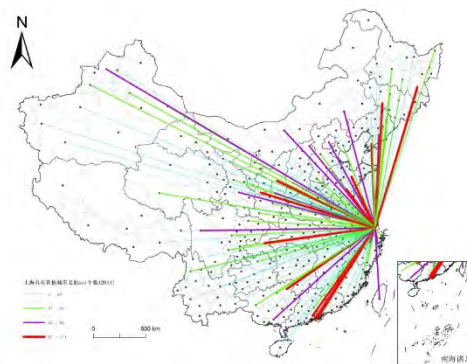
高等级城市为核心的中国城市网络出度格局



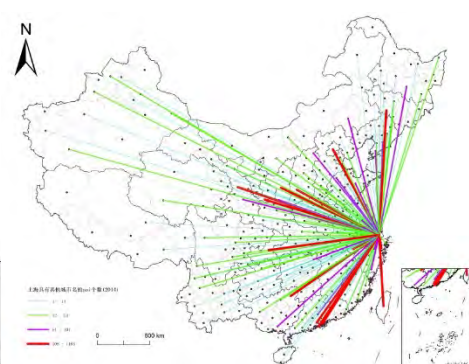
北京2011年



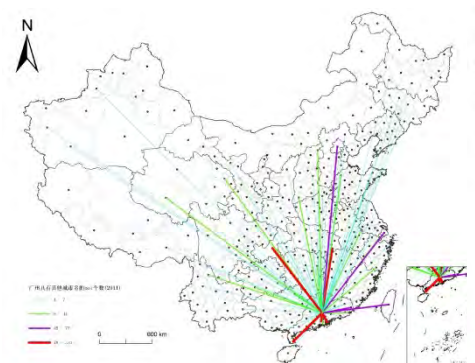
北京2014年



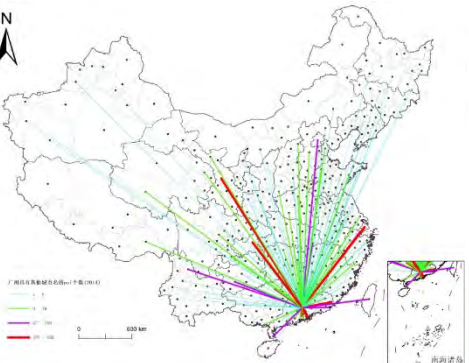
上海2011年



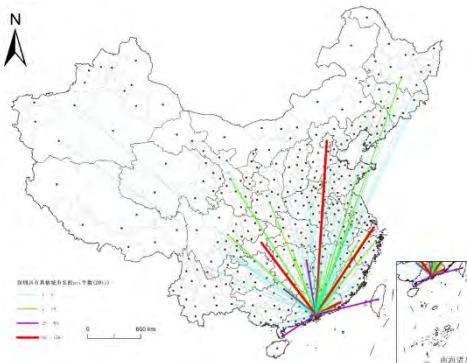
上海2014年



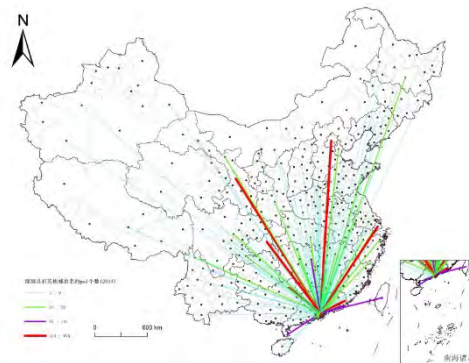
广州2011年



广州2014年



深圳2011年



深圳2014年

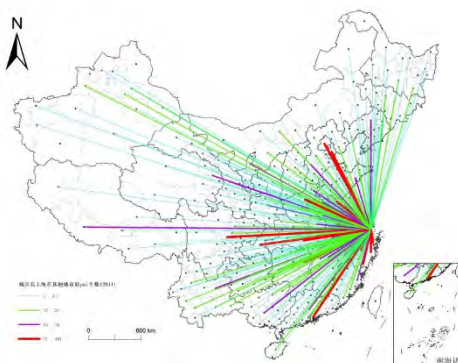
高等级城市为核心的中国城市网络入度格局



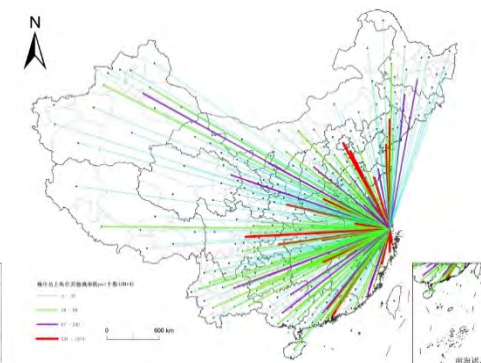
北京2011年



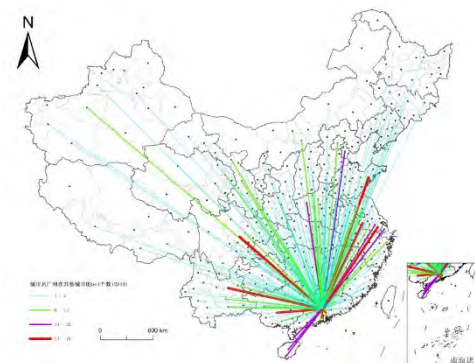
北京2014年



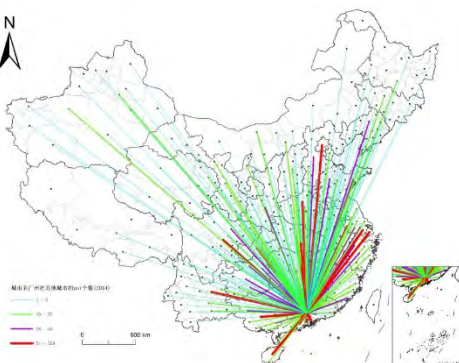
上海2011年



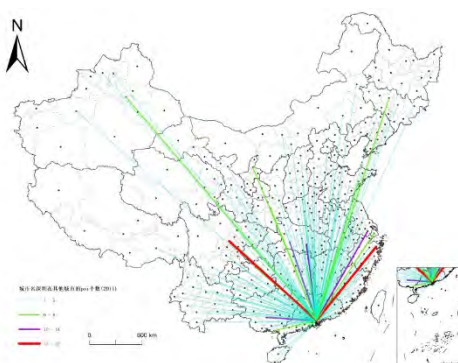
上海2014年



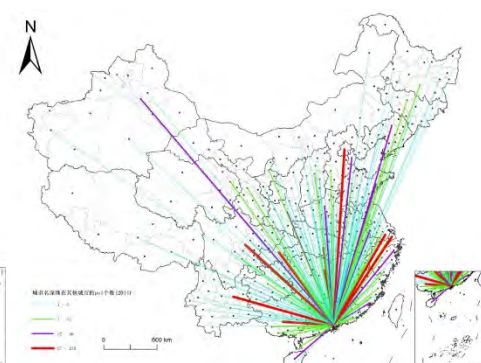
广州2011年



广州2014年

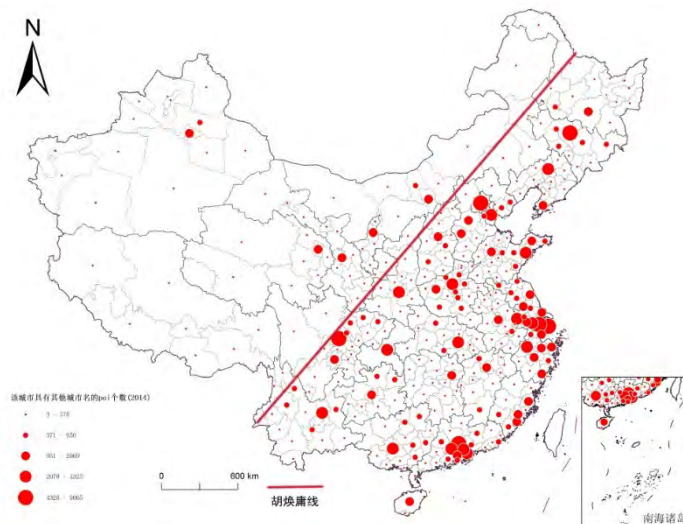
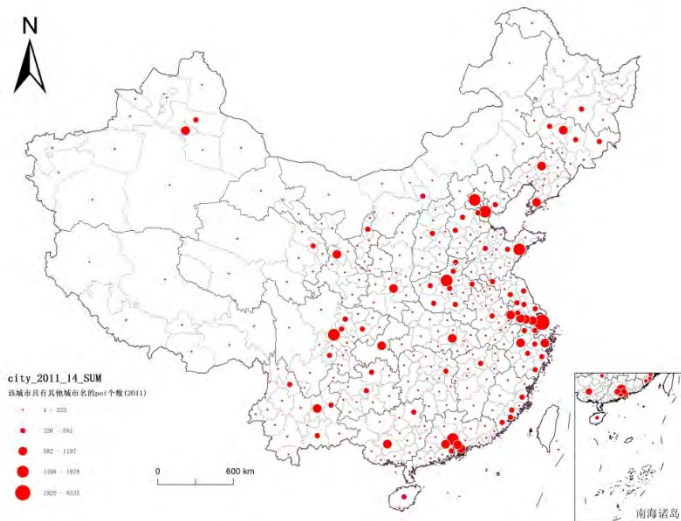


深圳2011年

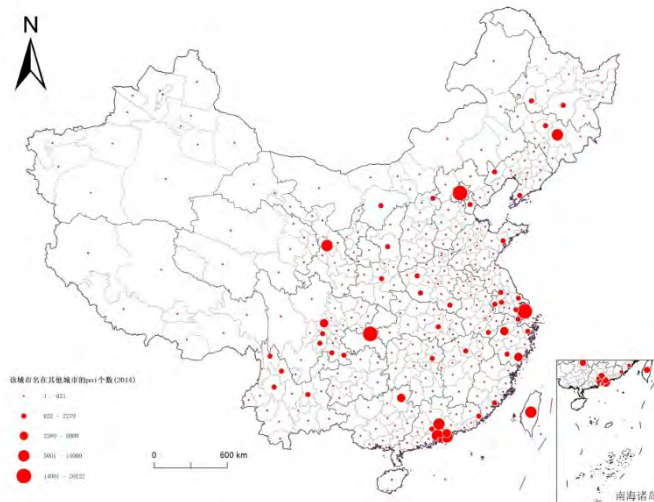
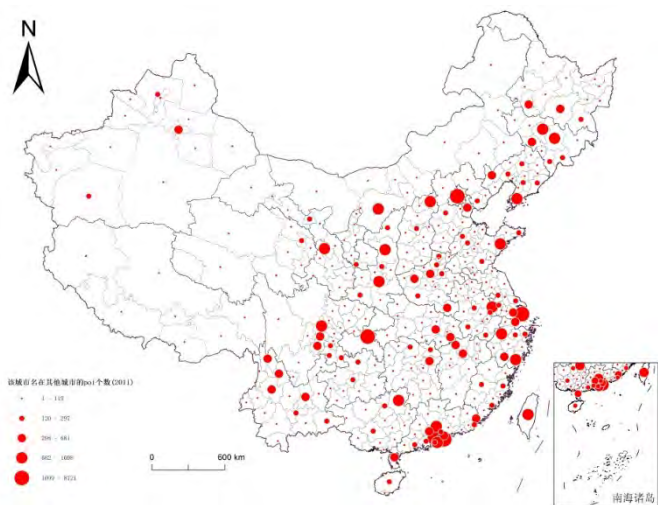


深圳2014年

城市网络层级入度/出度特征—基于POI个数

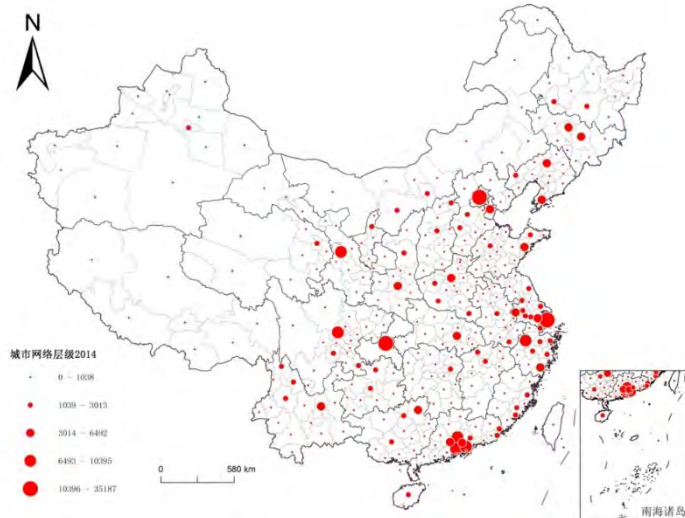
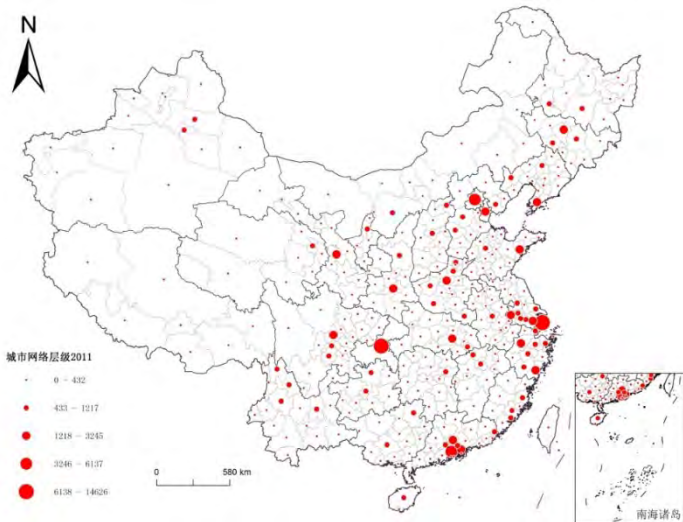


2011年和2014年该城市具有其他城市名的POI个数

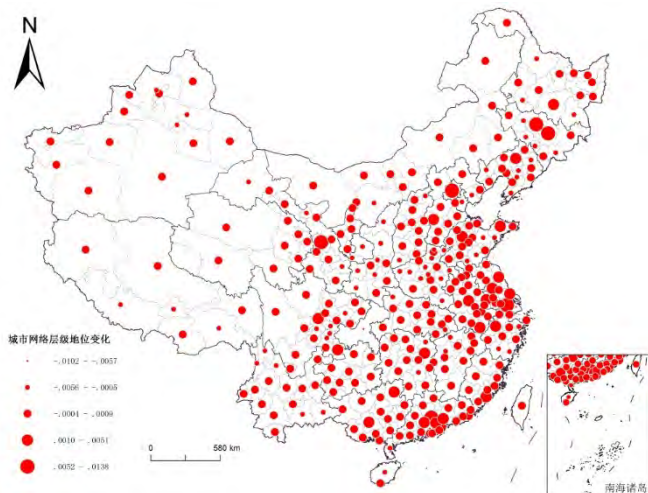
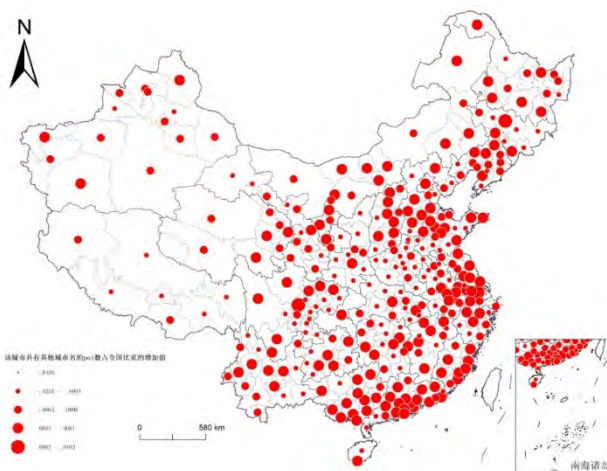
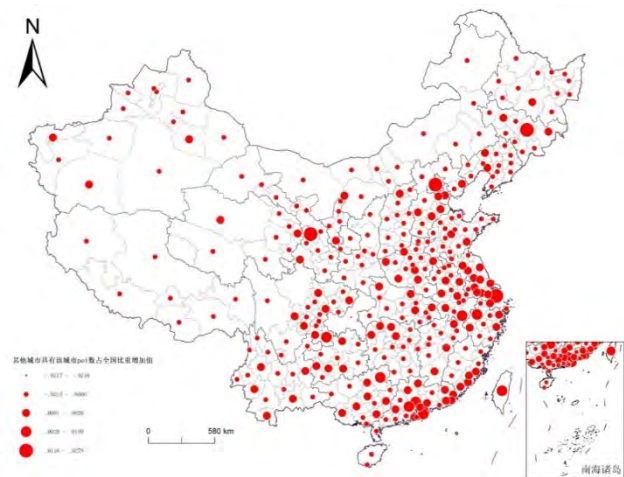


2011年和2014年该城市名在其他城市的POI个数

城市网络层级分析与比较—基于POI个数



2011年和2014年城市网络层级格局

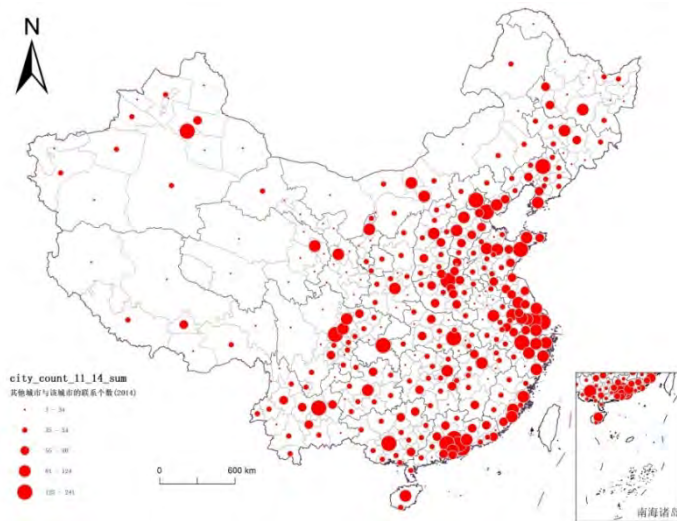
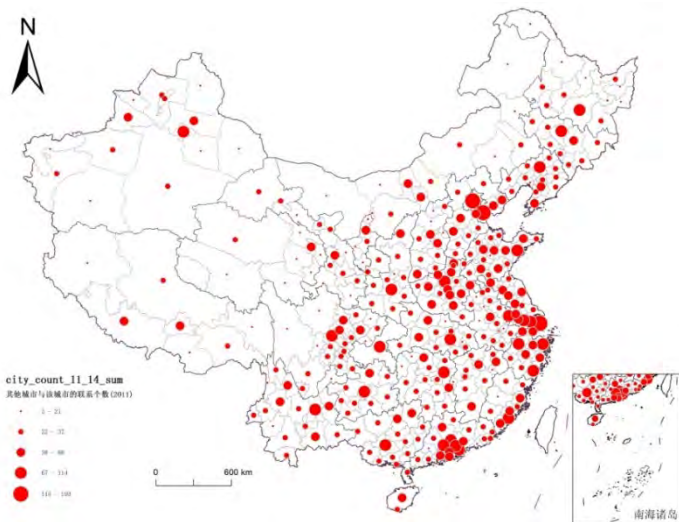


节点城市出度中心性变化

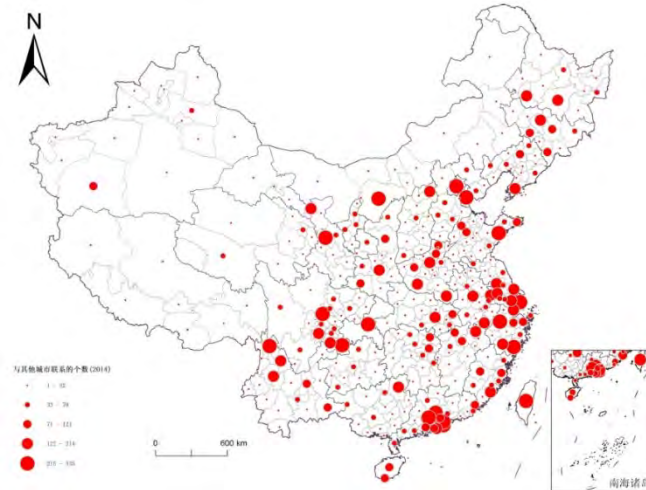
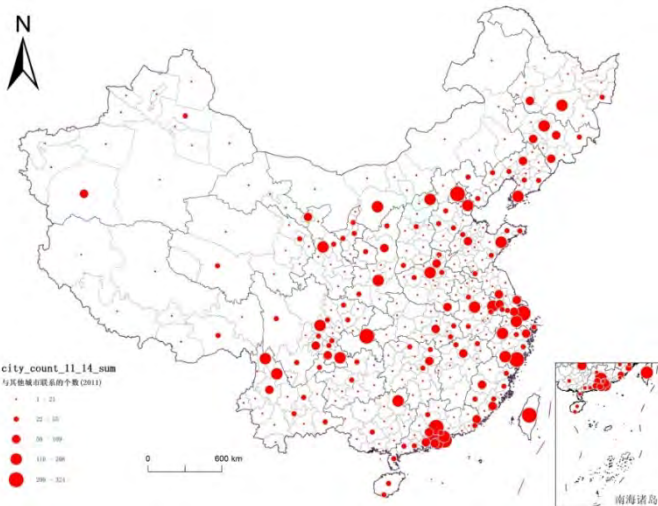
节点城市入度中心性变化

节点城市总中心性变化

城市网络层级入度/出度特征—基于有联系城市个数

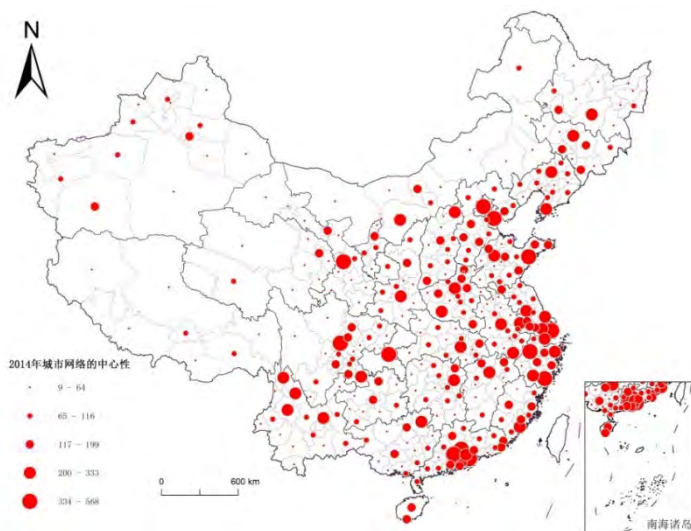
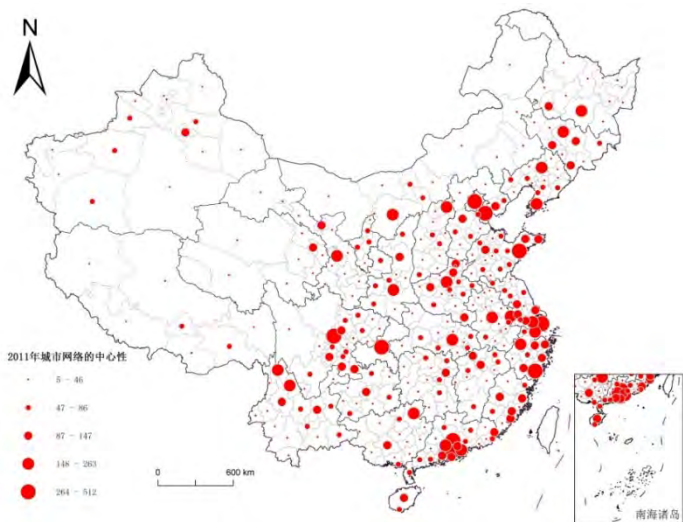


2011年和2014年其他城市与该城市有联系的城市个数



2011年和2014年与其他城市有联系的城市个数

城市网络层级分析与比较—基于有联系城市个数



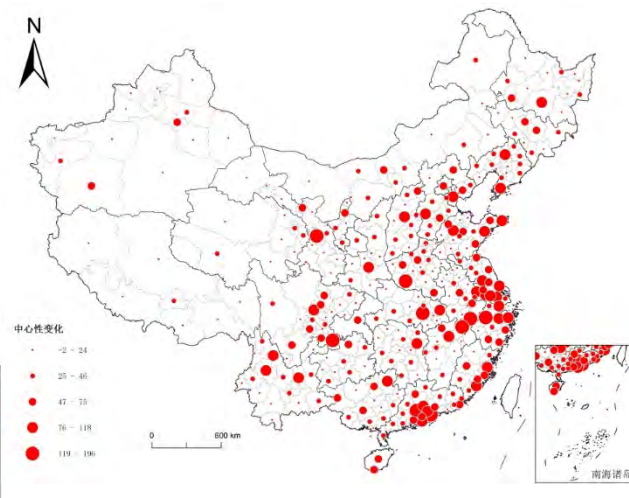
2011年和2014年城市的网络中心格局



节点城市出度中心性变化

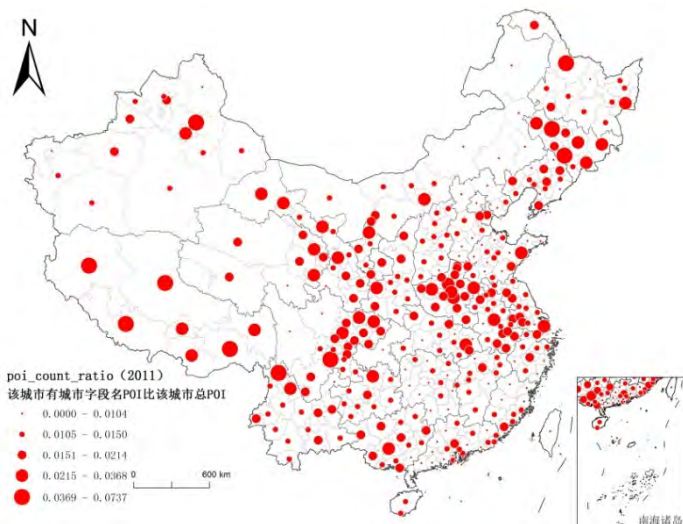


节点城市入度变化

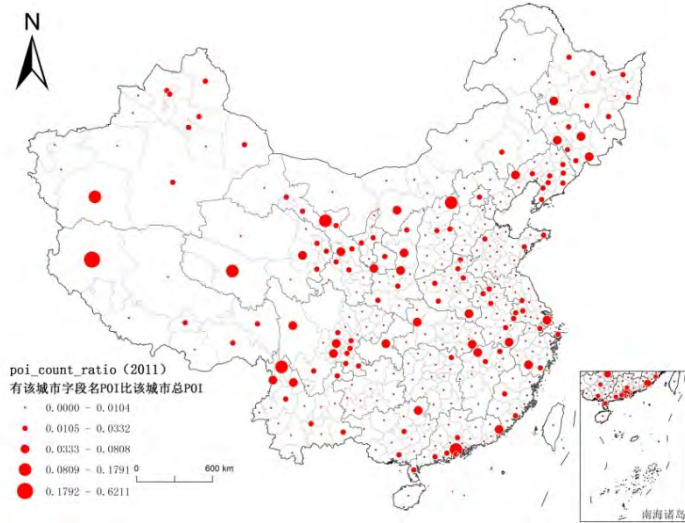


节点城市总中心性变化

城市多元性分析



2011年和2014年该城市有城市字段名POI比该城市总POI



2011年和2014年其他城市有该城市名POI比该城市总POI

关于做研究的小窍门

- 第一讲/第二讲：参考文献的重要性
 - 外国人的姓名写法、认真与否、文献等级
- 第三讲：
 - 论文与报告的区别（是否有科学问题）
 - 问题：Problem vs Question
- 第四讲：两类论文
 - 方法：证明方法优于已有的其他方法（效率/科学性、规划师/公众/同行评价？）
 - 实证：证明发现，与其他人发现的异同，对理论的贡献
- 第五讲：善用Google Scholar
 - <https://scholar.google.com/citations?user=4KAatI4AAAAJ>
 - 文献检索、跟踪某个学者的新发表或新引用、查看自己领域的更新updates、参考文献格式生成
- 第六讲：千里之行，始于足下。勤奋
- 第八讲：是什么、为什么和怎么办？解释模型与预测模型
 - R square的故事
- 第十讲：识别identify、评估profile、解释reasoning与政策policy

课后安排

- 阅读材料：
 - Didi_THU_Public_Report_v1
 - 龙瀛和周垠 2017 规划师_知识产出与消费
- 参考资料
 - 许留记和龙瀛 2016 北京规划建设_中国城市网络分析
- 答疑
 - ylong@tsinghua.edu.cn
 - 建筑学院新501办公室（默认每周下课后-11AM、Open Office Hour）
 - 建议提前两天通过info系统进行预约
 - 助教：徐婉庭，wantingsu1031@qq.com
 - 欢迎前来讨论大作业初步思路

终期成果要求（征求意见稿）

- 最终汇报时间：第16周整个上午
 - 鼓励在线可视化（如基于GeoHey）、网站等其他形式的成果
 - 每位同学一篇短文，关于上课感受（至少800字）
 - 每组一篇学术论文（达到可发表水准）
 - 初步第17周结束上交
-
- 有任何意见，请反馈给课代表。谢谢！



龙瀛, ylong@tsinghua.edu.cn, 新建筑馆501, 13661386623



北京城市实验室
Beijing City Lab

<http://www.beijingcitylab.com>



新浪微博: 龙瀛a1_b2 北京城市实验室BCL

微信公众号: beijingcitylab

清华大学

