

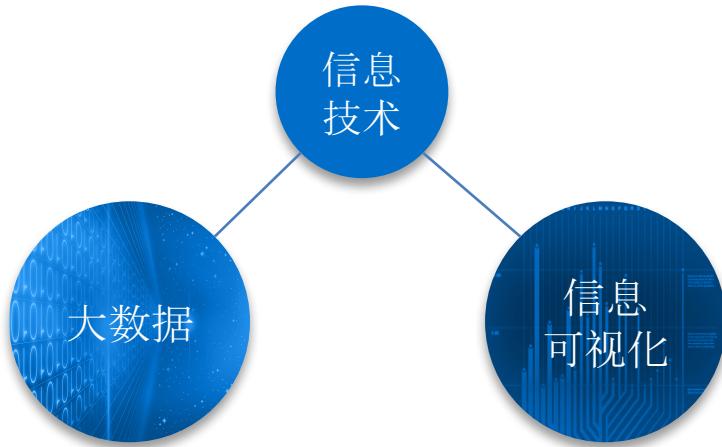


基于信息节点的复杂数字
界面信息关系研究

杨宝玲
2016.3.7



课题背景



大数据

一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合。具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

信息可视化

将抽象信息可视化为形象的图文。

复杂网络

具有自组织、自相似、吸引子、小世界、无标度中部分或全部性质的网络称为复杂网络。

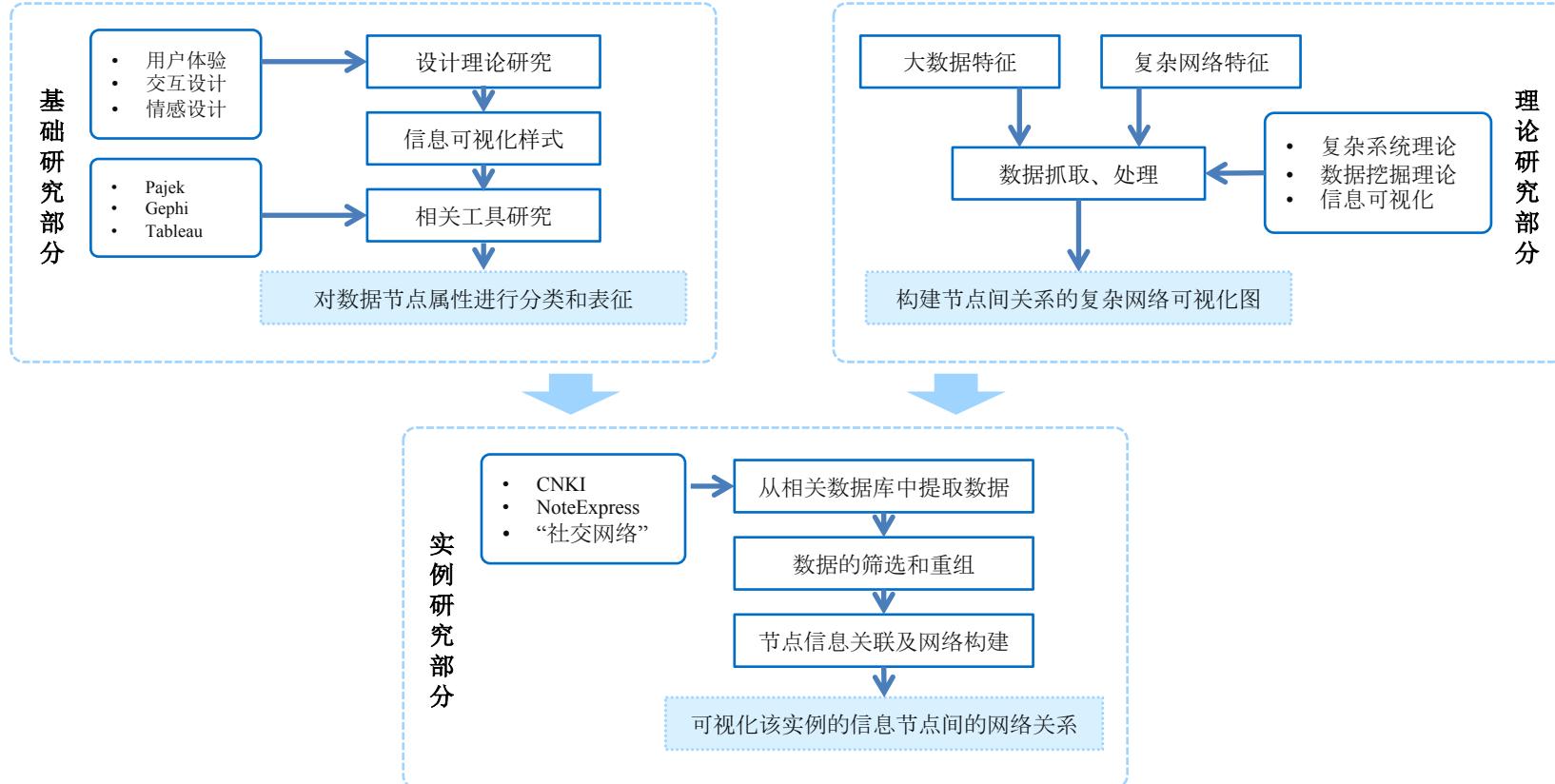
合著网络

合著网络是指学术论文的作者之间合著写作的关系网络，具有自组织、无标度特性，属于复杂网络的一种。

无标度：scale - free network, 现实世界的网络大部分都不是随机网络，少数的节点往往拥有大量的连接，而大部分节点却很少，一般而言他们符合80/20马太定律

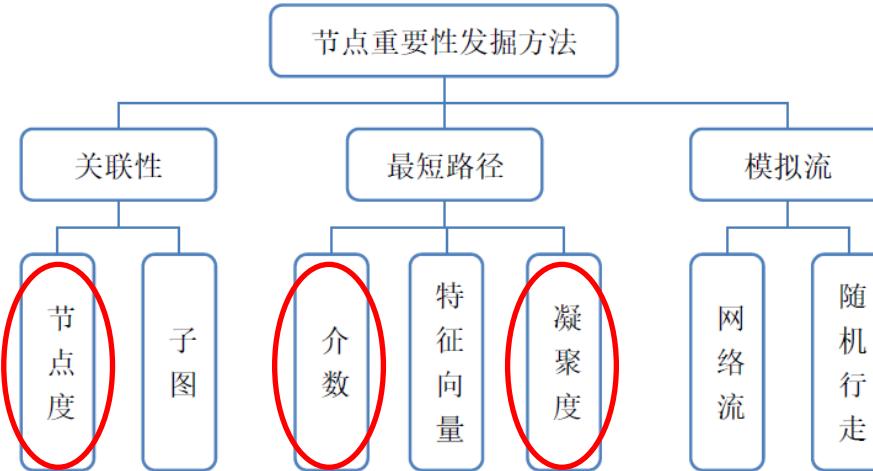
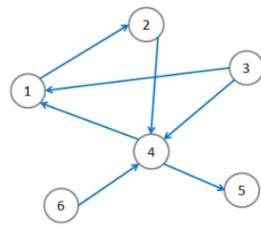


研究内容





节点重要度评价指标



$$\bar{k} = \frac{1}{N} \sum_i k_i = \sum_{n \neq m \neq v} \delta_{nm}(v), \quad \delta_{nm} = \frac{\partial_{nm}(v)}{\partial_{nm}} \frac{2e_i}{k(i)-1}$$

聚集系数：指该节点的邻居节点间连接的紧密程度。定义为与该节点连接的节点之间的真实连接边数与那些节点之间最大可能存在的连接边数之比。



信息可视化模式



图 3.8 TreeMap 可视化图

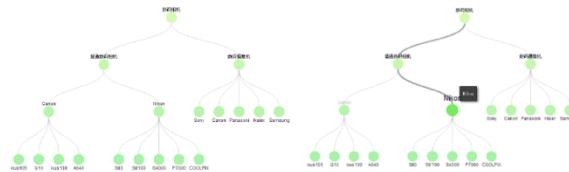


图 3.7 Tree 图

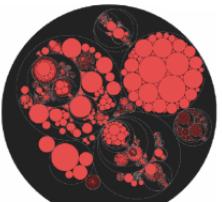
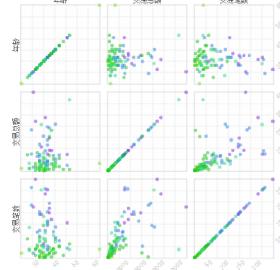


图 3.9 气泡 TreeMap

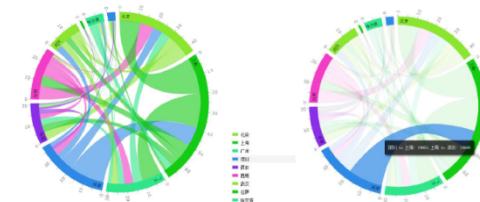


图 3.11 Chord 弦图

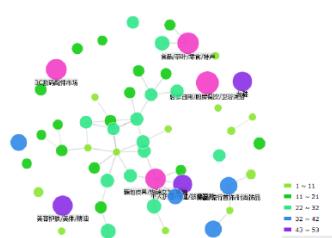
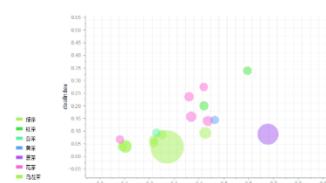


图 3.10 Force 网

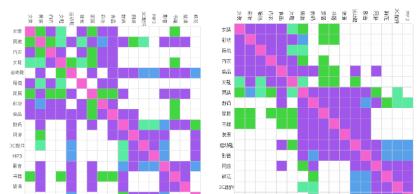


图 3.4 Matrix 图

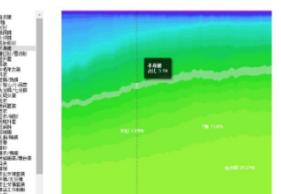


Fig. 2-12. *Carbonyl IR*



用户体验

信息的可视性设计需要满足

- 可察觉性 (detectability)
- 可区辨性 (discriminability)
- 有意义性 (meaningfulness)
- 标准化 (standardization)
- 信息的余备性 (redundancy)

(1)显示信息的**刺激量和刺激强度**等应符合人的感知特性，使用户能够敏锐感觉；

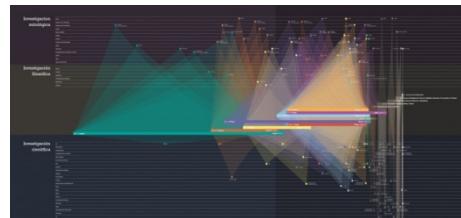
(2)信息的**编码和标志要明确且醒目**，使用户在译码时不会产生错误的感知和判断；



图形：直观，形象，简洁性，易于识别和记忆，具有丰富的情绪感染力



文字：直接而具体的元素，是叙述性符号



色彩：带有情感色彩。可通过深浅、饱和度、透明度的调整，突出主要信息



数据来源

- CNKI中国知网
- NoteExpress
- “社交网络”
- 2010-2015
- 19790条检索结果

The screenshot shows the CNKI search interface. At the top, there are search parameters: '数据库' set to '数据库选择' (Is), '主题' set to '精确' (Exact) and '并且' (And), and '发表时间' set to '从 19XX-XX-XX' (Is). Below the search bar is a table of search results:

| 作者 | 标题 | 来源 |
|----------|----------------------|----|
| 2013 吴保来 | 基于互联网的社交网络研究 | |
| 2013 邢星 | 社交媒体个性化推荐方法研究 | |
| 2013 谈峰 | 位置隐私保护及其在基于位置的社交... | |
| 2013 张春英 | 基于属性限制的社交网络建模与态势分... | |
| 2013 吴联仁 | 基于人类动力学的社交网络信息传播... | |
| 2013 王喆来 | 社交媒体服务平台用户参与激励机制... | |
| 2013 程军军 | 社交媒体中的信息传播模式及话题趋势... | |
| 2012 张彦超 | 社交媒体服务中信息传播模式与舆论... | |

At the bottom, it says '找到: 19790 条 (每页 10 条) 已截图: 850 条 已勾选: 850 条 页面 200: 正在获取数据...'.

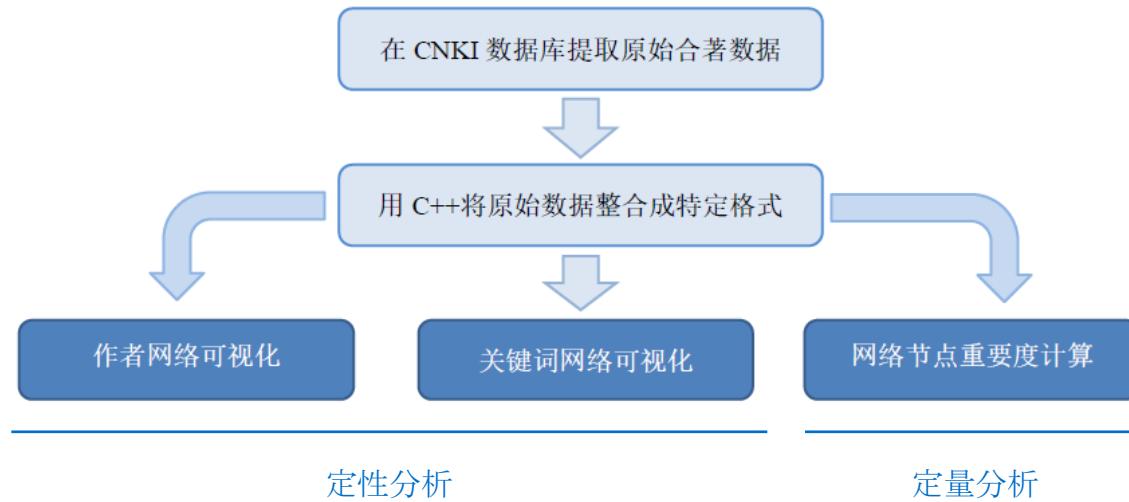
选取2127条检索结果作为研究数据

Reference Type、Author、Year、Title、Publisher、Journal、Volume、Issue、Pages、Date、Keywords、Abstract、Author Affiliation、Database

| Reference Type | Author | Year | Title | |
|-----------------|---------------|-------|--------------------------------|-------|
| Journal Article | 刘娴娴; 封兴中 | 2015 | “互联网+”背景下社交网络发展和营销模式研究 | |
| Journal Article | 梁力予; 任露凌; 万艳华 | 2011 | “基于 LBS 的社交网络”在城市公共空间管理与运营中的应用 | |
| Journal Article | 李惠芳; 殷俊 | 2011 | “宅”文化视角下的社交网络用户体验 | |
| Thesis | 刘一阳 | 2011 | 安全本体在社交网络中的应用研究 | |
| Journal Article | 高华 | 2013 | Web 数据挖掘在社交网络分析的应用研究 | |
| | | | | |



研究方法





统计学分析

论文篇数统计

表 5.2 2010-2015 年“社交网络”的论文篇数

| 文献篇数\年份 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------|------|------|------|------|------|------|
| 学位论文 | 6 | 41 | 69 | 132 | 132 | 71 |
| 期刊论文 | 33 | 56 | 108 | 151 | 323 | 962 |
| 会议论文 | 2 | 3 | 4 | 8 | 3 | 0 |
| 总数 | 41 | 100 | 181 | 291 | 458 | 1033 |

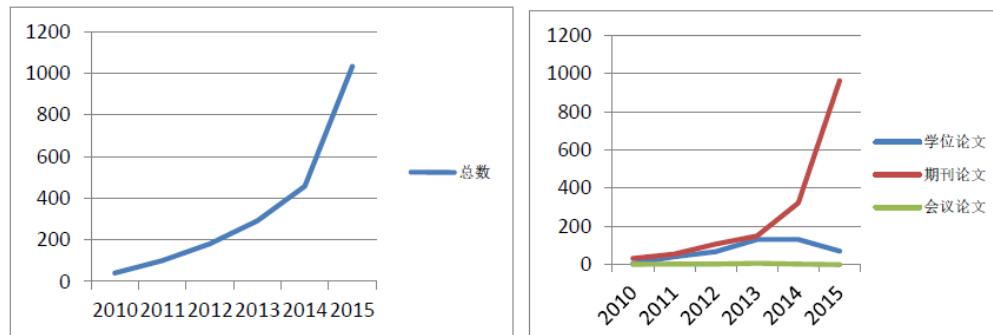


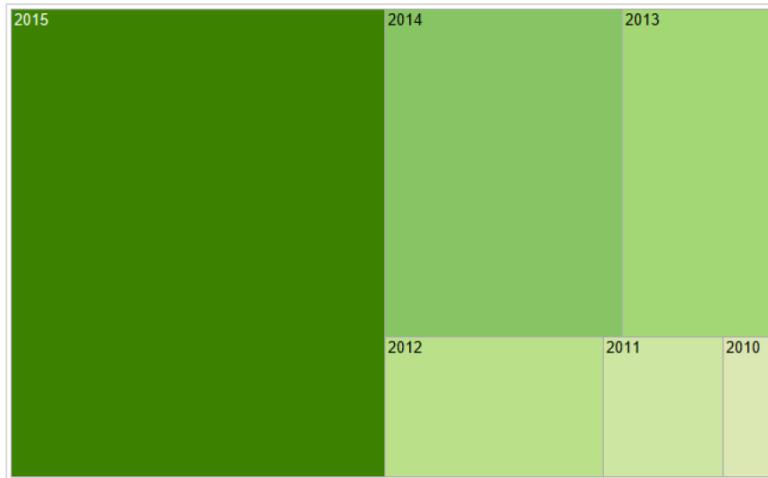
图 5.3 论文发表篇数随年份变化折线图



统计学分析

论文篇数可视化 (Tableau)

以**形状大小**和**颜色深浅**来表示论文数量



Treemap



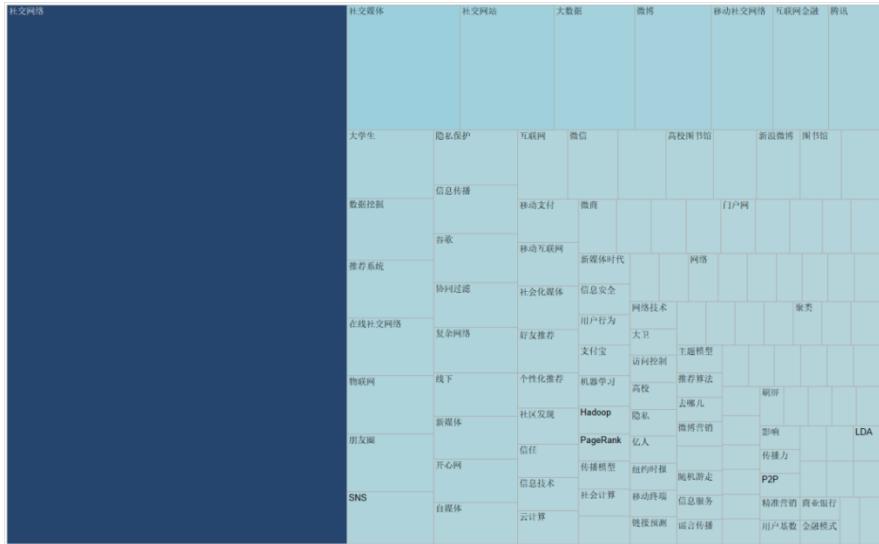
Bubble



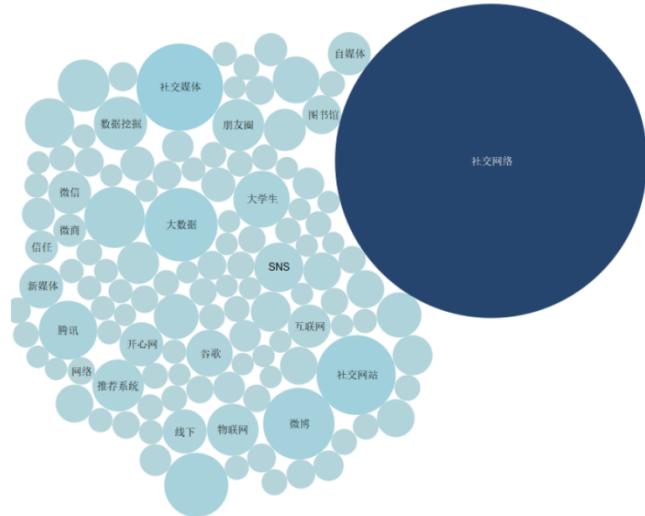
统计学分析

关键词词频可视化 (Tableau)

以**形状大小**和**颜色深浅**来表示论文数量



Treemap



Bubble

The diagram illustrates a complex network of connections between various entities across different categories. The main categories include:

- Students**: Connected to Privacy Protection, Internet, WeChat, and University Library.
- Data Mining**: Connected to Information Propagation, Mobile Payment, Micro-commerce, and Portal Network.
- Recommendation Systems**: Connected to Google, Mobile Internet, New Media Era, and Network.
- Online Social Networks**: Connected to Collaborative Filtering, Social Media, Information Security, User Behavior, and Clustering.
- Internet of Things**: Connected to Offline, Personalized Recommendation, Machine Learning, and刷屏 (Screen Scrolling).
- Friends Circle**: Connected to New Media, Community Discovery, Hadoop, Privacy, Microblog Marketing, Influence, and Spreadability.
- SNS**: Connected to Joyous Net, Trust, PageRank, Ego Person, and P2P.

Specific entities mentioned include: 大学生 (Student), 数据挖掘 (Data Mining), 推荐系统 (Recommendation System), 在线社交网络 (Online Social Network), 物联网 (Internet of Things), 朋友圈 (Friends Circle), SNS, 高校图书馆 (University Library), 移动支付 (Mobile Payment), 微商 (Micro-commerce), 门户网 (Portal Network), 新媒体时代 (New Media Era), 网络 (Network), 聚类 (Clustering), 网络技术 (Network Technology), 大卫 (David), 主题模型 (Topic Model), 访问控制 (Access Control), 推荐算法 (Recommendation Algorithm), 去哪儿 (WhereToGo), 微博营销 (Weibo Marketing), 影响 (Influence), 传播力 (Spreadability), 纽约时报 (New York Times), 随机游走 (Random Walk), 移动终端 (Mobile Terminal), 信息服务 (Information Service), 精准营销 (Precision Marketing), 商业银行 (Commercial Bank), 用户基数 (User Base), 金融模式 (Financial Model), 云计算 (Cloud Computing), 自媒体 (Self-Media), and LDA.



合著可视化 (Pajek)

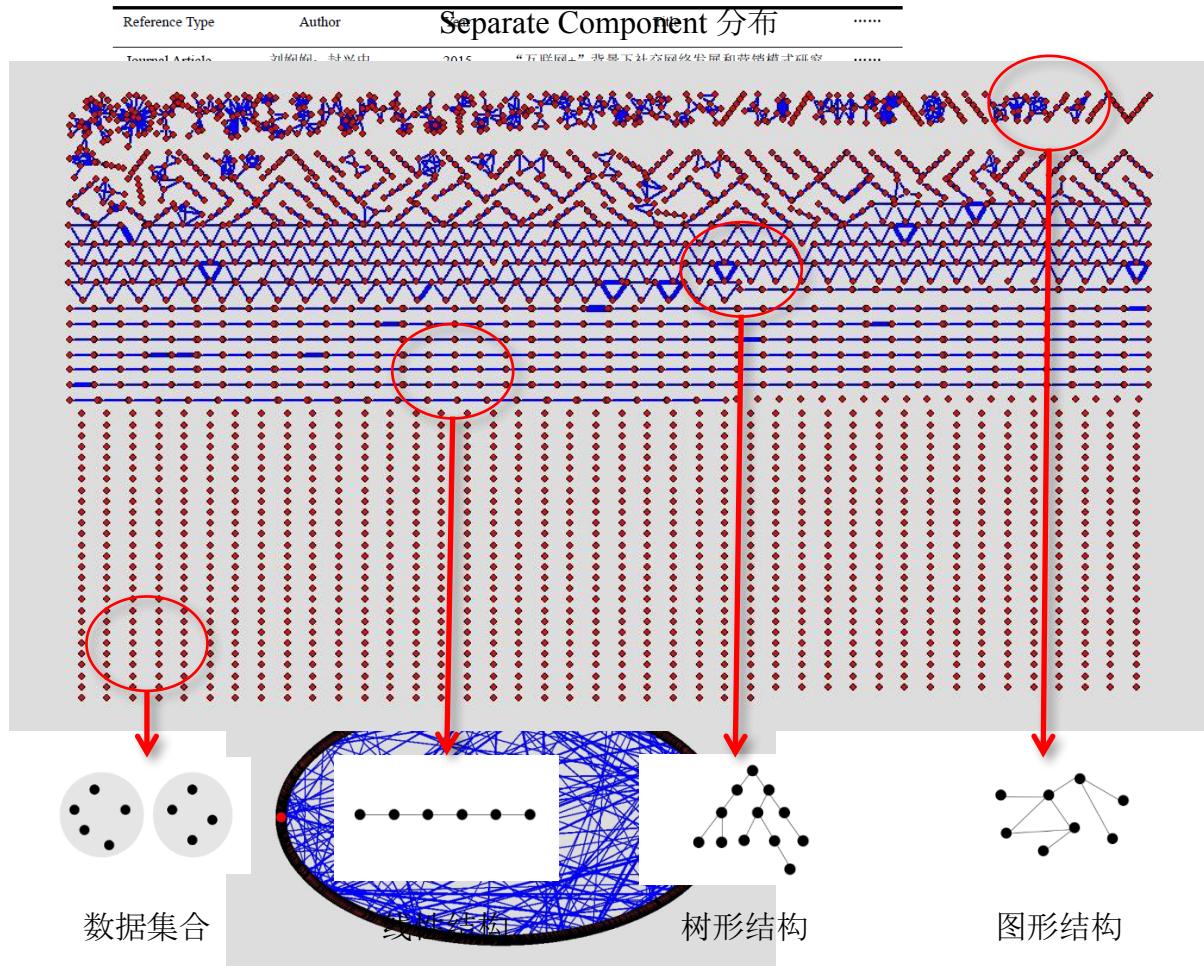
NoteExpress 提取原始数据



C++处理数据为.net格式



Pajek读取绘图





节点重要度计算

节点度前20名

| 作者节点 | 节点的度 | 作者节点 | 节点的度 |
|------|------|------|------|
| 汤庸 | 22 | 冯登国 | 9 |
| 程学旗 | 15 | 贾焰 | 9 |
| 王磊 | 11 | 詹剑锋 | 8 |
| 周斌 | 11 | 高婉铃 | 8 |
| 袁成哲 | 11 | 李经伟 | 8 |
| 张莹 | 11 | 魏凯 | 8 |
| 王剑 | 10 | 罗纯杰 | 8 |
| 方滨兴 | 9 | 韩锐 | 8 |
| 周涛 | 9 | 田昕晖 | 8 |
| 马建峰 | 9 | 姜春宇 | 8 |

节点介数前20名

| 作者节点 | 节点介数 | 作者节点 | 节点介数 |
|------|----------|------|----------|
| 汤庸 | 3.55E-05 | 刘晶 | 1.10E-05 |
| 方滨兴 | 3.46E-05 | 潘善亮 | 1.02E-05 |
| 周斌 | 2.69E-05 | 郭静 | 1.02E-05 |
| 贾焰 | 1.79E-05 | 李石君 | 9.78E-06 |
| 程学旗 | 1.57E-05 | 周涛 | 9.37E-06 |
| 吴信东 | 1.49E-05 | 刘艳 | 8.96E-06 |
| 许进 | 1.47E-05 | 覃俊 | 8.96E-06 |
| 张鹏 | 1.25E-05 | 张莹 | 8.67E-06 |
| 汪浩 | 1.22E-05 | 吴斌 | 8.55E-06 |
| 何静 | 1.14E-05 | 王剑 | 8.55E-06 |

节点度前20名的聚集系数

| 作者节点 | 聚集系数 | 作者节点 | 聚集系数 |
|------|----------|------|----------|
| 汤庸 | 0.186147 | 冯登国 | 0.388889 |
| 程学旗 | 0.209524 | 贾焰 | 0.333333 |
| 王磊 | 0.563636 | 詹剑锋 | 1 |
| 周斌 | 0.236364 | 高婉铃 | 1 |
| 袁成哲 | 0.4 | 李经伟 | 1 |
| 张莹 | 0.218182 | 魏凯 | 1 |
| 王剑 | 0.533333 | 罗纯杰 | 1 |
| 方滨兴 | 0.444444 | 韩锐 | 1 |
| 周涛 | 0.277778 | 田昕晖 | 1 |
| 马建峰 | 0.333333 | 姜春宇 | 1 |



合著可视化 (Gephi)

表 5.5 Gephi 读取的节点文件

| ID | Label | Modularity class |
|-------|-------|------------------|
| 1 | 李影 | |
| 2 | 邵爽 | |
| 3 | 陈伟英 | |
| 4 | 石丹 | |
| 5 | 卜玉敏 | |
| 6 | 高正波 | |
| | | |

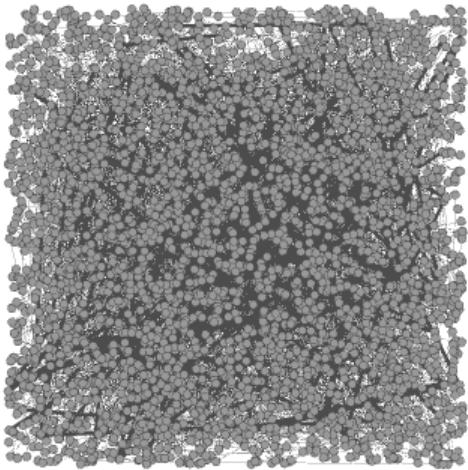
表 5.6 Gephi 读取的边文件

| Source | Target | Type | ID | Label | Weight |
|--------|--------|------------|-------|-------|--------|
| 1 | 李影 | Undirected | 1 | | 3 |
| 2 | 邵爽 | Undirected | 2 | | 3 |
| 3 | 陈伟英 | Undirected | 3 | | 3 |
| 4 | 石丹 | Undirected | 4 | | 3 |
| 5 | 卜玉敏 | Undirected | 5 | | 2 |
| 6 | 高正波 | Undirected | 6 | | 2 |
| | | | | | |

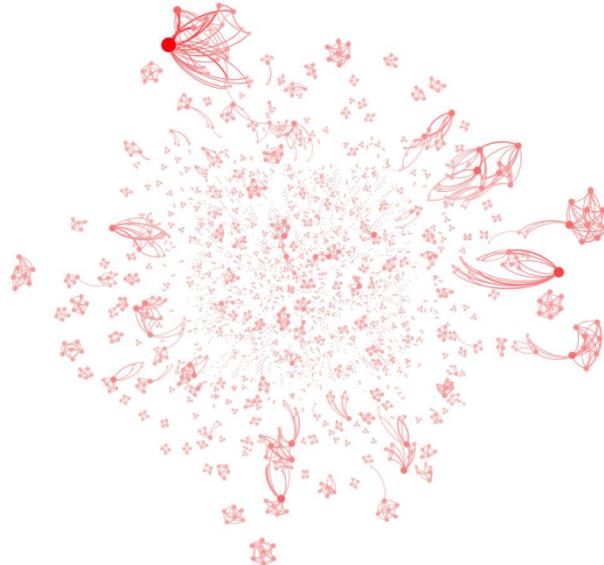


合著可视化 (Gephi)

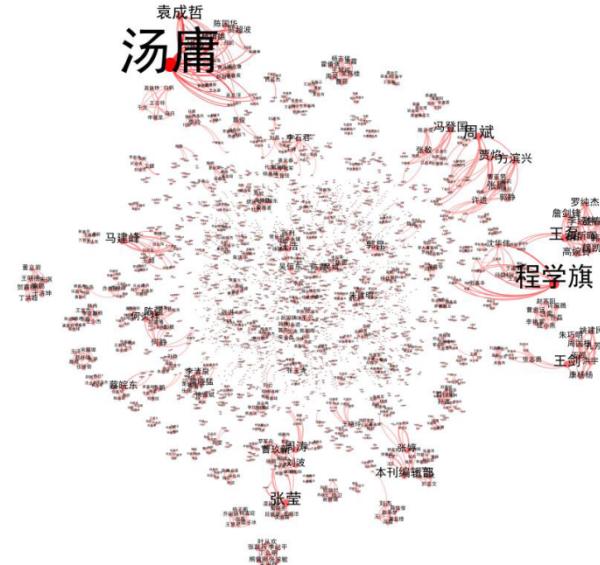
原始可视化图



颜色深浅和节点大小表示节点度数

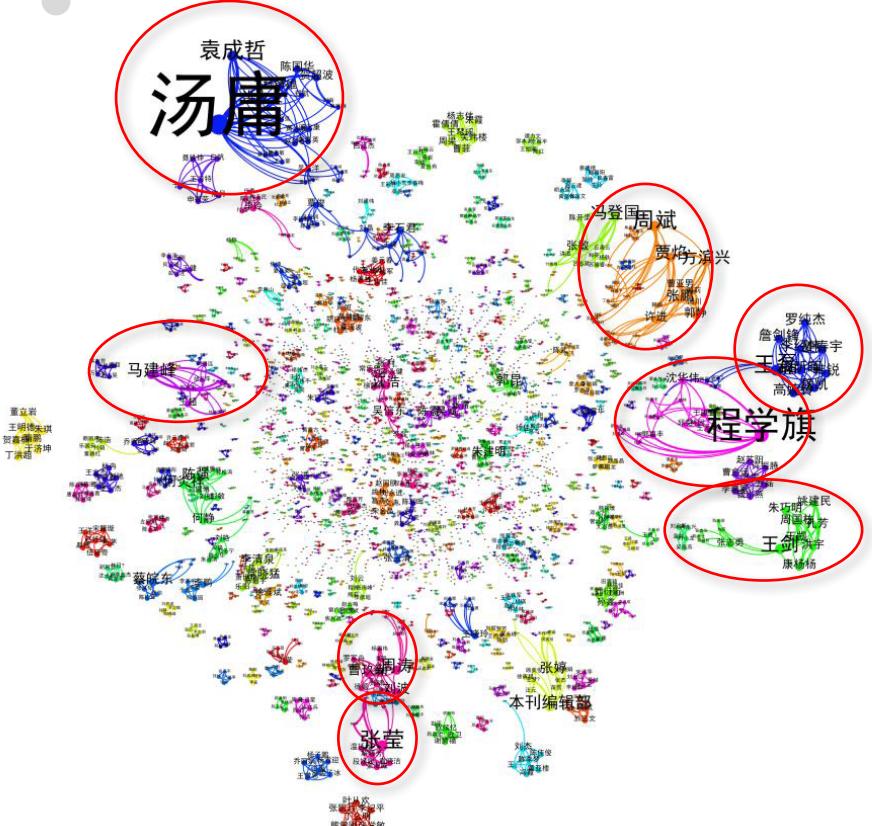


带标签网络

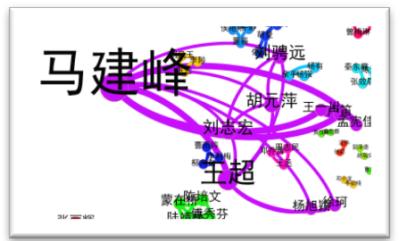
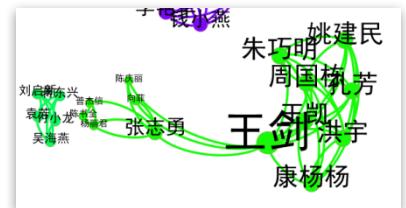
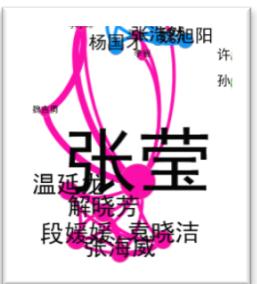
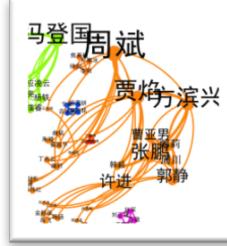
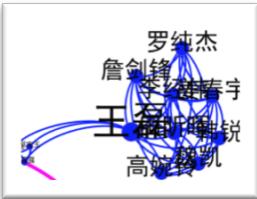
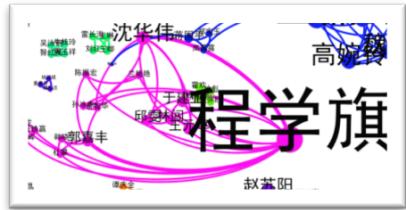
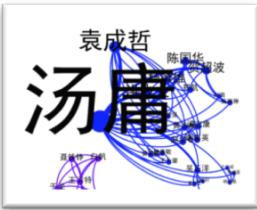




合著可视化 (Gephi)



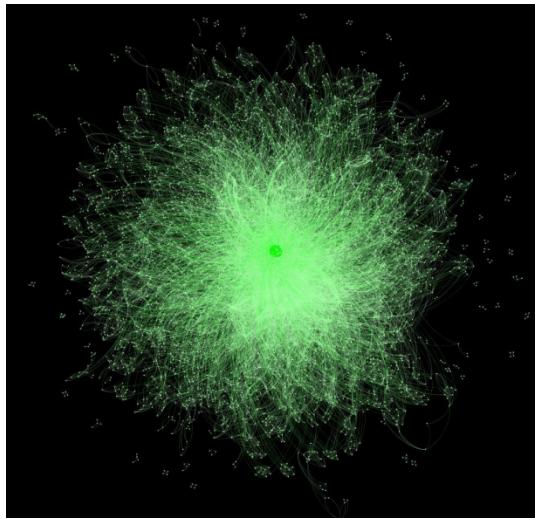
用颜色区分不同团队



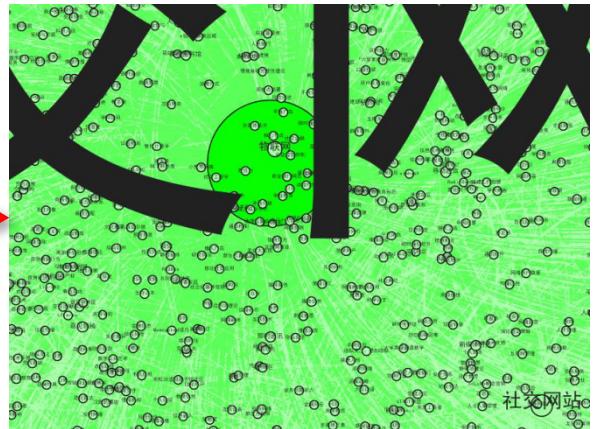
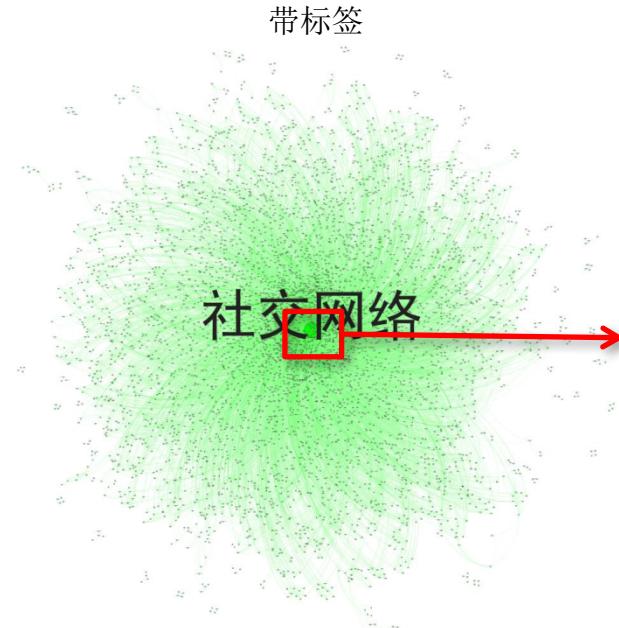


关键词可视化

无标签



带标签





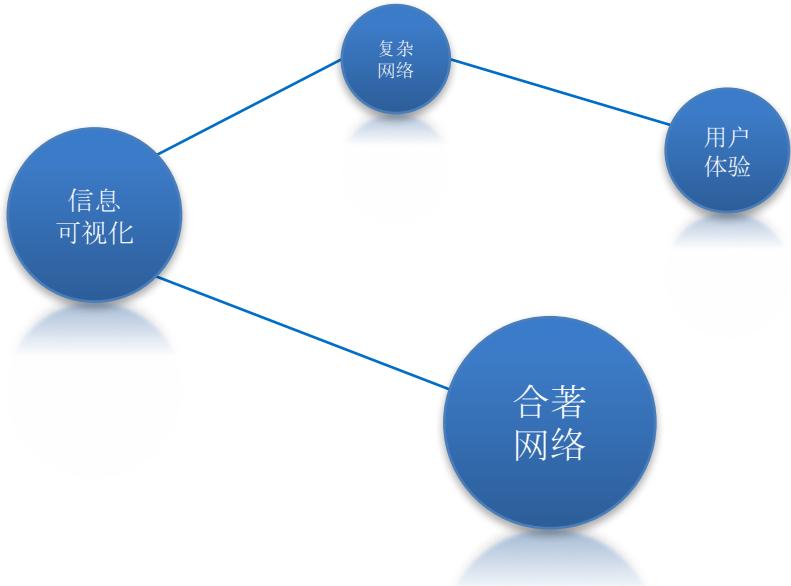
总结分析

检索到文献**19790篇**

提取**2127篇**用于研究

涉及**作者1377人**

关键词5283个



定性分析

+

定量分析

可视化

节点参数计算

该网络中较重要的作者有：

汤庸
程学旗
周斌
周涛
方滨兴
张莹
王剑

论文中出现次数较多的关键词：

社交网络
社交媒体
社交网站
大数据
微博
移动社交网络
互联网金融

Thank you

杨宝玲
2016.3.7