

# 基于大数据的城市网络分析

Urban Network Analysis using Big Data

龙瀛，博士

清华大学建筑学院

2016年10月28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（整合进9月23日）
  2. 大数据在城市规划中应用的研究进展（9月23日）
  3. 城市大数据的获取（9月30日）
  4. 城市大数据的分析与统计（10月9日）
  5. 城市大数据的可视化（10月14日）
  6. 大数据时代的空间句法（10月21日）
  7. 基于大数据的城市网络分析（10月28日）
  8. 图片城市主义：城市规划设计与研究的新思路（11月4日）
  9. 学生作业中期汇报与点评（11月11日）
- 
1. 数据增强设计（11月18日）
  2. 战略及总体规划中的大数据应用（11月25日）
  3. 控制性详细规划中的大数据应用（12月2日）
  4. 城市设计中的大数据应用（12月9日）
  5. 参与式规划中的大数据应用（12月16日）
  6. 大模型：新数据环境下的城市研究新方法（12月23日）
  7. 学生作业终期汇报与点评（12月30日）



## 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
  
- 大数据与空间句法关系的调查（online）

# 第二届数据增强设计学术研讨会

Symposium on Data Augmented Design 2016 (DAD2)

## 主旨报告 (按姓氏首字母排列)

题目待定 边兰春, 清华大学建筑学院  
谈谈中国的城市数据问题 毛其智, 清华大学建筑学院  
数据、规划框架与规划师的思考 沈振江, 金泽大学环境设计学院 (日本)  
题目待定 徐磊青, 同济大学建筑与城市规划学院

## 特邀报告 (按姓氏首字母排列)

基于室内定位数据的建筑尺度行为分析 黄蔚欣, 清华大学建筑学院  
基于150亿条数据的通州城镇化规划研究 姜鹏、徐飞, 国家发改委城市和小城镇改革发展中心  
规划变革下一程: 从数据驱动到算法驱动 李栋, 北京清华同衡规划设计研究院  
数据增强设计的研究、教学与实践 龙瀛, 清华大学建筑学院  
大栅栏人本观测: 从外表到内心的街道观测实验 茅明睿, 北京市城市规划设计研究院  
数据化设计教学实践 盛强, 北京交通大学建筑与艺术学院  
大尺度城市设计的数据支撑探讨 (暂定) 唐燕, 清华大学建筑学院  
多源大数据在城市设计中的应用: 本土实践与挑战 杨俊宴、史宜, 东南大学建筑学院  
网络聚集的厚度: 区域空间设计 杨滔, 北京市建筑设计研究院  
建筑密度与形态在城市空间活力上的影响测度 叶宇, 同济大学建筑与城市规划学院  
集成大数据与BIM的数据增强设计实践: 以城市设计为例 张鸿辉, 长沙市规划信息服务中心  
基于开放数据的城市公共中心规划设计方法 郑晓伟, 西安建筑科技大学建筑学院

## 清华大学建筑学院学生报告及点评 (按姓氏首字母排列)

基于多源数据的中国荒野识别与保护规划初探 曹越  
数据自适应的历史街区慢行系统设计 曹哲静  
新数据环境下总体城市设计空间形态生成的案例借鉴新方法 甘欣悦  
中国城市的中心识别与演变及其规划设计响应 李娟  
基于多时相街景图片的街道空间品质测度与变化识别: 以东四历史街区为例 李诗卉  
数据增强设计支持下的城市街道类型化设计 马尧天  
特大城市中心区街道空间品质测度 唐婧娴

主办: 中国城市科学研究会城市大数据专业委员会 (筹)  
清华大学建筑学院、北京城市实验室 (www.beijingscitylab.com)

时间: 2016年12月10日 (周六) 8:30-18:30

地点: 清华大学建筑学院王泽生报告厅

主持: 党安荣/龙瀛, 清华大学建筑学院 (联系邮件ylong@tsinghua.edu.cn)



(报名链接)

# 本讲大纲

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

# 共享的联系数据

OBJECTID *	userid	tripid	time1	lon1	lat1	poiid1	time2	lon2	lat2	poiid2	poiid	poiid_1
3243186	3996664072	3	9/17/14 20:	116.450119	39.9326247	B2094654D1	9/24/14 18:	116.387423	39.9377257	B2094757D5	B2094654D1	B2094757D5
3243187	3996664072	4	9/24/14 18:	116.387423	39.9377257	B2094757D5	10/11/14 21	116.447138	39.9311364	B2094757D1	B2094757D5	B2094757D1
3243188	3996666565	1	9/7/14 17:5	116.445650	39.9165266	B2094757D0	9/13/14 14:	116.447812	39.9350800	B2094757D0	B2094757D0	B2094757D0
3243189	3996666565	2	9/13/14 14:	116.447812	39.9350800	B2094757D0	9/15/14 12:	116.452895	39.9158370	B209465DD2	B2094757D0	B209465DD2
3243190	3996666565	3	9/15/14 12:	116.452895	39.9158370	B209465DD2	9/28/14 22:	116.452072	39.9057028	B2094757D4	B209465DD2	B2094757D4
3243191	3996666565	4	9/28/14 22:	116.452072	39.9057028	B2094757D4	12/28/14 01	116.415046	39.9472578	B2094654D1	B2094757D4	B2094654D1
3243192	3996666565	5	12/28/14 01	116.415046	39.9472578	B2094654D1	12/28/14 07	116.440246	39.9320165	B2094654D1	B2094654D1	B2094654D1
3243193	3996670020	1	12/28/14 08	116.437280	39.9259820	B2094654D0	4/15/15 19:	116.384918	39.9381767	B2094757D0	B2094654D0	B2094757D0
3243194	3996670382	10	9/25/14 14:	112.152369	37.5288635	B2094656D7	9/25/14 20:	116.415865	40.0317297	B2094750D5		B2094750D5
3243195	3996670382	11	9/25/14 20:	116.415865	40.0317297	B2094750D5	9/30/14 20:	111.490673	37.8213689	B209475DD7	B2094750D5	
3243196	3996680185	1	9/3/14 19:5	116.313824	40.0317482	B2094757D0	9/13/14 20:	116.405395	39.9167657	B2094654D0	B2094757D0	B2094654D0
3243197	3996680185	2	9/13/14 20:	116.405395	39.9167657	B2094654D0	9/22/14 20:	116.404965	39.9163657	B2094654D1	B2094654D0	B2094654D1
3243198	3996680185	3	9/22/14 20:	116.404965	39.9163657	B2094654D1	9/23/14 23:	116.447138	39.9311364	B2094757D1	B2094654D1	B2094757D1
3243199	3996680185	4	9/23/14 23:	116.447138	39.9311364	B2094757D1	9/24/14 16:	116.301724	40.0001693	B2094654D2	B2094757D1	B2094654D2
3243200	3996680185	5	9/24/14 16:	116.301724	40.0001693	B2094654D2	10/3/14 23:	116.408482	39.9469329	B2094757D0	B2094654D2	B2094757D0
3243201	3996680185	6	10/3/14 23:	116.408482	39.9469329	B2094757D0	10/6/14 19:	116.454827	39.8973961	B2094654D1	B2094757D0	B2094654D1
3243202	3996680185	7	10/6/14 19:	116.454827	39.8973961	B2094654D1	10/12/14 14	116.434331	39.9199106	B2094757D0	B2094654D1	B2094757D0
3243203	3996680185	8	10/12/14 14	116.434331	39.9199106	B2094757D0	10/12/14 18	116.397953	39.9081834	B209475CD2	B2094757D0	B209475CD2
3243204	3996680185	9	10/12/14 18	116.397953	39.9081834	B209475CD2	10/27/14 06	116.440586	39.9312071	B2094757D0	B209475CD2	B2094757D0
3243205	3996682006	1	1/17/15 23:	115.443033	38.8500868	B2094656D0	1/19/15 17:	116.421038	39.9023238	B2094654D3		B2094654D3
3243206	3996692417	1	9/10/14 10:	116.447138	39.9311364	B2094757D1	10/12/14 15	116.440586	39.9312071	B2094757D0	B2094757D1	B2094757D0
3243207	3996692417	2	10/12/14 15	116.440586	39.9312071	B2094757D0	10/27/14 01	116.434090	39.8781395	B2094654D3	B2094757D0	B2094654D3
3243208	3996692417	3	10/27/14 01	116.434090	39.8781395	B2094654D3	12/28/14 07	116.327436	39.9415903	B2094654D0	B2094654D3	B2094654D0
3243209	3996848482	2	1/31/15 15:	108.032318	24.6900413	B2094653D7	7/14/15 10:	116.378935	39.9364764	B2094757D0		B2094757D0
3243210	3996861396	1	7/4/14 01:0	118.290627	25.3272233	B2094750D1	7/10/15 13:	116.454626	39.8967757	B2094654D0		B2094654D0
3243211	3996848482	3	7/14/15 10:	116.378935	39.9364764	B2094757D0	7/21/15 13:	104.047504	30.6473090	B2094655DA	B2094757D0	
3243212	3996872965	25	8/16/15 05:	116.372740	39.9585221	B2094654D2	8/24/15 20:	116.445758	39.9320562	B2094654D1	B2094654D2	B2094654D1
3243213	3996883568	1	3/16/14 18:	116.337721	40.0078360	B2094654D0	5/15/14 09:	116.336504	40.0021589	B2094757D0	B2094654D0	B2094757D0
3243214	3996883568	2	5/15/14 09:	116.336504	40.0021589	B2094757D0	5/20/14 10:	116.337110	39.9491795	B2094654D1	B2094757D0	B2094654D1
3243215	3996891168	1	4/19/14 16:	116.393707	39.9393984	B2094757D0	6/2/14 17:2	116.346732	39.9397895	B2094757D0	B2094757D0	B2094757D0
3243216	3996894533	1	1/26/14 00:	116.435289	39.9053399	B2094654D0	1/26/14 00:	116.454562	39.8944107	B2094654D1	B2094654D0	B2094654D1
3243217	3996894533	2	1/26/14 00:	116.454562	39.8944107	B2094654D1	3/29/14 23:	116.444427	39.8723134	B2094654D1	B2094654D1	B2094654D1
3243218	3996872965	24	8/3/15 18:1	121.704715	39.0230817	B2094657D3	8/16/15 05:	116.372740	39.9585221	B2094654D2		B2094654D2
3243219	3996893542	3	4/18/15 18:	121.438426	31.2233519	B2094757DB	5/29/15 03:	116.460538	39.8811672	B2094757D0		B2094757D0
3243220	3996893542	5	8/8/15 17:2	121.438955	31.2233682	B2094757D0	8/26/15 09:	116.437919	39.8678057	B2094654D1		B2094654D1
3243221	3996895585	3	3/14/15 13:	113.110942	31.0159958	B2094652D7	3/27/15 13:	116.334842	39.9496188	B2094757D0		B2094757D0
3243222	3996883568	3	5/20/14 10:	116.337110	39.9491795	B2094654D1	5/2/15 18:5	125.420781	43.8248378	B2094757D0	B2094654D1	

# 共享的联系数据

- Trajectory\_Beijing.gdb中的TrajectoryTable\_final，共约324万记录
  - 其中的poiid等对应所共享的Weibo\_POI\_Checkins1127jp空间点图层中的poiid字段
- 来自2015年微博的所有签到记录，根据每个人的不同时间的签到记录推导对应的轨迹（同李栋的黄金周研究）
  - 并没有考虑两次签到之间的时间关系（bias）
  - 提供的数据涵盖了至少一个轨迹点在研究范围（Ext图层）内的轨迹（最后两个字段与之前的poiid1和poiid2重复，空表示不在研究范围内）
- 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poiid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

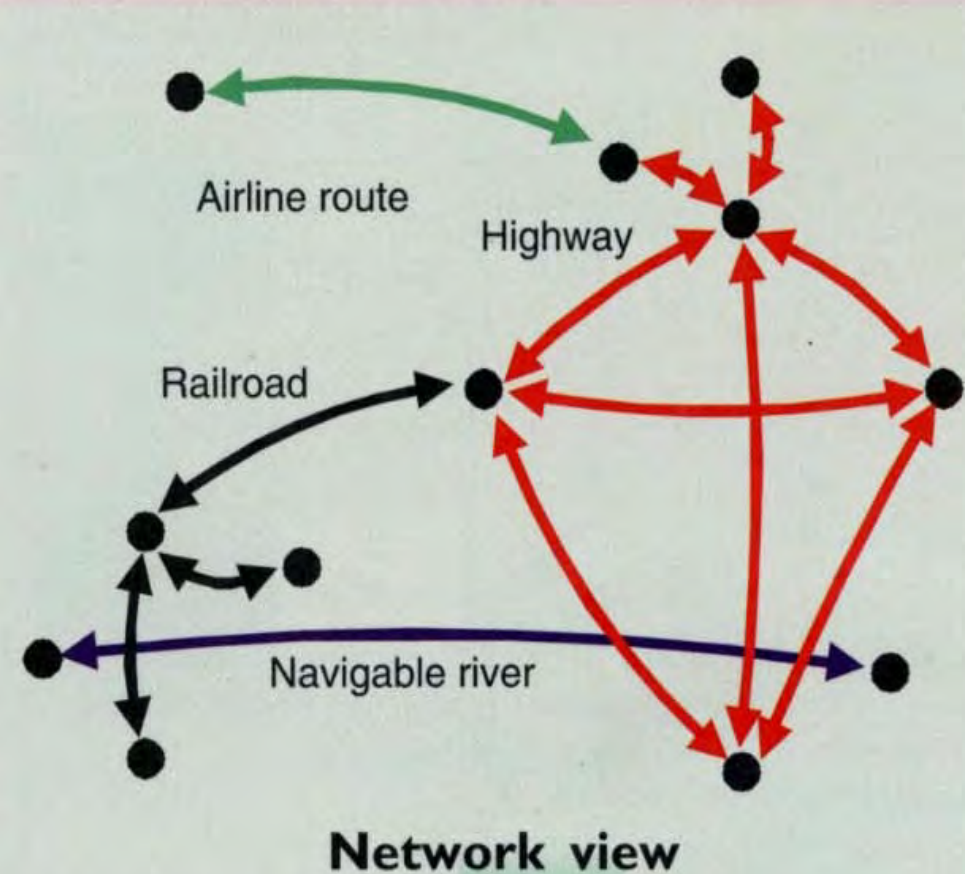
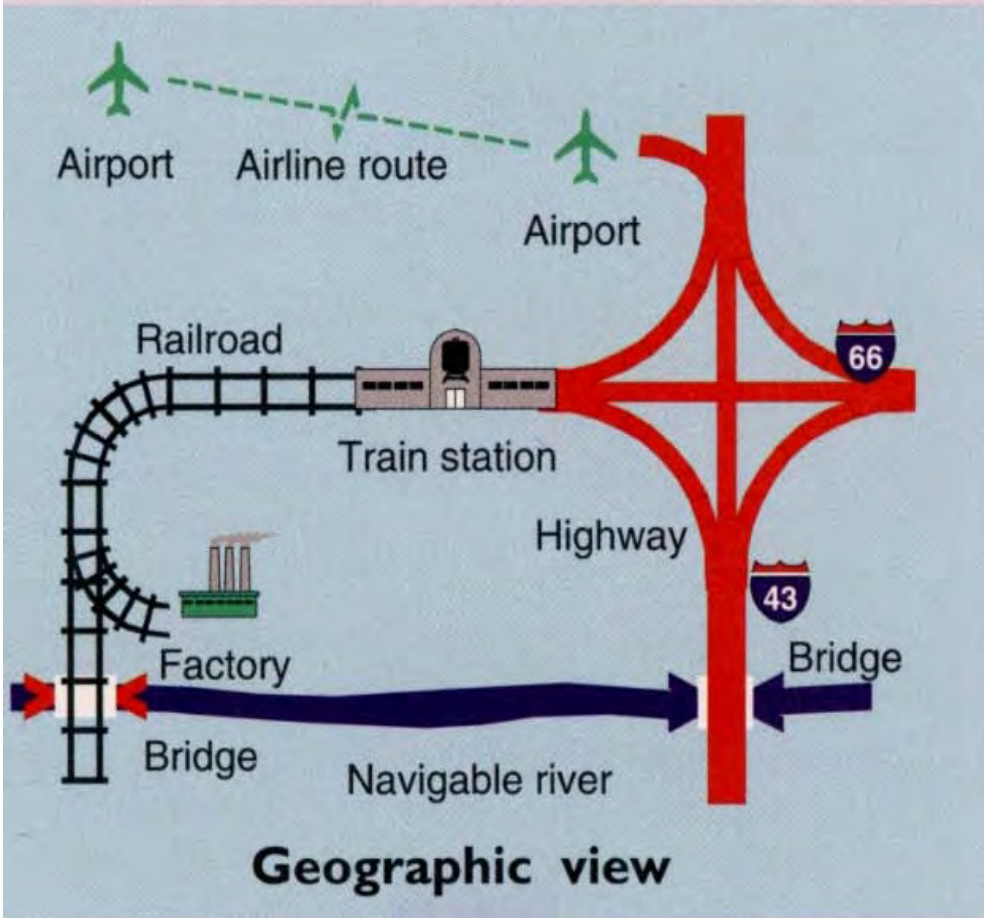


# 城市大数据可以主要归纳为两类

- Density based
- Flow based

# 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.

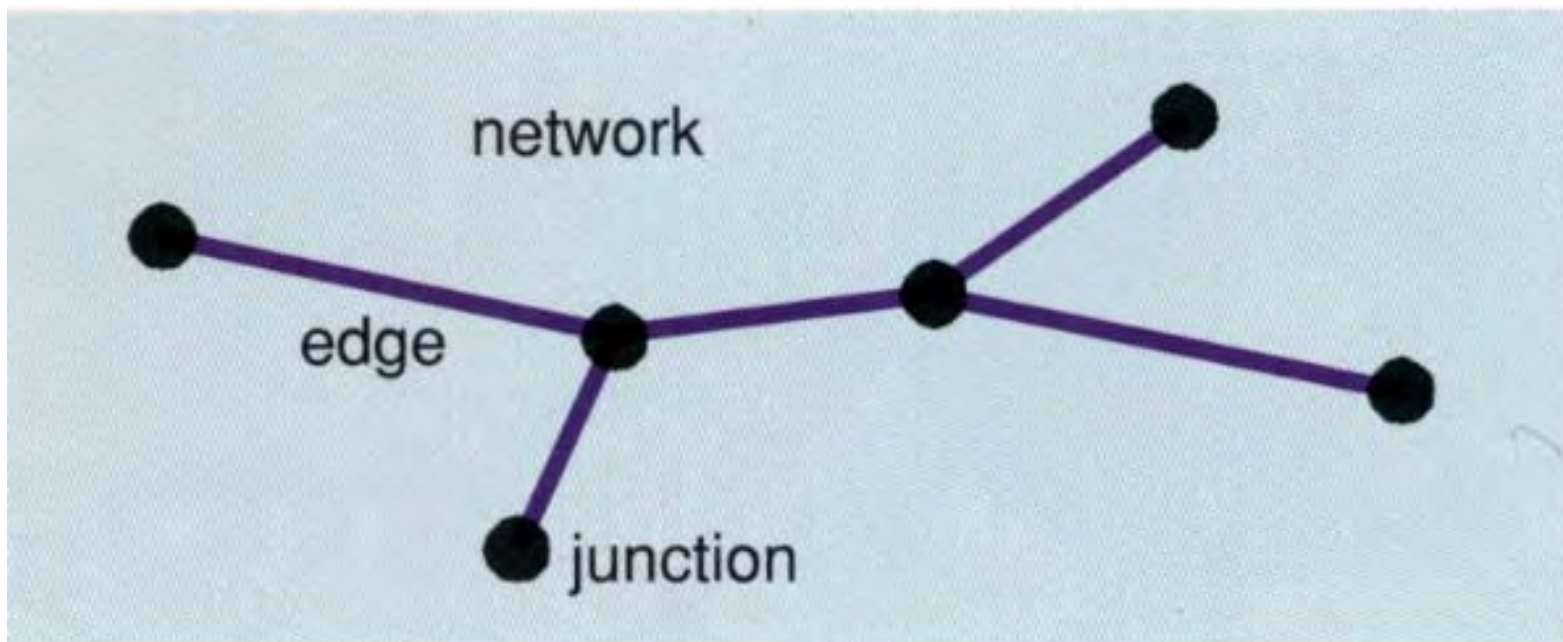
**Geometric network**

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

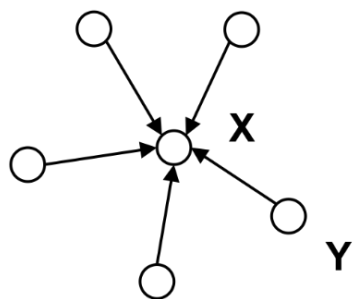
## • 度分布(degree distribution)

- 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。

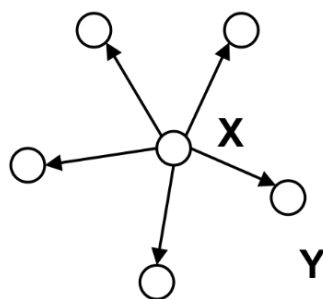
# 基本的统计特征

- 介数 (betweenness)

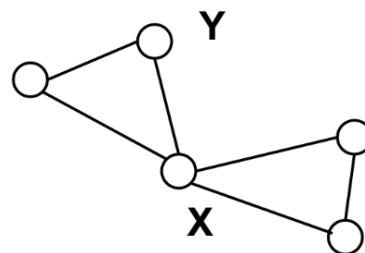
- 介数反映了节点或边的作用和影响力，如果一对节点间共有 $B$ 条不同的路径，其中有 $b$ 条经过节点 $i$ ，那么节点 $i$ 对这对节点的介数的贡献为 $b/B$ 。把节点 $i$ 对所有节点对的贡献累加起来，再除以节点对总数，就可得到节点 $i$ 的介数。
- 类似的，边的介数定义为所有节点对的最短路径中经过该边的数量比例。
- 介数越大 $\rightarrow$ 经过该节点/边的最短路径越多 $\rightarrow$ 经过该节点/边的信息量越大



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  $i$  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  $C^\lambda$ )**: 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

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

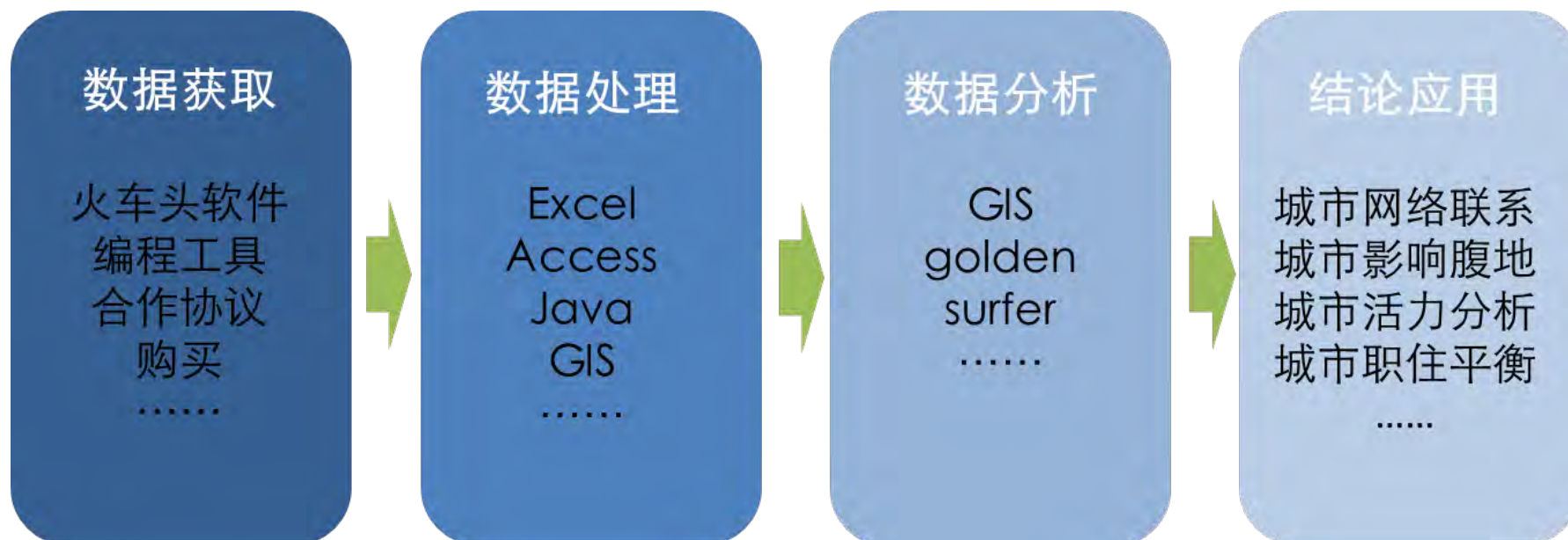
- 这就直接引导出了众所周知的特征向量计算问题  $\lambda S = AS$ , 并且最大特征值对应的特征向量就是这个向量中心性 ( $C^\lambda = S$ ).

# 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
  - 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)

**Globalization and World Cities Research Network**

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://dx.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](http://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](http://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**

---

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 Witlox](#) (Ghent)

[Email GaWC](#)

4,590 visits  
Oct 1st - Oct 31st

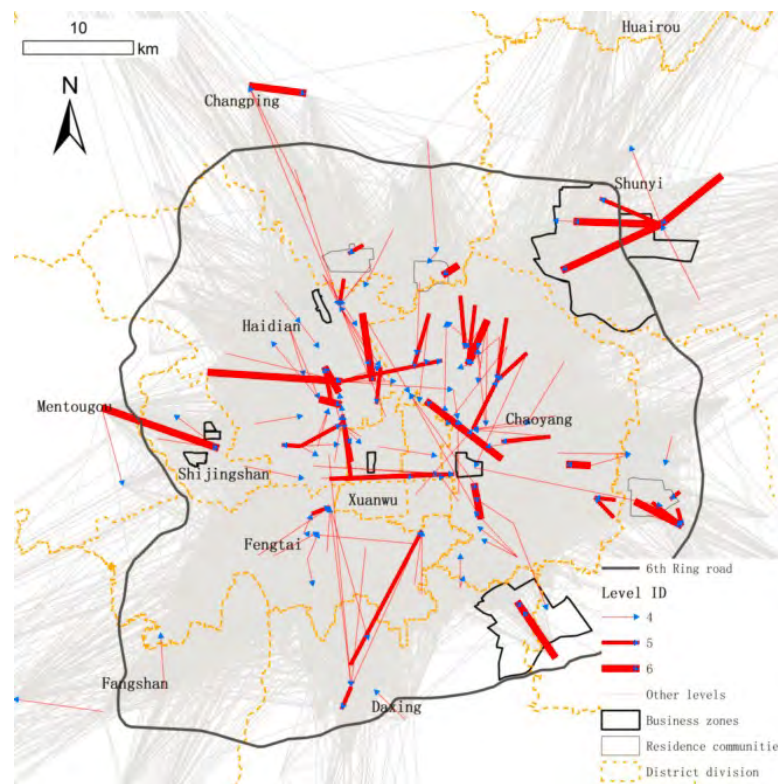
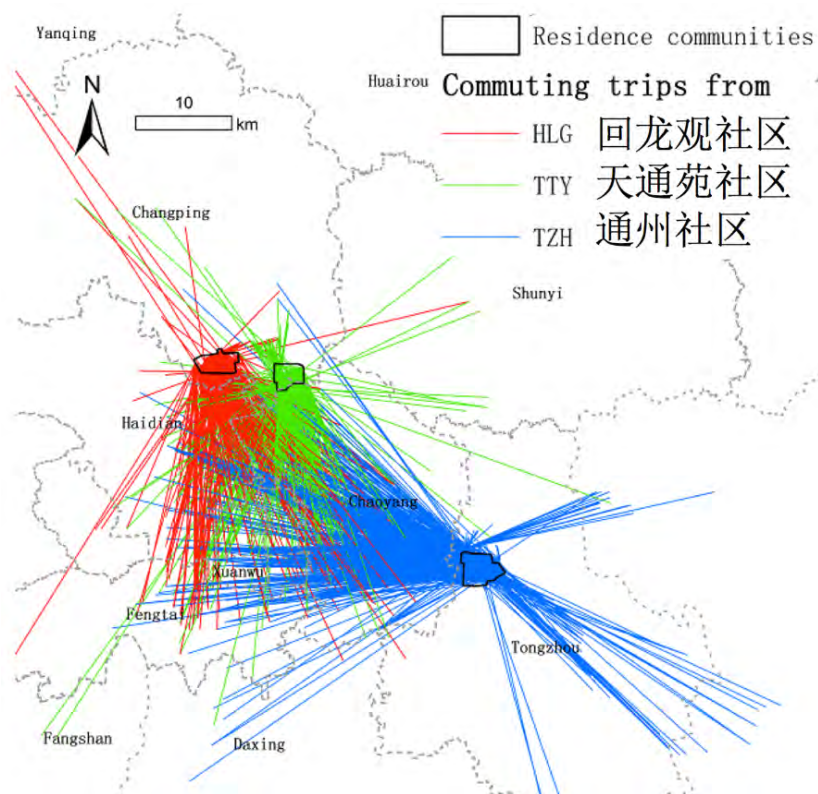


[Click to see details](#)

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

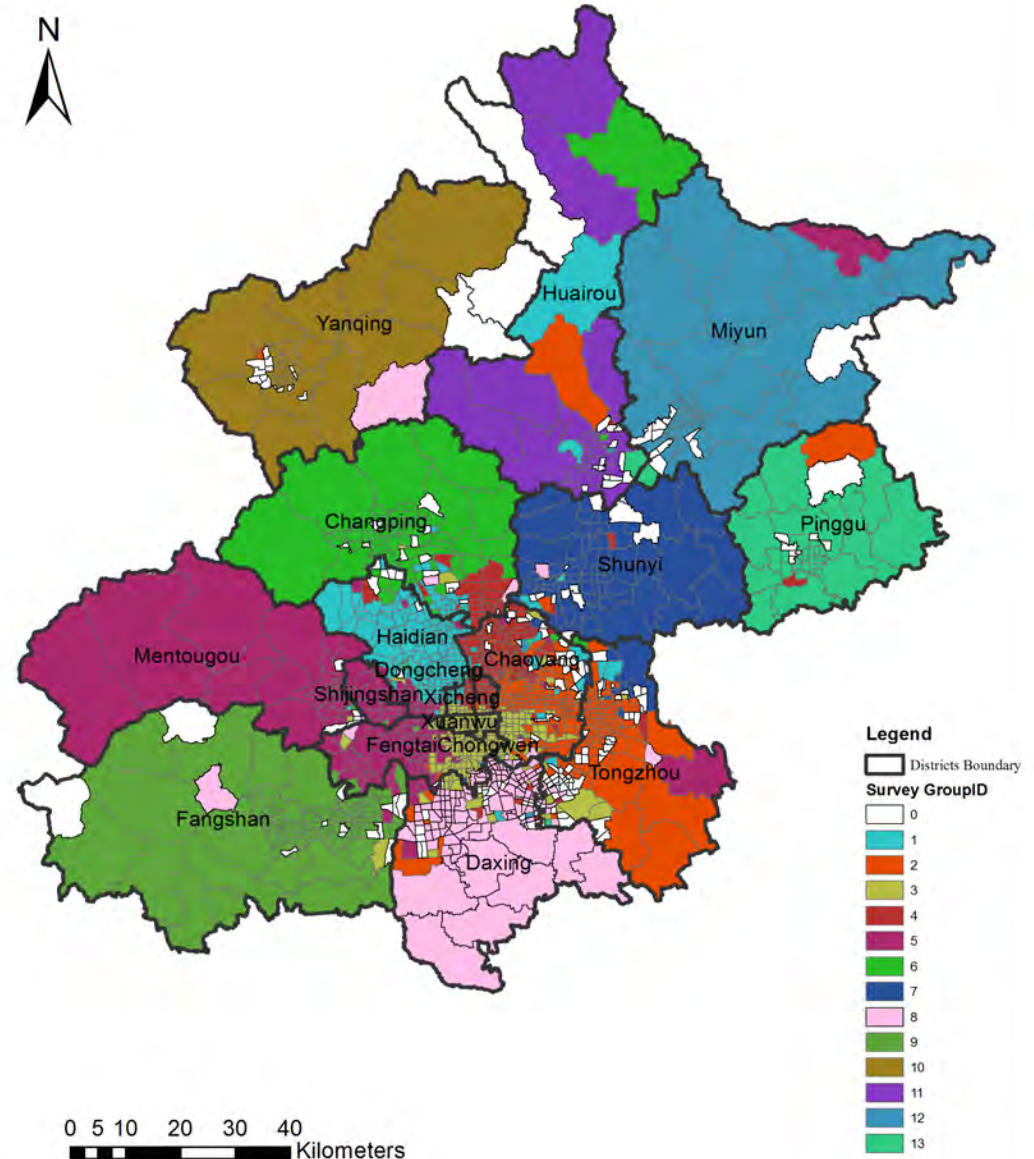
# 城市内部网络

- 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



# 城市内部网络

- 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

# 数据处理

## • 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

# 数据处理

## • 武汉城市圈企业矩阵表

企业分支

企业总部

	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 (0)
- 目的地点 (0)
- 点障碍 (0)
- 线障碍 (0)
- 面障碍 (0)

加载位置

加载自: 城市点

排序字段: FID

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 千米

使用网络位置字段(N)

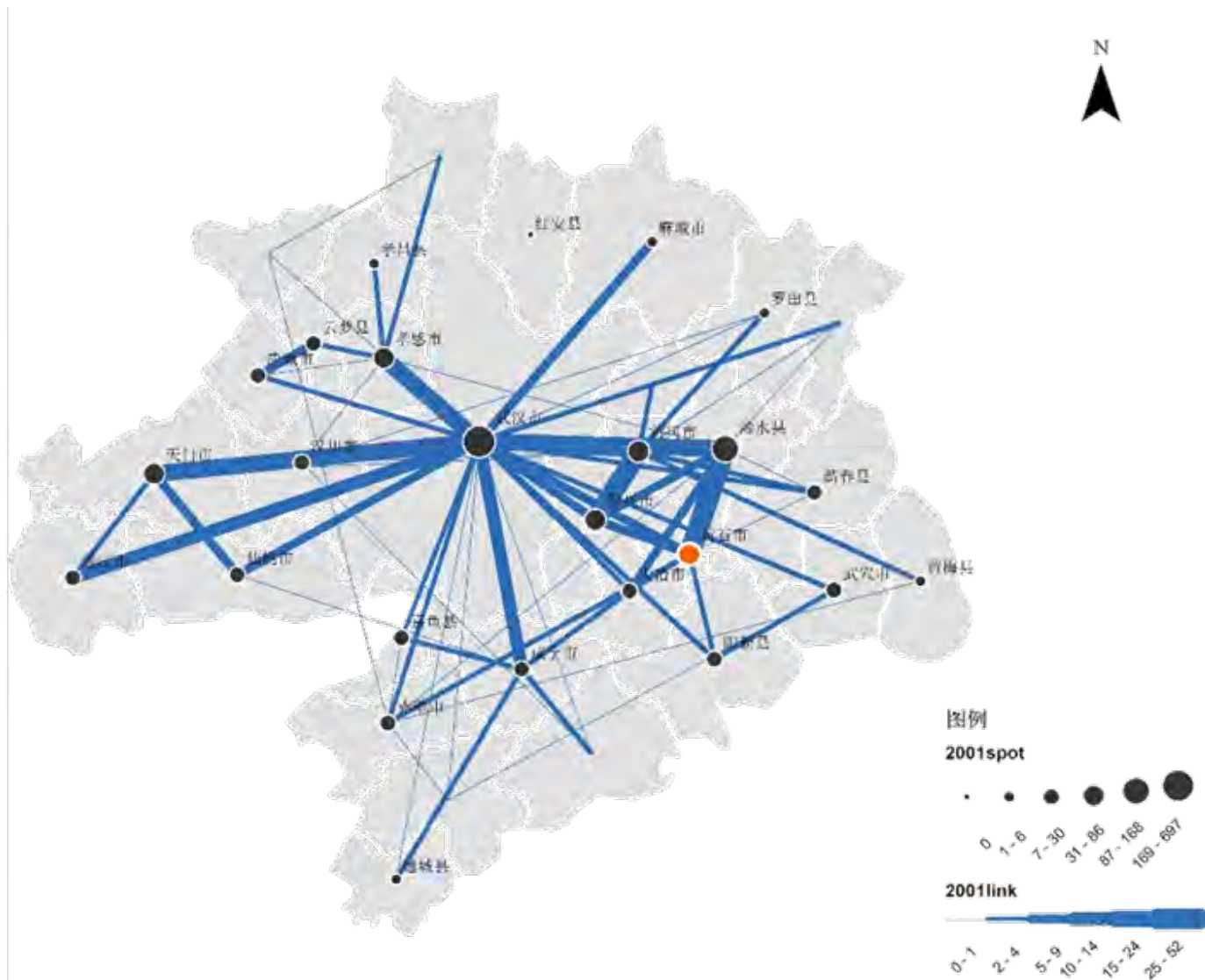
属性	字段
SourceID	
SourceOID	
PosAlong	
SideOfEdge	

高级(A)... 关于加载位置 确定 取消

# 数据分析

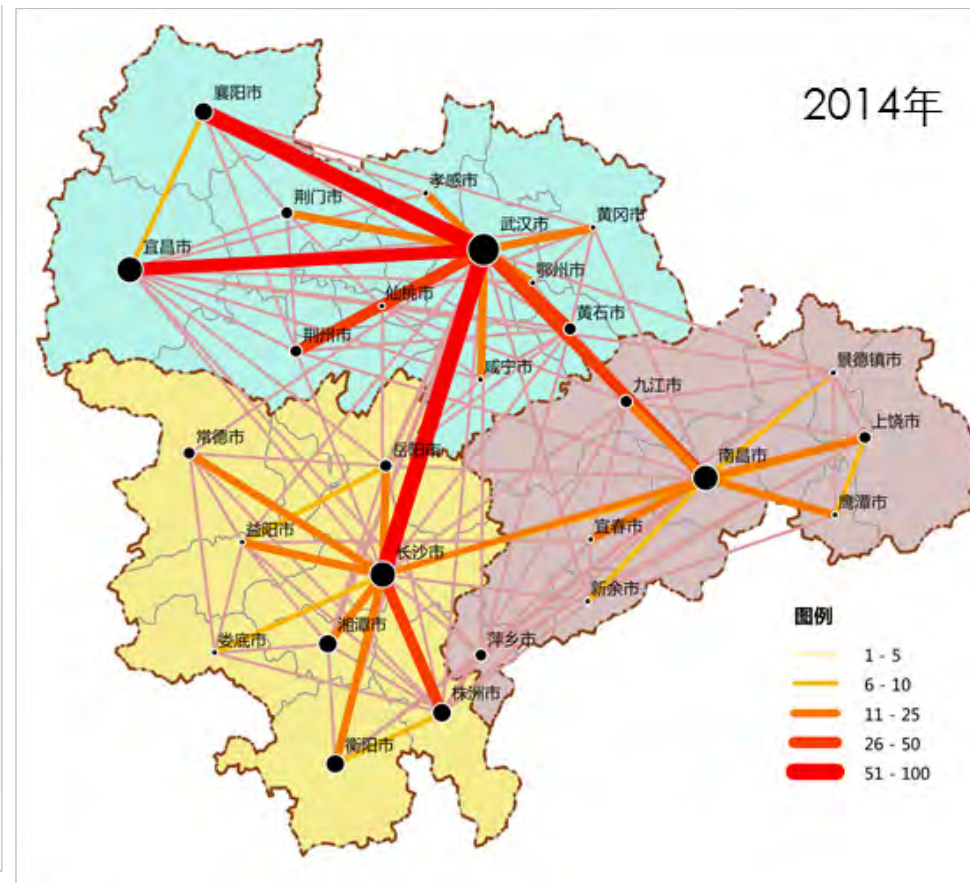
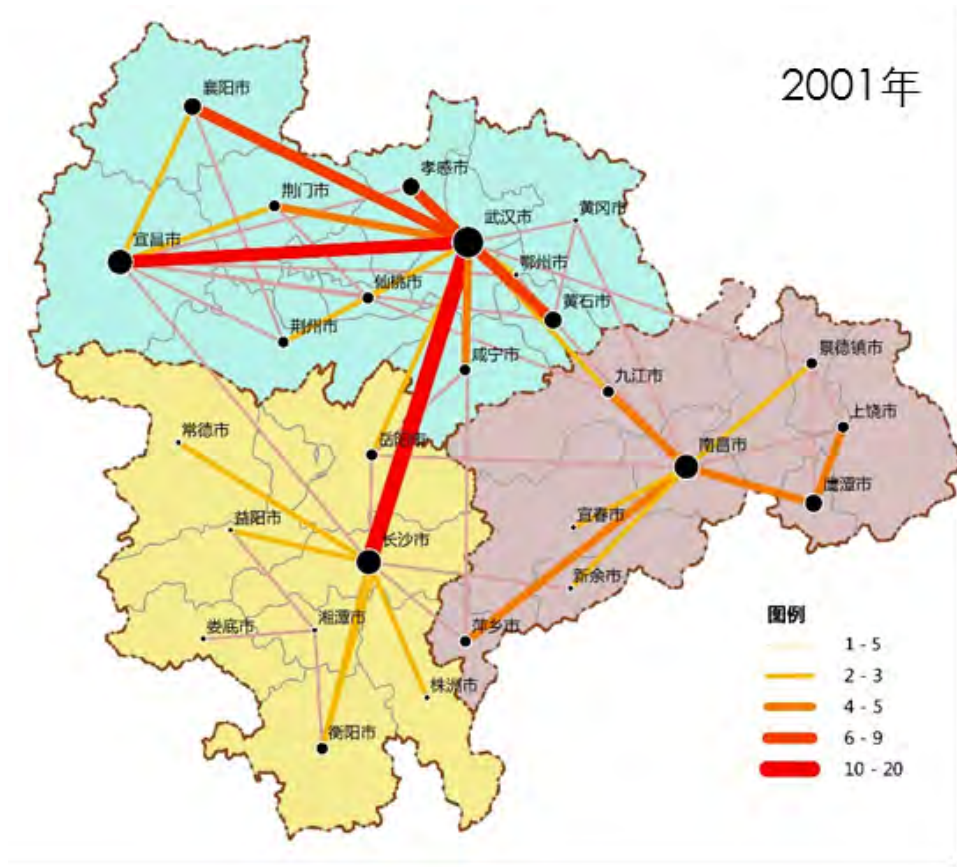
	位置 21 安陆市	位置 22 赤壁市	位置 23 崇阳县	位置 24 大悟县	位置 28 大冶市	位置 31 鄂州市	位置 25 汉川市	位置 2 红安县	位置 3 黄梅县	位置 30 黄石市	位置 18 黄冈市	位置 4 嘉鱼县	位置 5 罗田县	位置 6 麻城市	位置	联系度
位置 21	140														位置 21 - 位置 21	140
位置 22		227	1												位置 21 - 位置 22	0
位置 23			12	25											位置 21 - 位置 23	0
位置 24	1				21										位置 21 - 位置 24	0
位置 28		1				302				32					位置 21 - 位置 28	0
位置 31						1	269			1	2				位置 21 - 位置 31	0
位置 25							141								位置 21 - 位置 25	0
位置 2								21							位置 21 - 位置 2	0
位置 3									36			1			位置 21 - 位置 3	0
位置 30										395					位置 21 - 位置 30	0
位置 18											151		6		位置 21 - 位置 18	0
位置 4			3									74			位置 21 - 位置 4	0
位置 5													40		位置 21 - 位置 5	0
位置 6				1										124	位置 21 - 位置 6	0
位置 7											1				位置 21 - 位置 7	0
位置 8															位置 21 - 位置 8	0
位置 9							2								位置 21 - 位置 9	0
位置 1															位置 21 - 位置 1	0
位置 10															位置 21 - 位置 10	0
位置 17											2				位置 21 - 位置 17	0
位置 27	3	3	3	2	3	3	10	2		1	3	5			位置 21 - 位置 27	1
位置 19			1												位置 21 - 位置 19	0
位置 11							1			1	2				位置 21 - 位置 11	0
位置 12															位置 21 - 位置 12	0
位置 13					2							2			位置 21 - 位置 13	0
位置 26															位置 21 - 位置 26	0
位置 20	1							1							位置 21 - 位置 20	0
位置 29															位置 21 - 位置 29	0
位置 15							1					1			位置 21 - 位置 15	0
位置 14															位置 21 - 位置 14	0
位置 16															位置 21 - 位置 16	0

# 数据分析



武汉城市圈2001年企业分支机构图

# 结论应用



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络
  - “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

# 数据获取



机票预订 > 机票查询 > 航班查询

## 航班时刻查询

航线	出发城市	<input type="text"/>	换
航班号	到达城市	<input type="text"/>	
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 巴彦卓尔航班 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

# 数据处理

- 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 (0)
- 目的地 (0)
- 线 (0)
- 点障碍 (0)
  - 禁止型 (0)
  - 成本增加型 (0)
- 线障碍 (0)
  - 禁止型 (0)
  - 成本按比例增加型 (0)
- 面障碍 (0)
  - 禁止型 (0)
  - 成本按比例增加型 (0)

加载位置

加载自: 城市点

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		
TargetIDestinationCount		
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

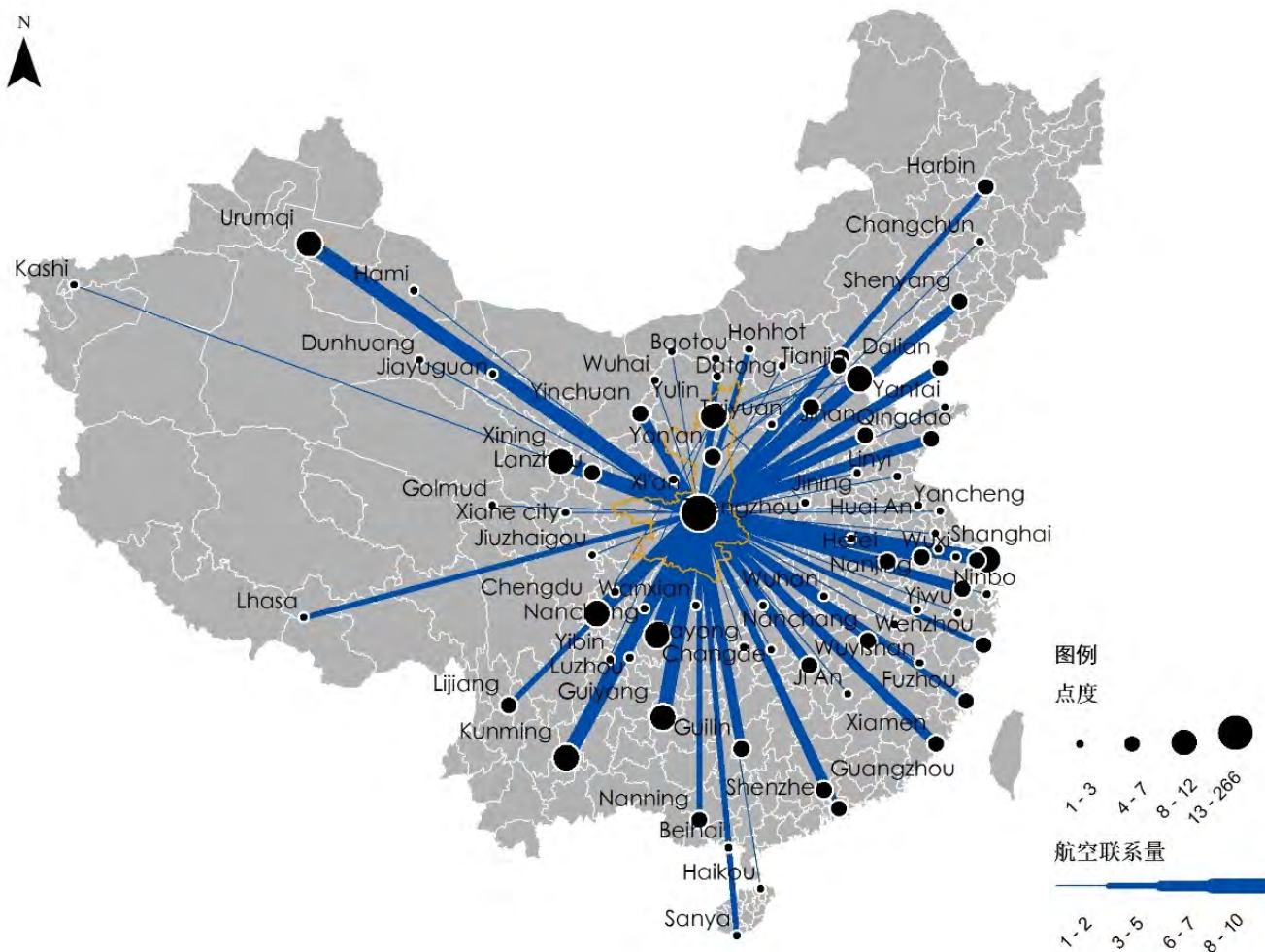
使用网络位置字段(N)

属性	字段
SourceID	
SourceOID	
PosAlong	
sideOfEdge	

高级(A)... 关于加载位置 确定 取消



# 结论应用



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

# 数据获取

The screenshot shows a web API testing tool interface. At the top, there is a navigation bar with links: 开放平台, 首页, 网站接入, 应用开发, 文档, 支持, 管理中心. Below the navigation bar, there is a section for API tools. On the left, there is a form for configuring a request. The form includes a dropdown menu for the API endpoint, a green button labeled "获取Access Token", and a text input field containing the API key "2LUHVE0cc11d8f33RgPJOD". Below this, there is a dropdown menu for the API method, currently set to "GET". There are also radio buttons for "GET" and "POST". At the bottom of the form, there are two input fields for latitude and longitude, with values "39.98437" and "116.30987" respectively. On the right side of the interface, there is a section for the request details. It shows the request method as "get", the request URL as "https://api.weibo.com/2/place/nearby/pois.json", and the request parameters as "lat=39.98437&long=116.30987&access\_token=2.009ac21CdLUHVE0cc11d8f33RgPJOD". Below this, there is a section for the returned content, which is a JSON object: 

```
{  "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

# 数据处理



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

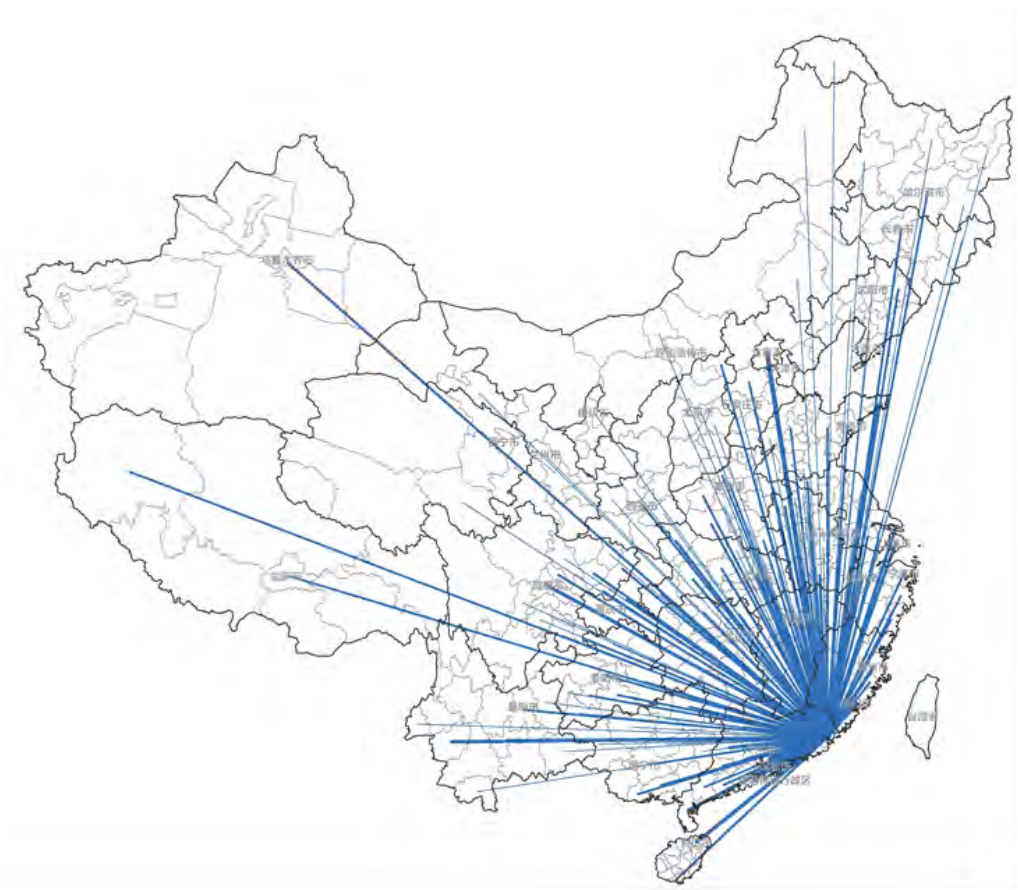
# 数据分析

The screenshot shows the ArcMap interface with the Network Analyst extension. The 'Add Location' dialog box is open, and the 'Location Analysis Properties' table is visible. The table has the following columns: 属性 (Property), 字段 (Field), and 默认值 (Default Value).

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

The 'Location Selection' section shows the 'Use Geometry' option selected, with a search tolerance of 5000 meters. The 'Advanced' button is visible at the bottom of the dialog.

# 应用结论



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

# 背景

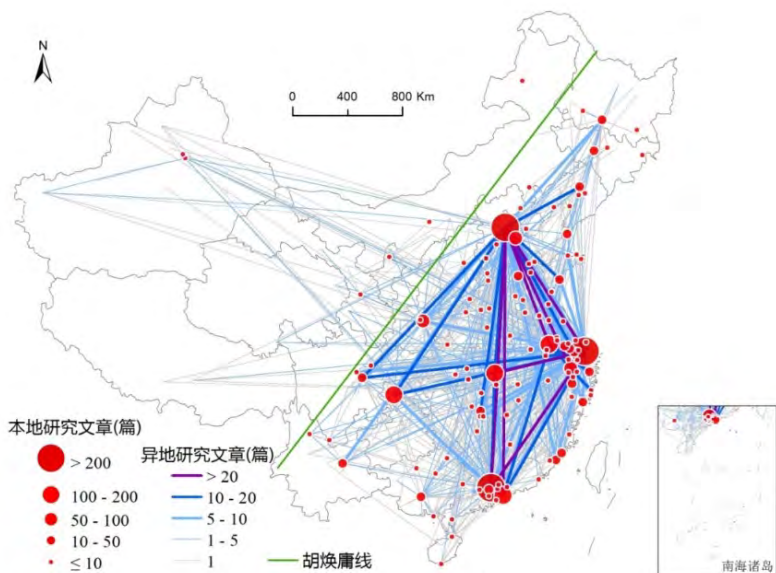
- **知识的产出和消费**是后工业社会的中心问题
  - **知识的消费地区**：大部分城市规划研究以实际项目为依托，有指定的研究区域。将被研究的地区视为城市规划知识的消费地区；
  - **知识的产出者**：与研究项目相关的高校科研工作者、规划设计院/所、规划管理部门、咨询公司的研究人员。
- 国内，城市规划领域已有文献计量研究不足
  - **没能充分利用**互联网丰富的信息资源；
  - **不足以清晰的刻画**出城市规划研究的区域规律；
  - 与其他领域的计量分析**较为相似**。
- 国外，Matthiessen等人利用计量文献分析的方法，对全球知识网络展开了一系列的研究
  - 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。



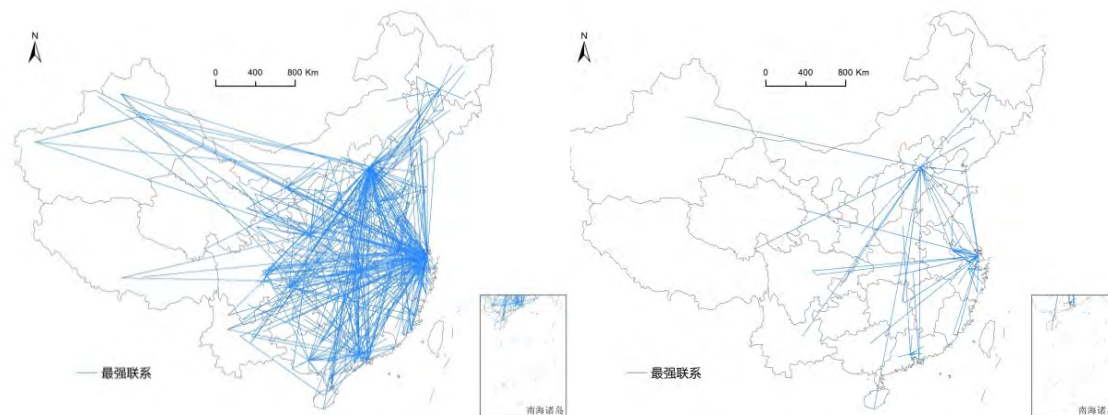
# 数据



# 城市网络规律

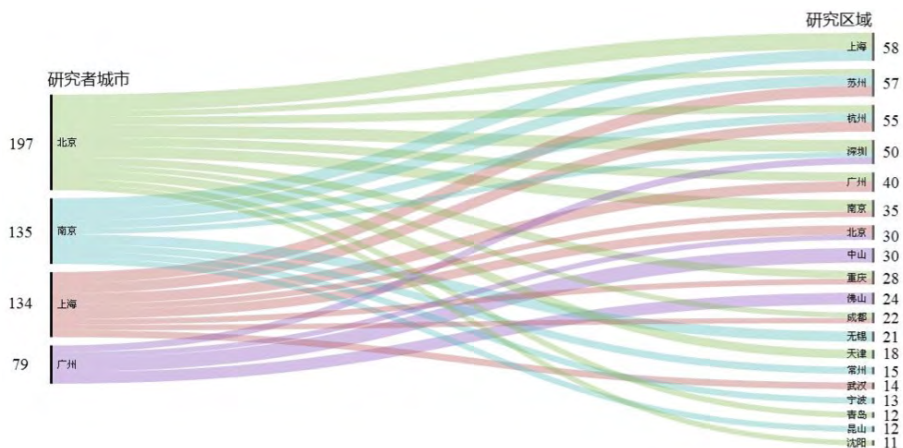


知识产出—消费网络格局（无向连接）

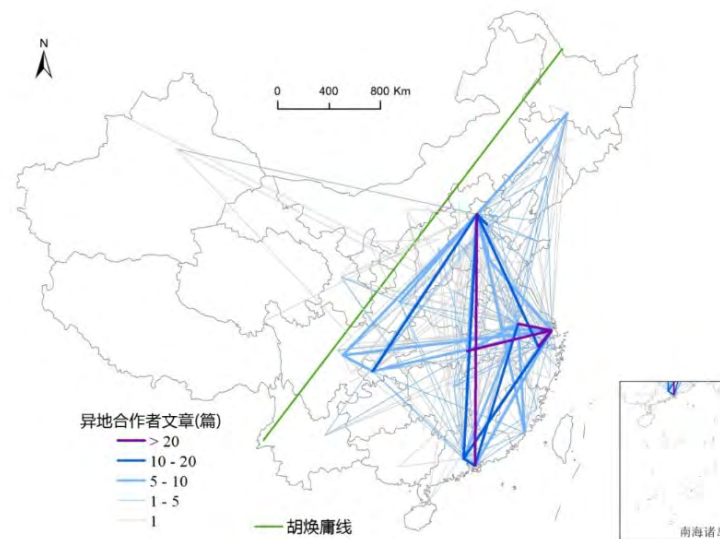


a. 所有最强联系网络 b. 最强联系大于3的网络

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

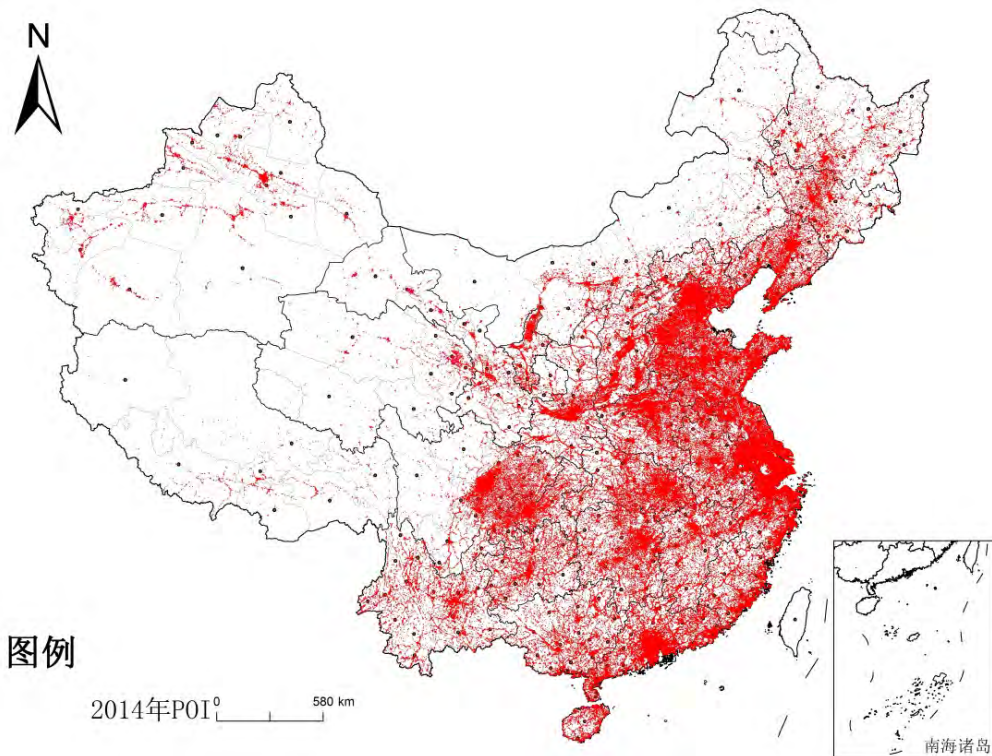
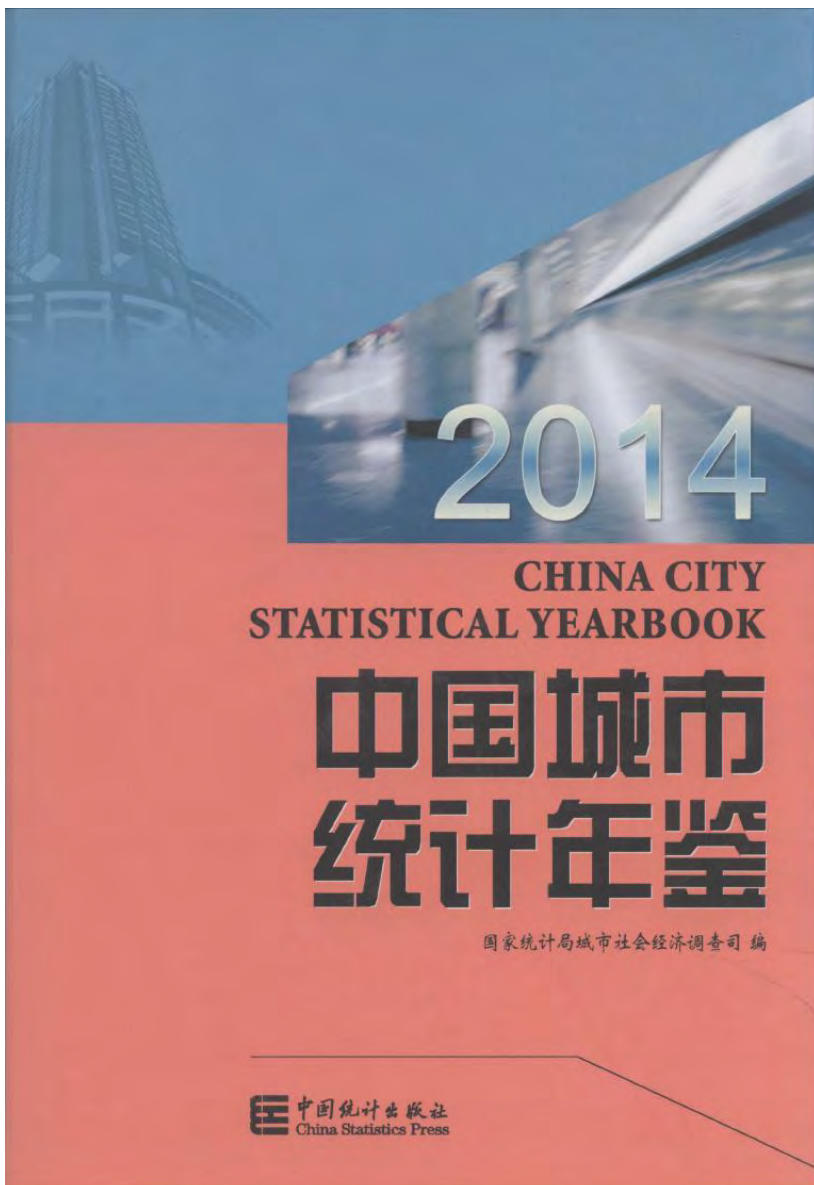
# 六、基于兴趣点位置和名称的 中国城市网络分析

课外阅读材料（辅以共享的论文）

# 背景

- 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。

# 数据



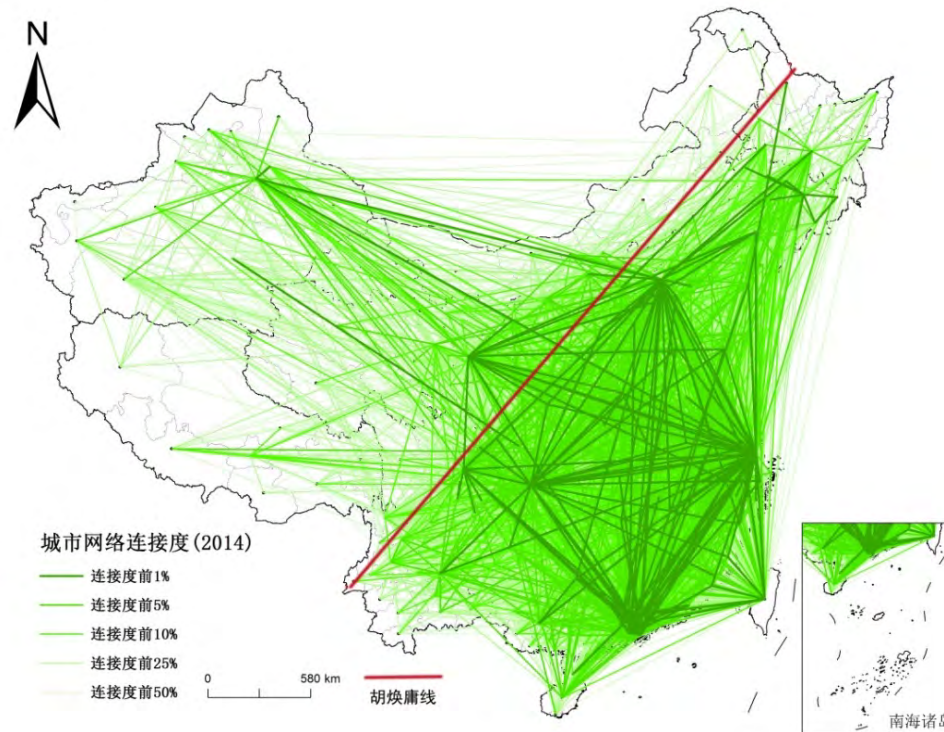
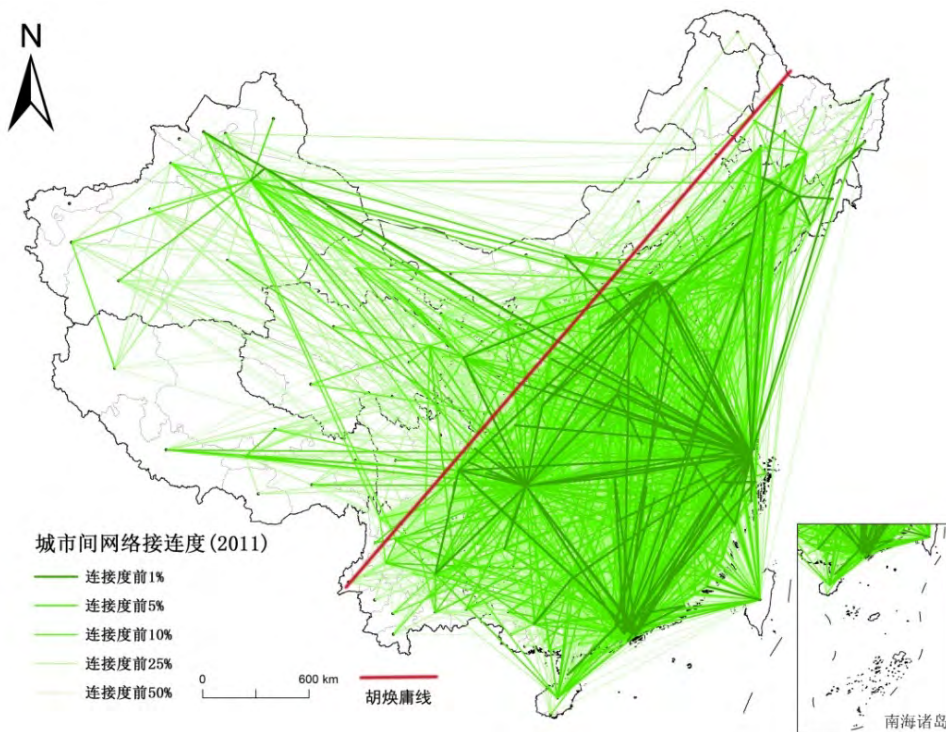
# 研究方法

- **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：**

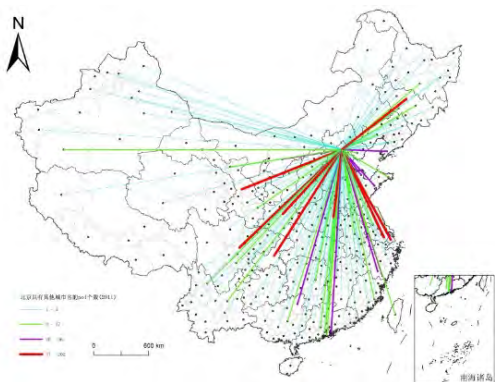
城市间联系流	<b>出度中心性</b>	<b>入度中心性</b>	总中心性
城市的多元性	<b>距离衰减系数</b>	<b>变异系数</b>	

# 中国城市网络空间格局特征

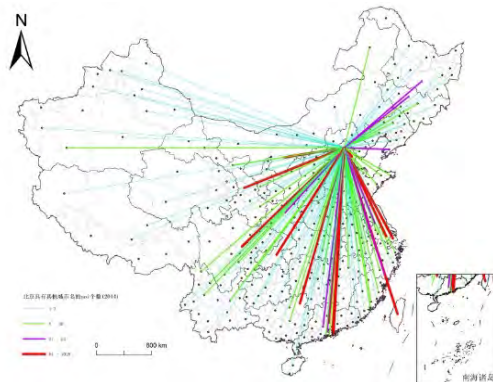


- 总体上，中国城市网络的空间格局差异较大，发展不平衡。

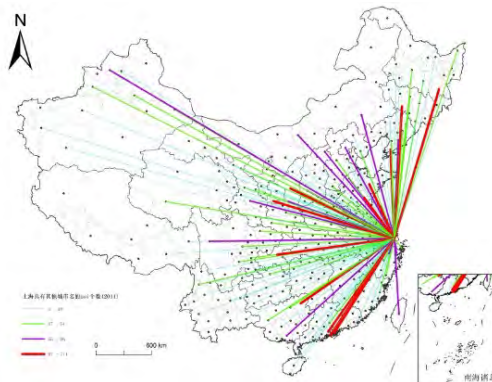
# 高等级城市为核心的中国城市网络出度格局



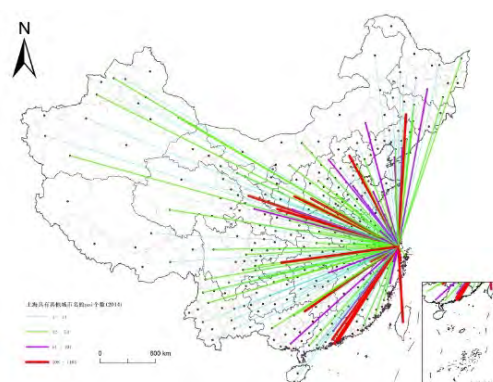
北京2011年



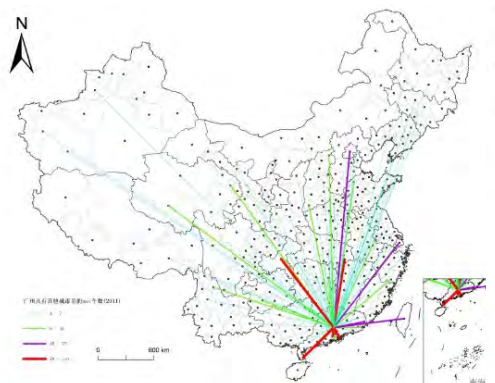
北京2014年



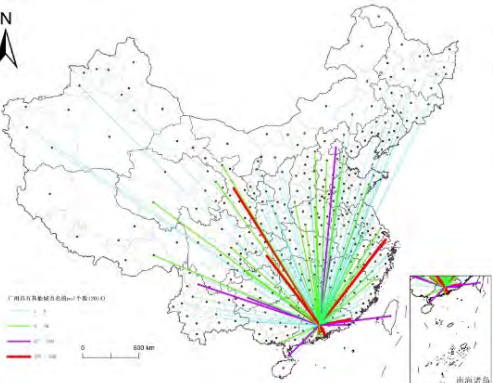
上海2011年



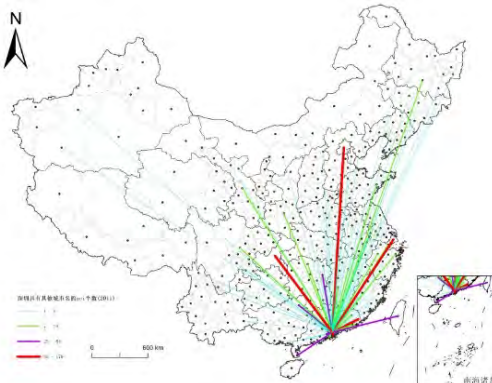
上海2014年



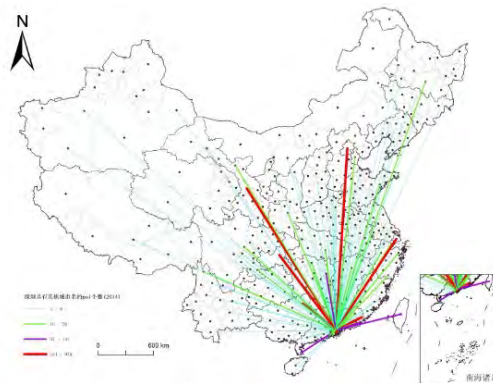
广州2011年



广州2014年



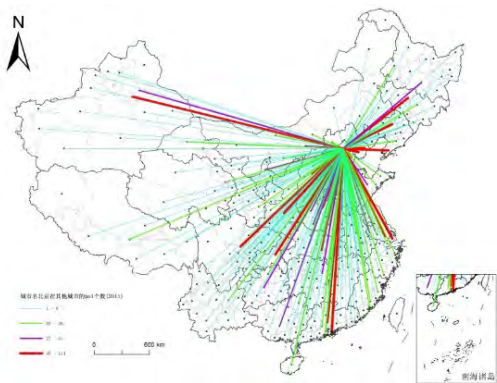
深圳2011年



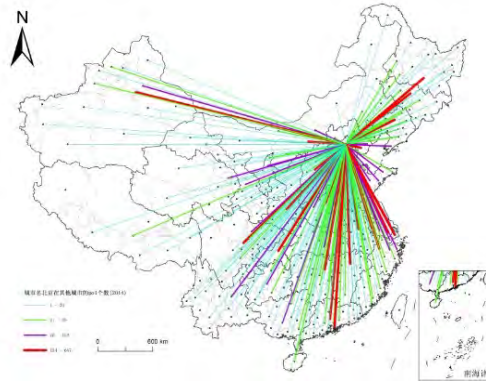
深圳2014年



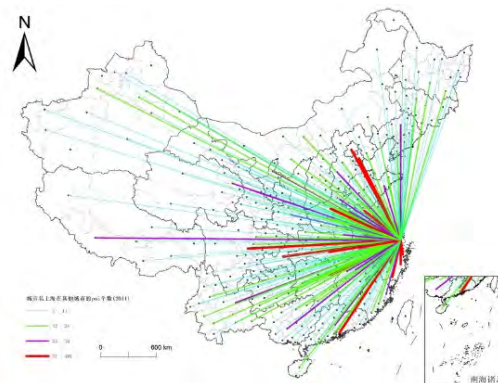
# 高等级城市为核心的中国城市网络入度格局



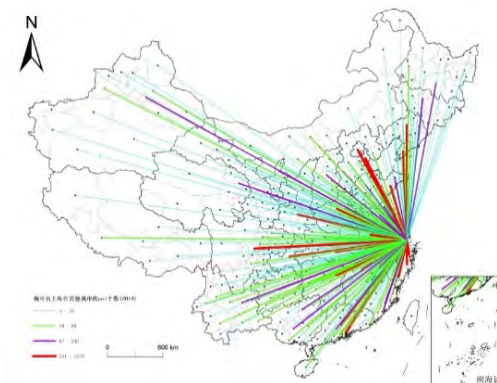
北京2011年



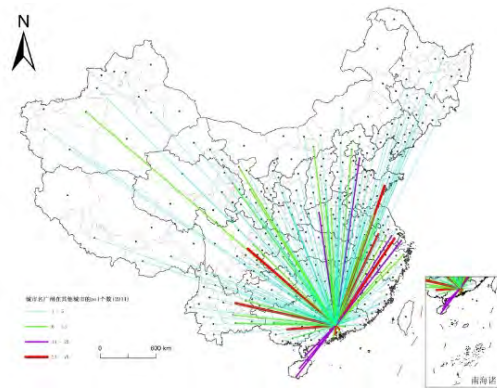
北京2014年



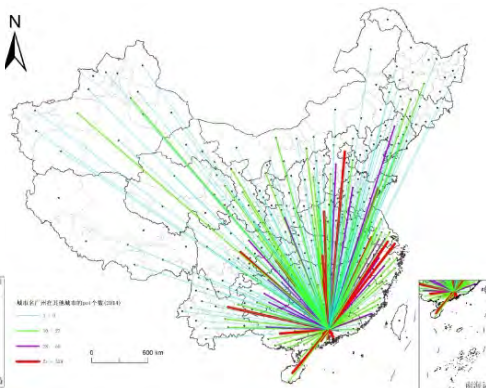
上海2011年



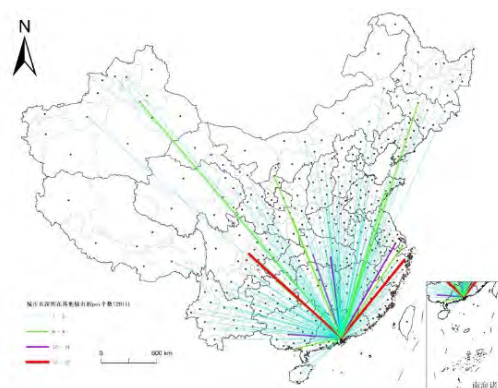
上海2014年



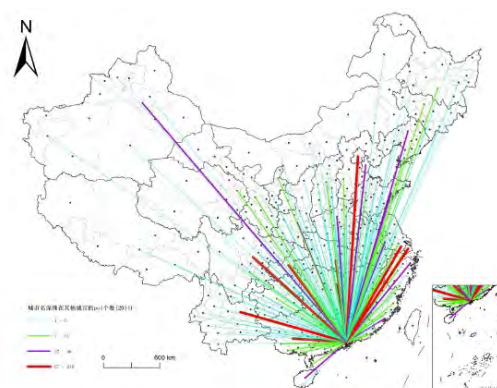
广州2011年



广州2014年

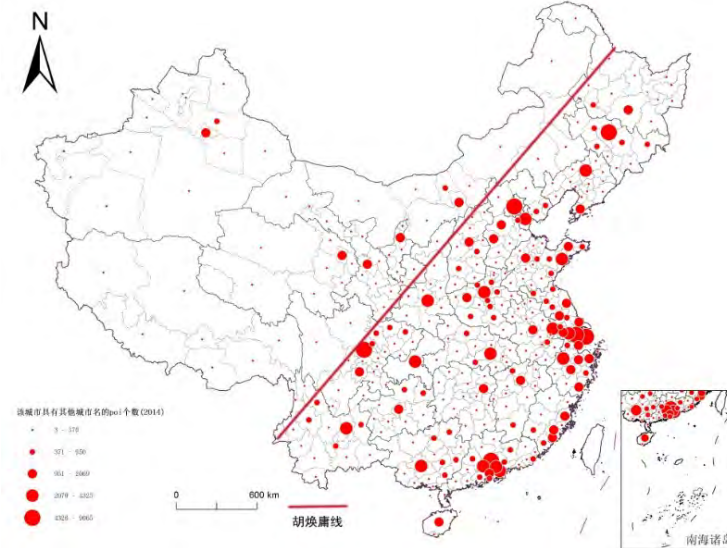
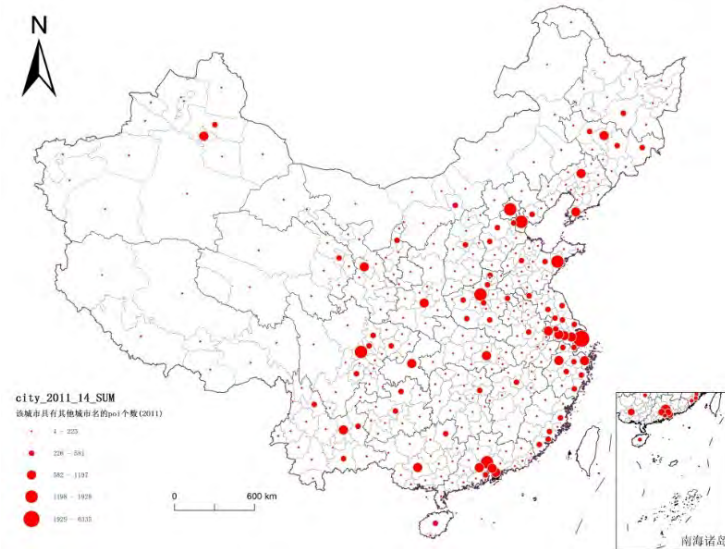


深圳2011年

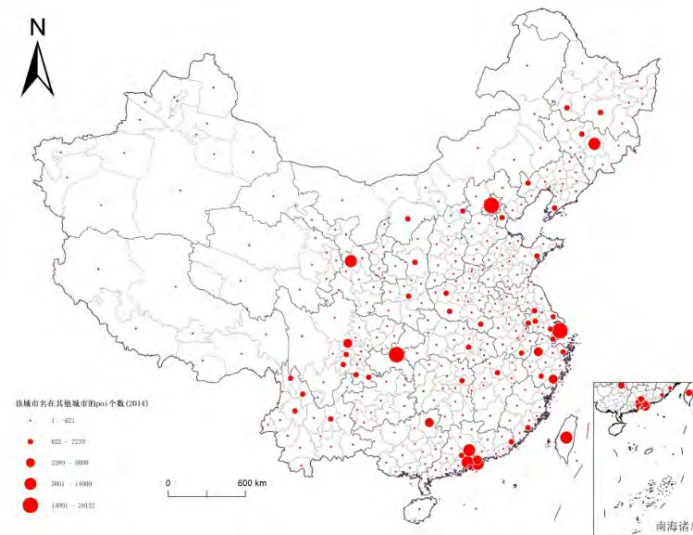
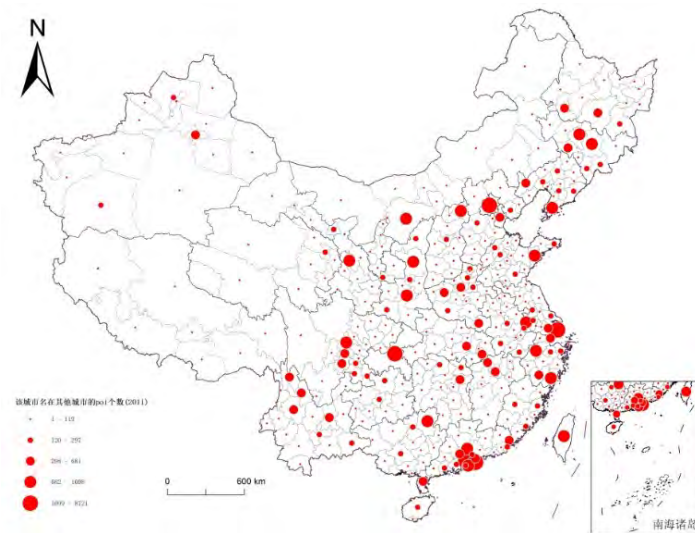


深圳2014年

# 城市网络层级入度/出度特征—基于POI个数

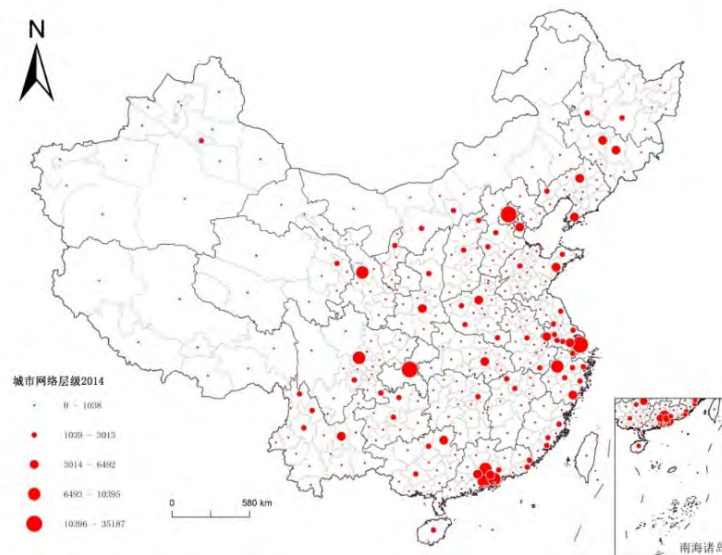
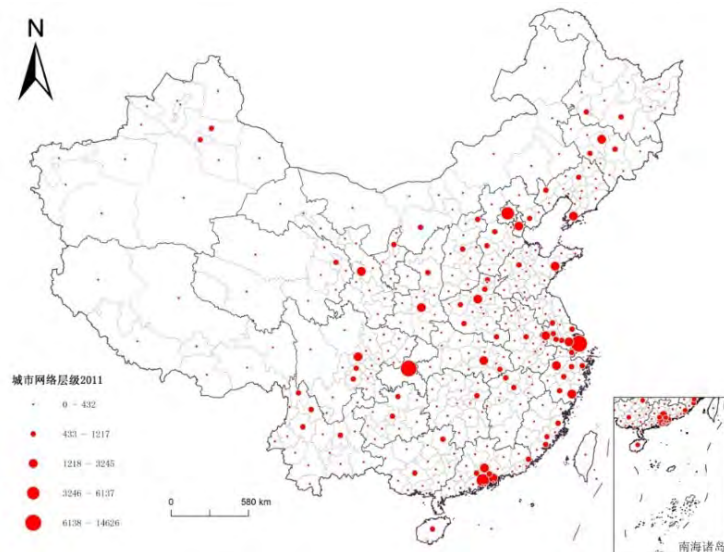


2011年和2014年该城市具有其他城市名的POI个数

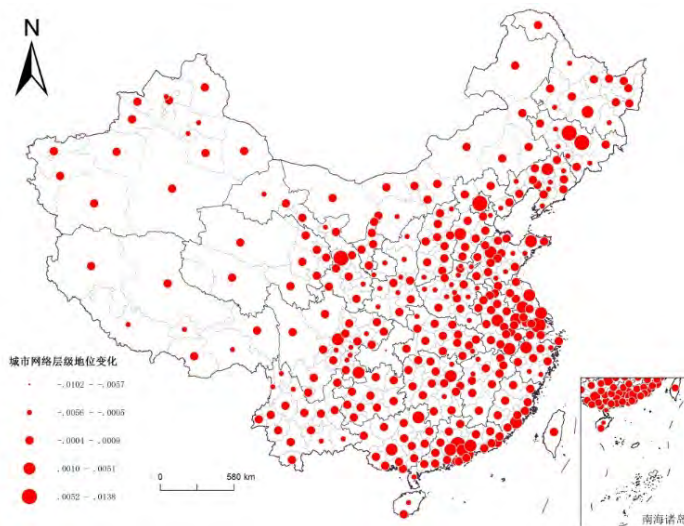
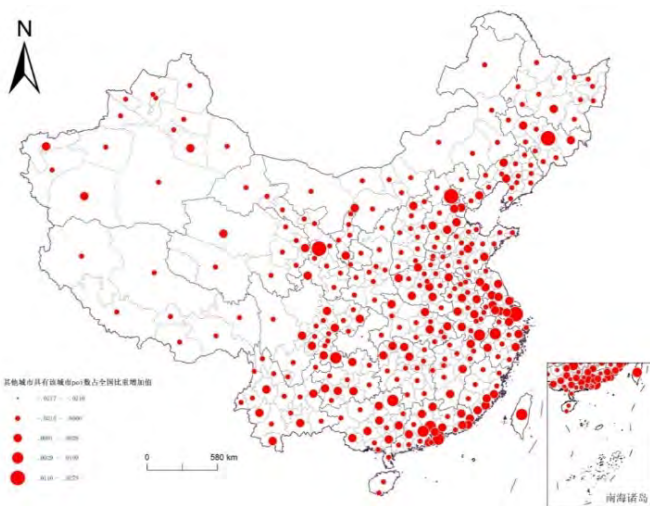


2011年和2014年该城市名在其他城市的POI个数

# 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

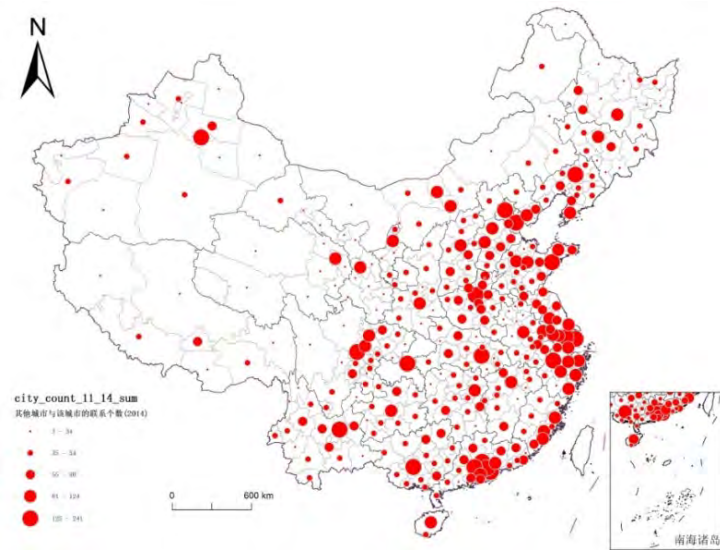
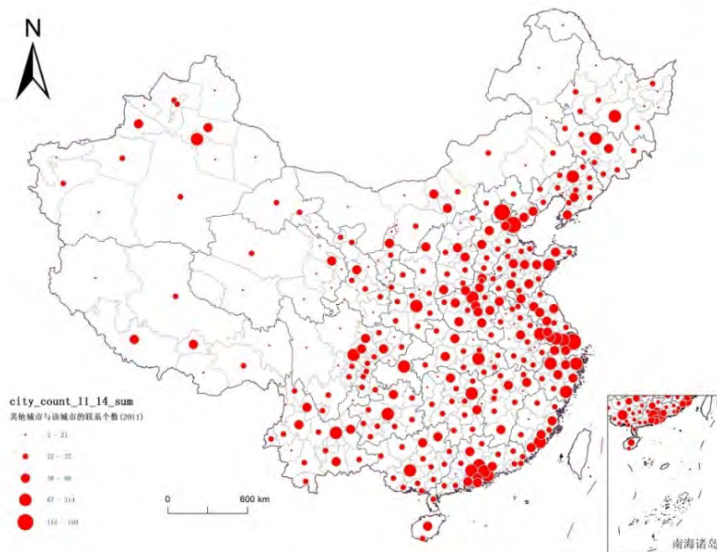


节点城市出度中心性变化

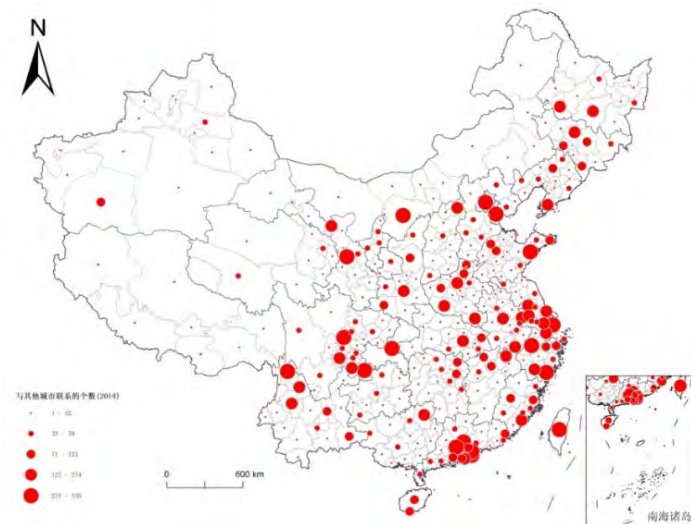
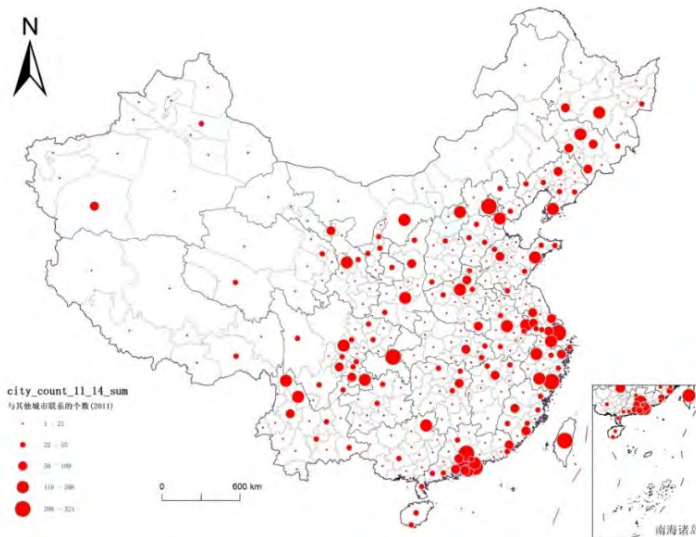
节点城市入度中心性变化

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

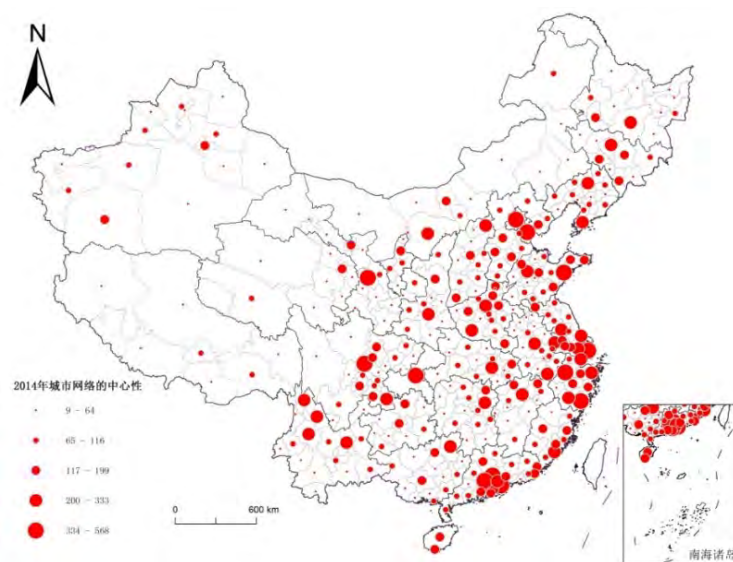
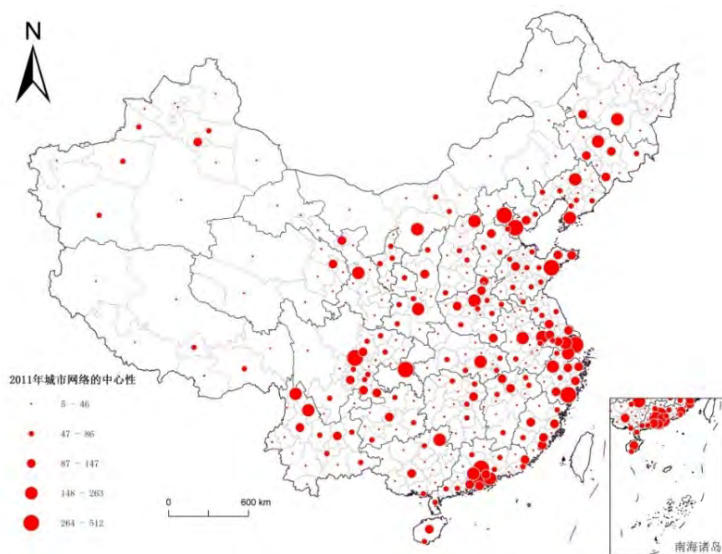


2011年和2014年其他城市与该城市有联系的城市个数

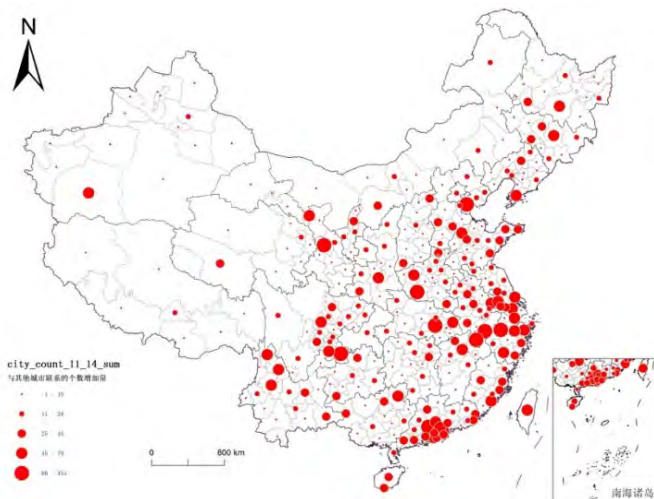


2011年和2014年与其他城市有联系的城市个数

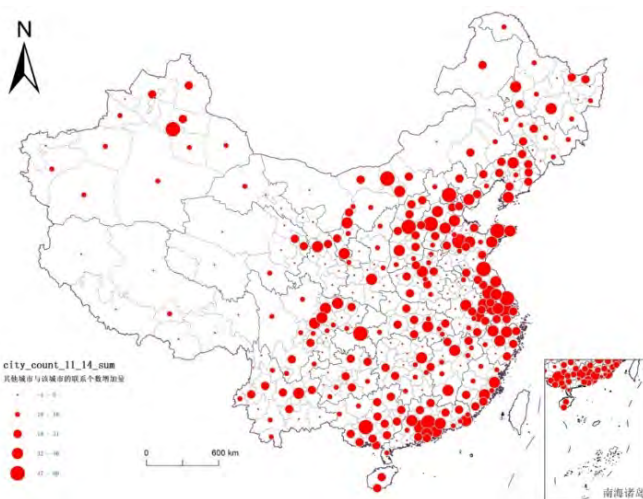
# 城市网络层级分析与比较—基于有联系城市个数



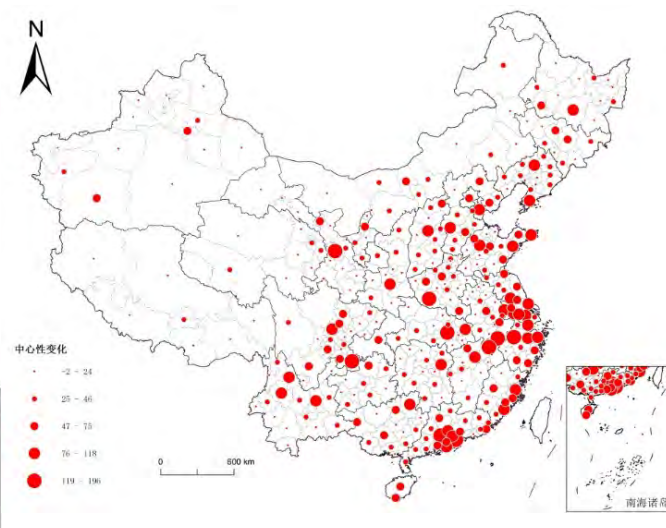
2011年和2014年城市的网络中心格局



节点城市出度中心性变化

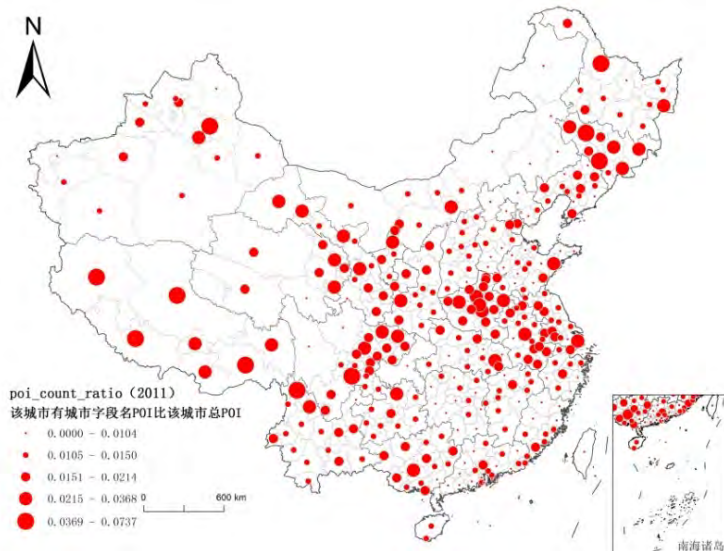


节点城市入度变化



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
- 第七讲：是什么、为什么和怎么办？解释模型与预测模型
  - R square的故事

# 课后安排

- 阅读材料：
  - 城市网络分析案例.doc（偏实际操作，华南理工大学赵渺希赞助）
  - 龙瀛的知识生产与消费论文
- 参考资料
  - 刘行健贡献的课件
  - 其他待补充
- 王江浩沙龙：城市时空大数据分析与可视化
- 答疑
  - [ylong@tsinghua.edu.cn](mailto:ylong@tsinghua.edu.cn)
  - 建筑学院新501办公室（默认每周五上午10:00-11:30）
  - 欢迎前来讨论大作业初步思路





龙瀛, [ylong@tsinghua.edu.cn](mailto: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](https://www.beijingcitylab.com)

清华大学

