



数据挖掘与商务分析

第8讲 社会网络分析与商务实践

主讲教师：肖升生

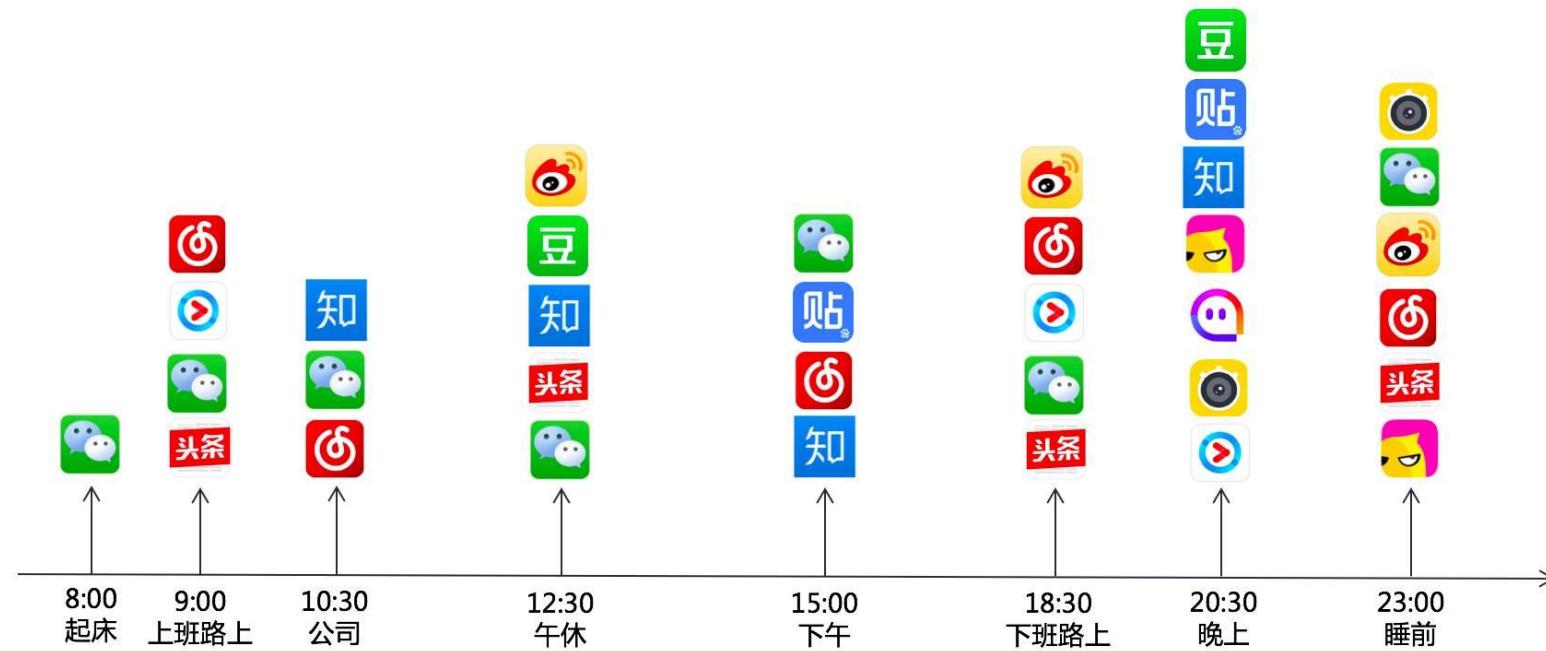


课程导入：身边的社交网络





课程导入：中国的社交媒体





讲授提纲

- 01 概述**
- 02 图论基础**
- 03 社群识别**
- 04 影响力度量**
- 05 商务案例-社会网络分析**



讲授提纲

- 01 概述**
- 02 图论基础**
- 03 社群识别**
- 04 影响力度量**
- 05 商务案例-社会网络分析**



社交网络分析

社交媒体（Social Media）指互联网上基于用户关系的内容生产与交换平台。

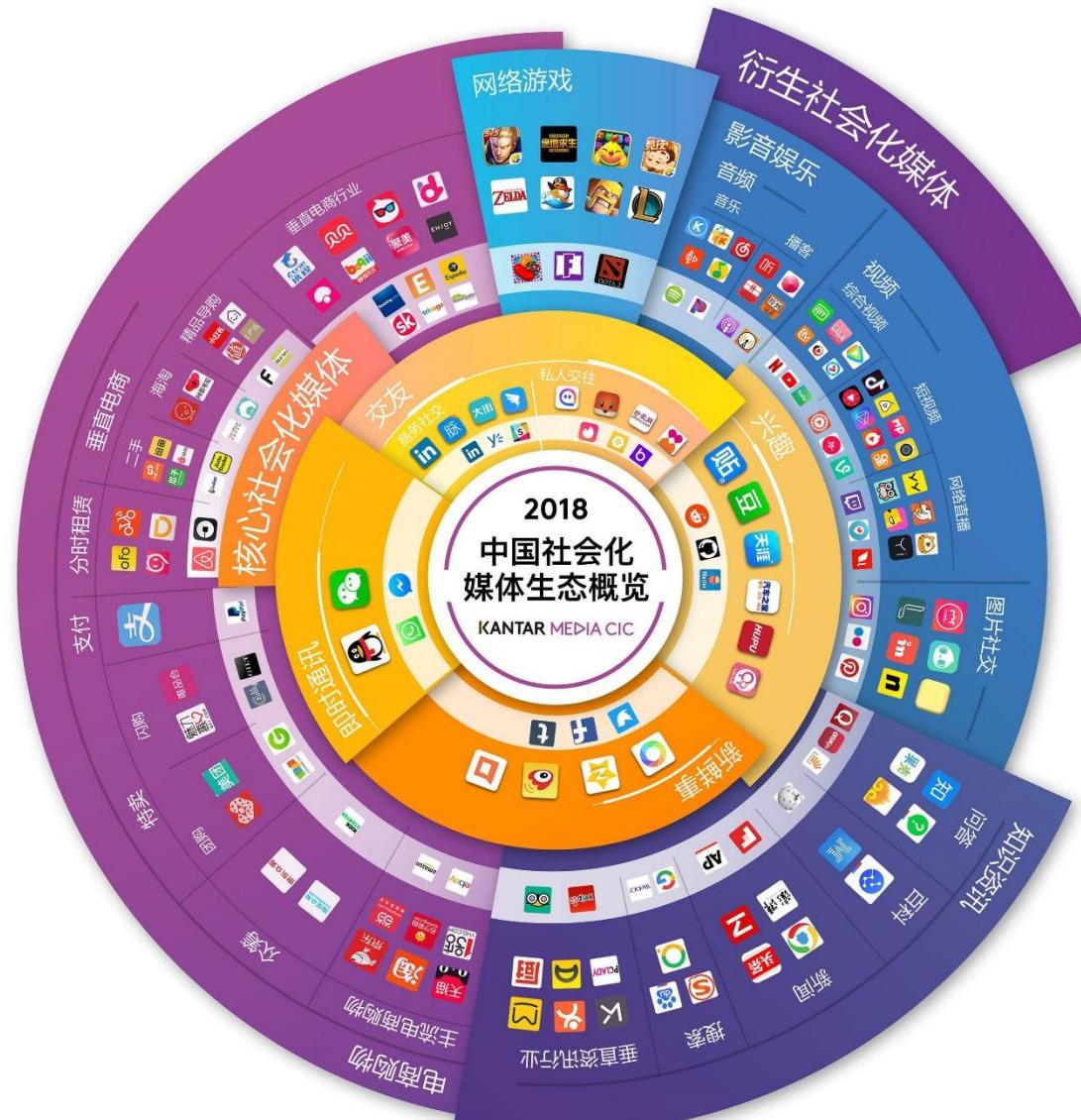
- 社交媒体是人们彼此之间用来分享意见、见解、经验和观点的重要渠道
- 社交媒体是企业进行广告宣传、提供客户服务、收集顾客反馈和竞争情报的重要平台

问：为什么研究社交媒体、分析社交网络？

- 社交媒体上用户行为及网络构成怎样，社交网络如何变化，消息是如何在社交网络上扩散的？
- 如何利用社交媒体数据实施智能服务？如广告营销等



社交媒体应用概览





社交媒体类型

传统社交网络：人人网，开心网



即时通讯：微信，陌陌



点评：大众点评，豆瓣



博客：新浪博客，CSDN博客



微博：新浪微博，腾讯微博



百科：百度百科，维基百科



问答：知乎，百度知道



视频：优酷，哔哩哔哩



短视频：抖音，快手





传统社交网络

■ 用户可以在线上与其线下好友、熟人交流

■ 交互方式

- 好友交流：评论、喜欢……
- 公开信息分享
- 私信

■ 例子

- 脸书
- 人人网
- 开心网





即时通讯

- 能够即时发送和接收互联网消息等的工具
- 早年集中于PC端应用；随智能手机普及，现主要为手机端应用
- 部分替代传统社交网络网站功能，尤其在移动社交领域
 - 熟人社交：微信“朋友圈”功能
 - 陌生人社交：陌陌“附近的人”功能





点评

- 点评网站用户对现有产品、服务、娱乐、商业、场所等进行主观评论和打分
- 例子

- 大众点评
- 马蜂窝
- 豆瓣

火宴山火锅(望京店)

商户详情

★★★★★ 人均: ¥61
口碑: 环境: 服务:

地址: 泰通东大街6号泰方恒购物中心4楼
电话: 84788888

用户点评

luoyao1110 | 喜 悅 | 2015-07-26
★★★★☆ 人均: ¥78
就在老公公司旁边，地方很大，我是盛子的粉丝，东西还算丰富吧，就是有的时候会有点不齐，要煮的东西多，肉质有点淡，但味道还是不错的。

商户图片

蜂蜂点评 (共有1861条真实评价)

全部 | 有图 (862条) | 好评 (1643条) | 中评 (134条) | 差评 (84条) | 人很多 (461人提及) | 排队30分钟 (366人提及)

适合小朋友 (340人提及) | 玩一天 (321人提及) | 强烈推荐 (288人提及) | 金牌点评 (138条)

神经病的世界太美好 ★★★★★
LV.5
4.30日去的，六点多就到了排队人也超多，真的没有淡旺季那一说，下车就各种跑，到了排队的地方就是等，好不容易进园了，拿到纸质门票通过迪士尼的app扫门票上的二维码只抢到了一张快速通行证，其他项目就是排队，到处都是人基本要排两三个小时左右，排个队就累够呛，虽然烟花秀挺精彩的，圆一圆朋友的儿时梦，但是真的不想去第二次了，景区未开封的水和食品是可以少量带进去的，里面的食物一般量很少，一根热狗40，冰激凌40，面条和一杯雪碧是套餐，90左右，量少根本没吃饱，迪士尼下来后面的行程都没缓过来，如果选择去最好放在后面的行程去

ZOL.com.cn

10



博客

■ 博客是一种日志网站，供用户(即博主)贡献文本和多媒体内容，通常按时间倒序排列

功能

- 与朋友/陌生人分享信息和观点
- 传播特定主题相关的内容

例子

- 新浪博客
- CSDN博客

The screenshot shows a CSDN blog post by Asia-Lee. The post title is "社交网络影响力最大化". It was published on March 23, 2018, at 22:26:36, with 10311 views and 15 likes. The author is an original member and an enterprise employee. The post has 8335 comments, 1205 likes, 894 shares, 444 reviews, and 4866收藏 (Collection). The content is organized into a目录 (Table of Contents) with sections: 1. 社交网络概述, 2. 影响力最大化问题分类, 3. 社交网络影响力最大化作用, 4. 传播模型, 4.1 独立级联模型 (Independent Cascade Model) 简称 IC 模型, and 4.2 线性阈值模型 (Linear Threshold Model) 简称 LT 模型. At the bottom, there are buttons for点赞 (Like), 评论 (Comment), 分享 (Share), 收藏 (Collection), 手机看 (Mobile View), 打赏 (Tip), and 关注 (Follow).



微博

■ 微博可以被视为发表内容受限的博客

■ 用途

- 传播媒介
- 粉丝互动
- 发布新闻

■ 例子

- 新浪微博
- 腾讯微博

Real-time Hot Search

考神超凡炒饭

微博热搜榜

考神超凡炒饭

相关用户

实时 好友 热点 名人 潮流

实时搜索热点，每分钟更新一次

1 考神超凡炒饭 6666666

2 唐鹤德 557457

3 江西惊现情侣鞋 545356

4 奇葩说 何炅 536455

5 张亮女儿 535786

6 徐娇被吐槽胖了 395787

7 史上尺度最大反腐剧 332263

81192 255567

微博热搜榜

3分钟前 来自微博 weibo.com

超凡炒饭今日，据网友爆料，被万千考生
考神的超凡炒饭终于开通微博了！万千粉丝纷纷
去考神超凡炒饭的微博下面祈求保佑，考神也沒
有让大家失望，一一回复了大家的困惑。粉丝们感
动落泪：厉害了我的考神！

转发 评论 赞

微博神吐槽



百科&问答

■ 知识类社交媒体应用，汇聚网络用户的集体智慧

- 百科：以主题组织，用户可协同编辑
- 问答：以提问组织，用户可分别回答

■ 例子

- 百科：百度百科，维基百科
- 问答：知乎，百度知道

Baidu 百科 社交网络 进入词条

★ 收藏

99+

编辑 讨论 上传视频

社交网络 (互联网概念名词)

社交网络即社交网络服务，源自英文SNS (Social Network Service) 的翻译，中文直译为社交网络服务，意译为。

社交网络含义包括硬件、软件、服务及应用，由于四字构成的词组更符合中国人的构词习惯，因此人们习惯上用代指SNS (Social Network Service)。

中文名	社交网络	又 名	社交网络服务
外文名	Social Network Service	源 自	英文SNS的翻译
		含 义	硬件、软件、服务及应用

知乎 首页 发现 等你来答 海边拍婚纱照被海浪卷走

社交网络 心理学 人际交往 社交产品 互联网社交

社交网络对真实生活中的人际关系会有什么影响？

圆桌收录 • 谁不孤独

最近贾樟柯导演拍了新短片，讲 3 个年轻人因为陌陌找到同类化解孤独。目前社交网络越来越发达这种突破地域和空间的限制，强调人与人之间的连接，是否会对人...显示全部

关注问题 写回答 邀请回答 好问题 8 5条评论 分享 修

查看全部 190 个回答

KnowYourself 心理学话题下的优秀回答者

编辑推荐

Byron 等 1,299 人赞同了该回答

1996年，Sherry Turkle第一次在TED演讲时，她的题目是“庆祝我们在互联网上的生活”，但如今，她却认为网络正在“带我们去到我们不愿意去的地方”。尽管互联网时代给人们提供的社交途径越来越多，人与人的关系却好像越来越疏远，我们更不愿意与人亲密，更少会邀请人来家里做



视频&短视频

■ 娱乐类社交媒体应用，用户可上传内容、并进行评论、点赞等交互（短视频应用限制上传视频长度）

■ 例子

- 视频：优酷，哔哩哔哩
- 短视频：抖音，快手

【核爆神曲】听听 现场版 燃爆全球的动漫神曲【主唱以及作曲者：泽野弘之 花泽香菜 南...
音乐 > 音乐综合 2019-01-02 00:51:11 全站排行榜最高第15名
258.2万播放 · 6.1万弹幕

一开头这一定是青鸟 哪都有你
温馨提示：打开字幕CC，有歌词
歌—青鸟
木叶飞舞之处。火亦生生不息
飛翔いだら 庚らないといって

飞翔(はばた)いたら 庚(もど)らないと言(い)って
如果振翅高飞 我说过我不会回来
惊了150个视频

实时热点
新鲜多元 一触即达

抖音热点 / 热门 新鲜 多元 /

热点榜	正能量
1. 阅兵车牌号是1949和2019	982.6w
2. 李现的第一个抖音	864.5w
3. 中国机长与当事乘客重逢	854.3w
4. 武汉军运会	799.3w
5. 把猫给我	753.2w
6. 跳精舞门召唤罗志祥	701.5w
7. 上海时装周	683.4w
8. 【化物语】主题曲 恋爱循环 现场版【主	650.6w
9. 【某科学的超电磁炮】主题曲 Only My	626.3w
10. 【火影忍者】主题曲 青鸟 现场版【主	



小结



社交媒体应用



社交网络

即时通讯

点评

博客&微博

百科&问答

视频&短视频

.....



讲授提纲

01 概述

02 图论基础

03 社群识别

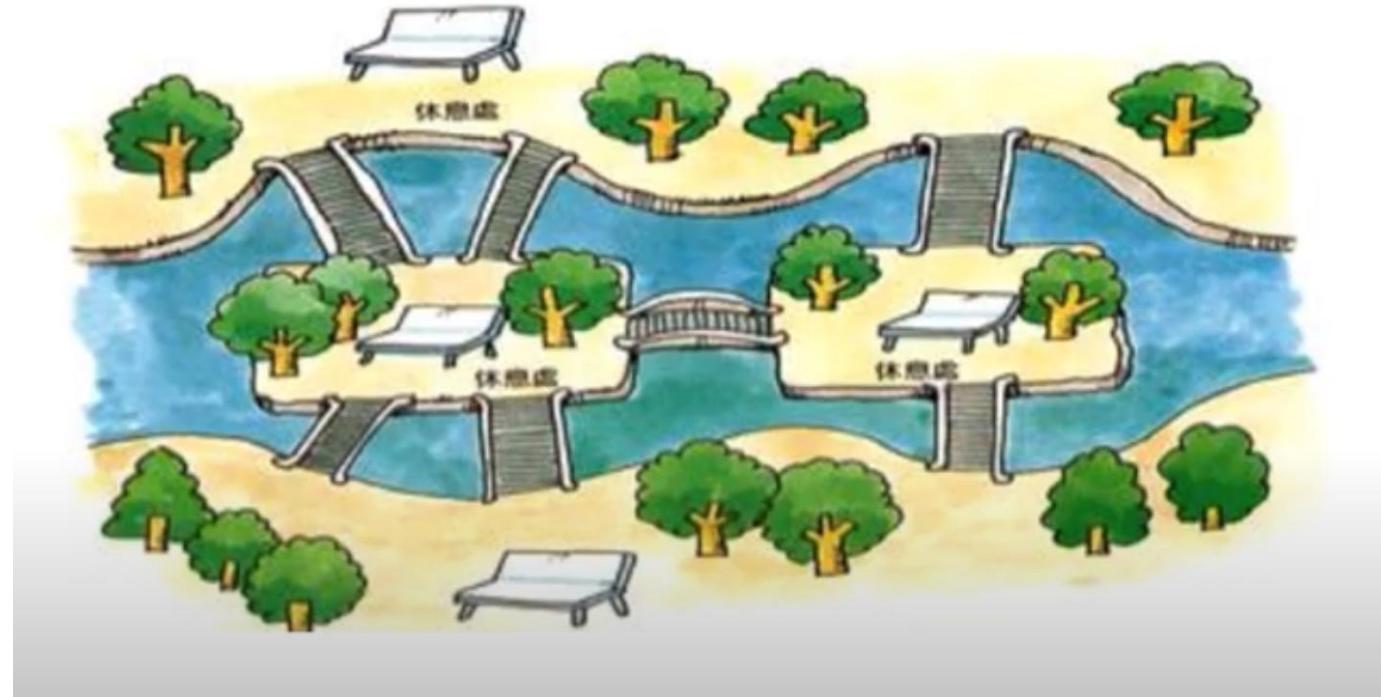
04 影响力度量

05 商务案例-社会网络分析



图论起源：柯尼斯堡七桥问题

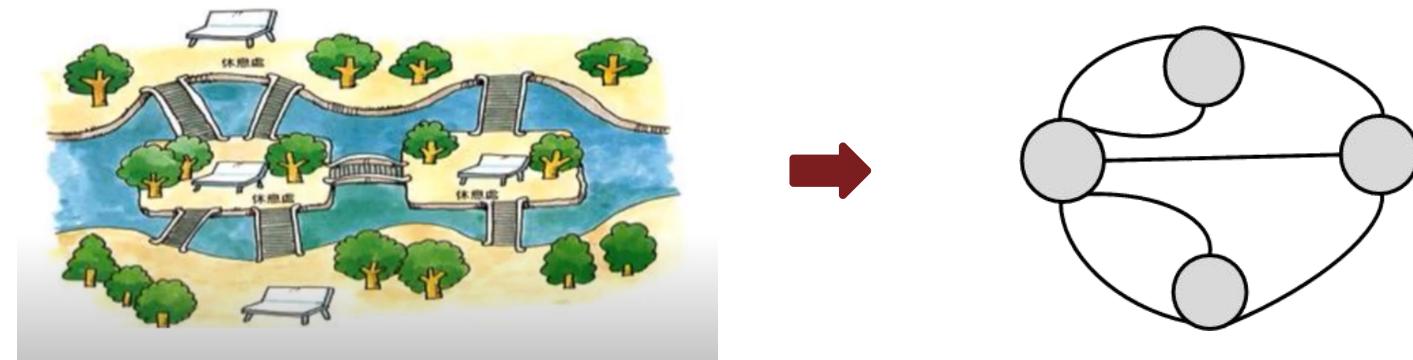
- 河中有2个岛，通过7座桥将2个岛以及岛和河岸连起来
- 一个步行者如何能不重复、不遗漏地一次走完7座桥，最后回到出发点？





图论建模

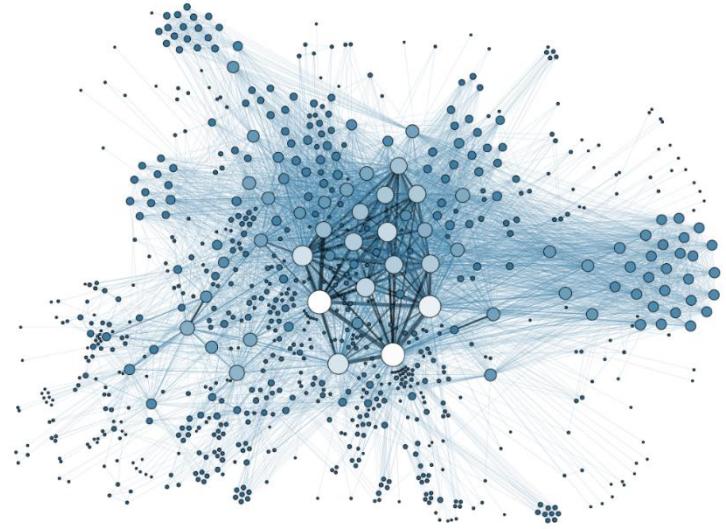
- 图表示：岛或河两岸 → 节点；桥 → 边（一笔画问题）
- “一笔画问题” 充要条件
 - 图形必须是连通的；
 - “奇点（有奇数条边相连的点）” 个数必须是0或2；
- 基于图论建模，欧拉证明不可能完成上述一次性不重复、不遗漏通过7桥的任务。





大规模社交网络分析基础：图

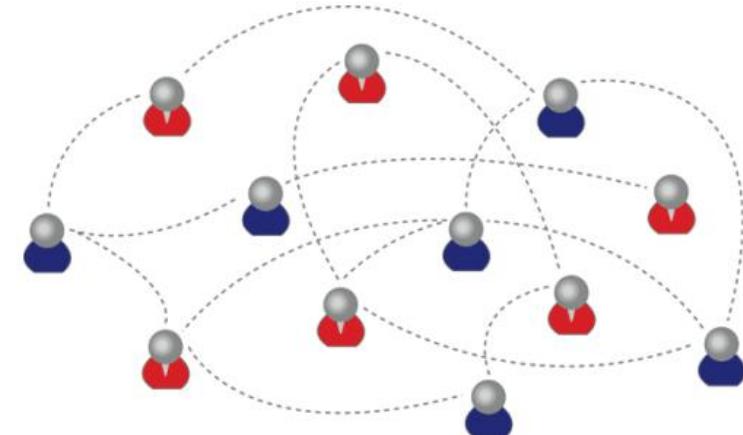
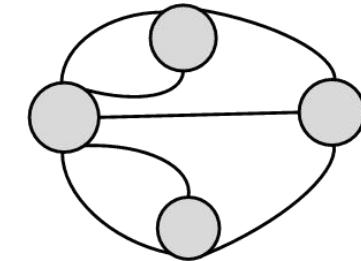
- 图论研究基础扎实、工具繁多，可以简化很多较难的实际问题；
- 社交网络天然就是一个图；
- 图可以转换成计算机可以识别和处理的形式，便于大规模社交网络分析。





社交网络的图表示

- 图：一组被**边**连接的**节点**的集合
- 社交网络的基本核心元素：**用户**（人）和**关系**（包括家人、朋友、同事等）
- 社交网络天然就是一个图
 - 图的**节点**是所有**用户的集合**
 - 节点之间的**连边**表示**用户之间**的**关系**





无向图与有向图

Find us on:
facebook.



WeChat
微信

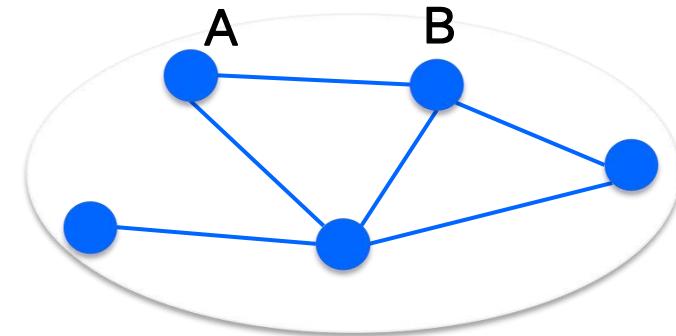


新浪微博
weibo.com

twitter



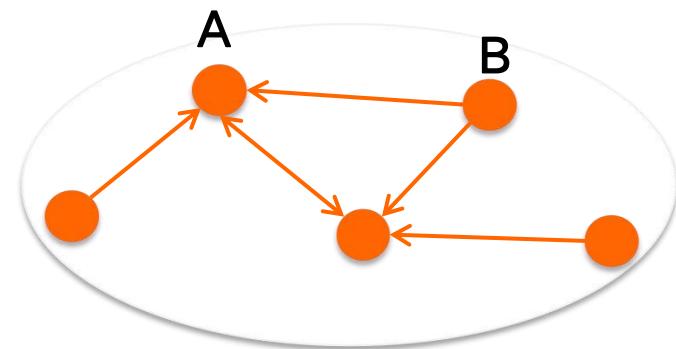
互为好友



无向图



单方面关注

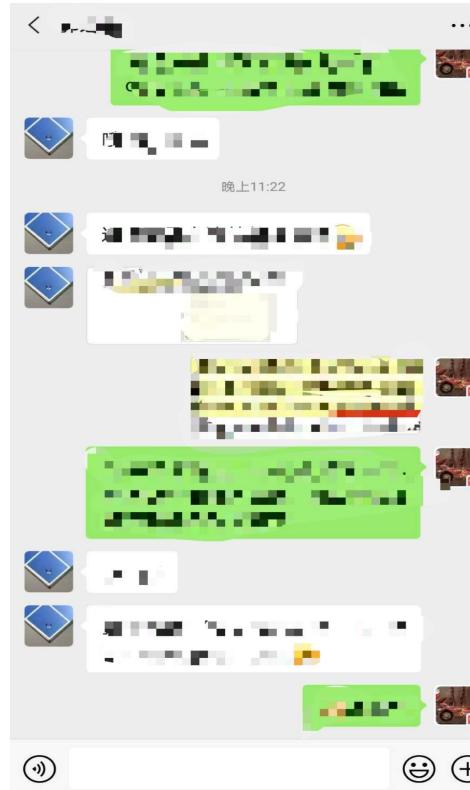


有向图

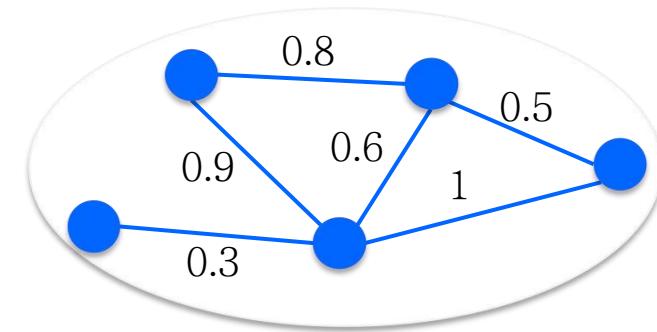
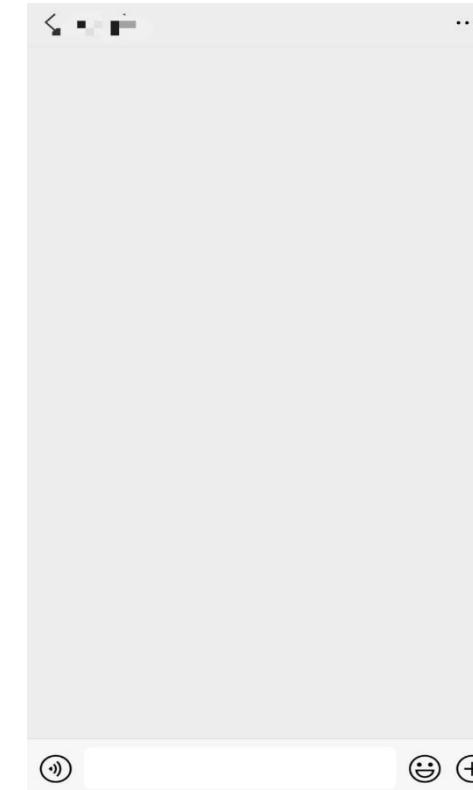


加权图

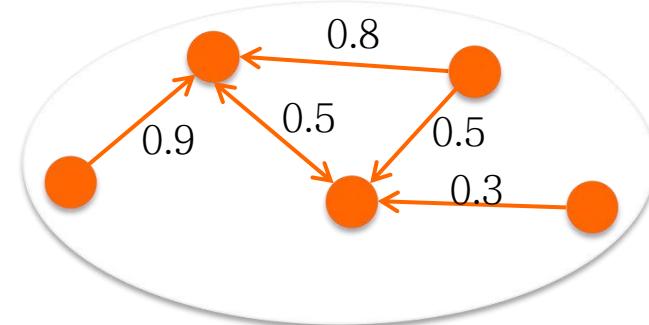
频繁交流



无交流



无向加权图



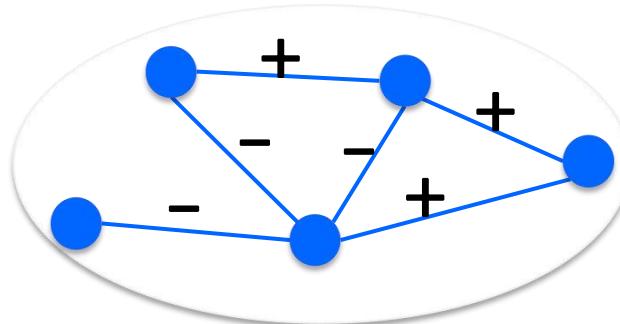
有向加权图

无向图权重度量：交流的频度，共同好友数，兴趣相似度

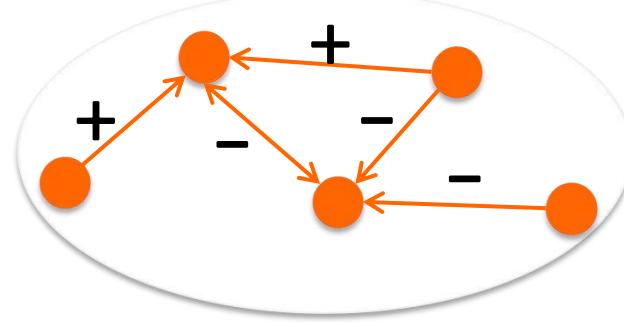
有向图权重度量：点赞或转发的次数



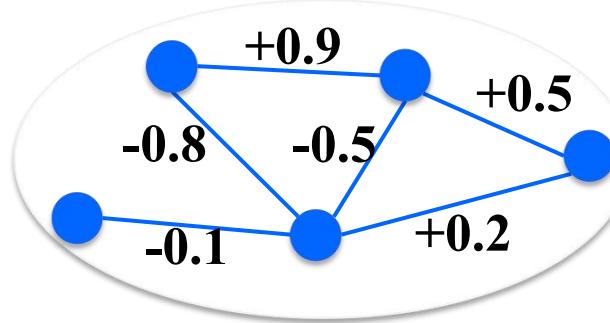
信号图



无向信号图



有向信号图



无向加权信号图



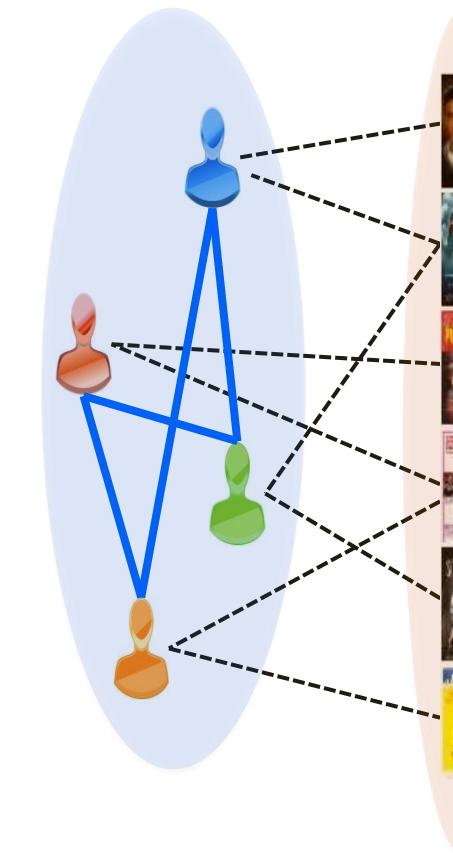
社交网络二分图/异构图表示

豆瓣 douban

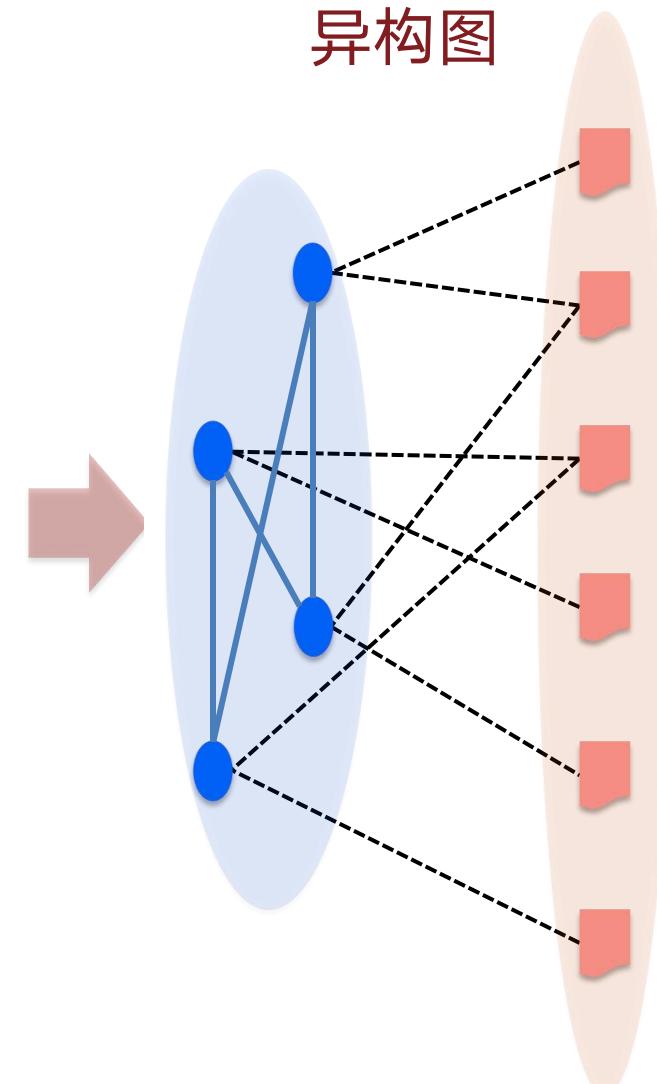


QQ 阅读
一生阅读伙伴

二分图



异构图



二分图：两类节点、一种边

异构图：多类节点、多种边



社交关系的矩阵表示

- 图的表示直观、便于理解和分析，但却不适合高效的数学表达和计算
- 目标：找一种新的表示方法，满足以下三点：
 - 不丢失任何信息；
 - 能被计算机有效识别和处理；
 - 已有的数学方法能方便处理；

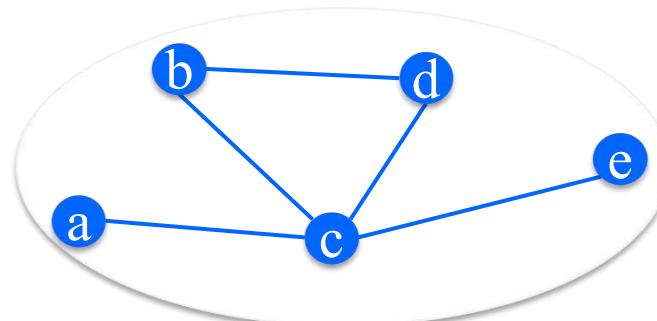
社交网络矩阵表示





社交关系的矩阵表示：邻接矩阵

邻接矩阵：表示图中顶点与顶点之间的边的关系



无向图

	a	b	c	d	e
a			1		
b			1	1	
c	1	1		1	1
d		1	1		
e			1		

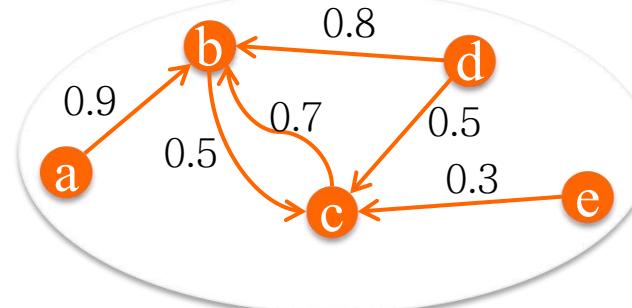


应用：谁的好友最多？

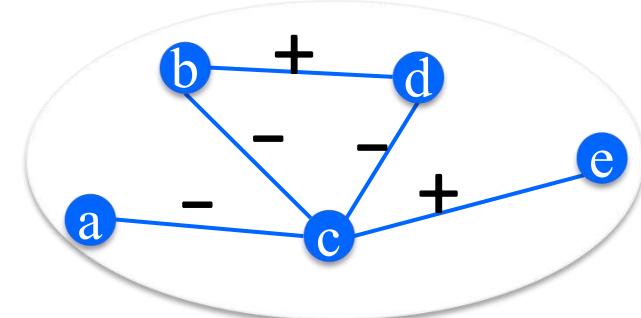
计算：按行（列）相加所得的和表示对应用户的好友数



有向加权信号图的邻接矩阵



有向加权图



无向信号图

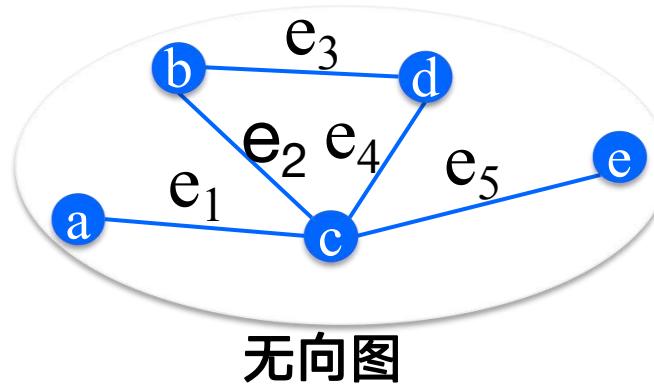
	a	b	c	d	e
a		0.9			
b			0.5		
c		0.7			
d		0.8	0.5		
e			0.3		

	a	b	c	d	e
a			-1		
b			-1	+1	
c	-1	-1		-1	+1
d		+1	-1		
e			+1		



社交关系的矩阵表示：关联矩阵

关联矩阵：表示图中顶点与边之间的关联关系



	e_1	e_2	e_3	e_4	e_5
a	1				
b		1	1		
c	1	1		1	1
d			1	1	
e					1

常用于社交网络边比较稀疏的情况



小结



社交网络表示



无向图、有向图

加权图

信号图

二分图/异构图

邻接矩阵

关联矩阵



讲授提纲

- 01 概述**
- 02 图论基础**
- 03 社群识别**
- 04 影响力度量**
- 05 商务案例-社会网络分析**



社群识别主要内容

定义：什么是组群？

组群识别的意义？

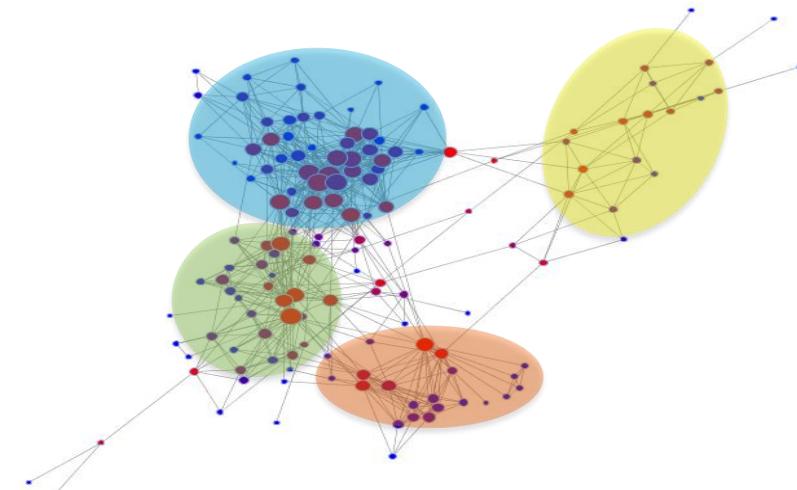
组群识别技术：如何发现组群？





什么是组群？

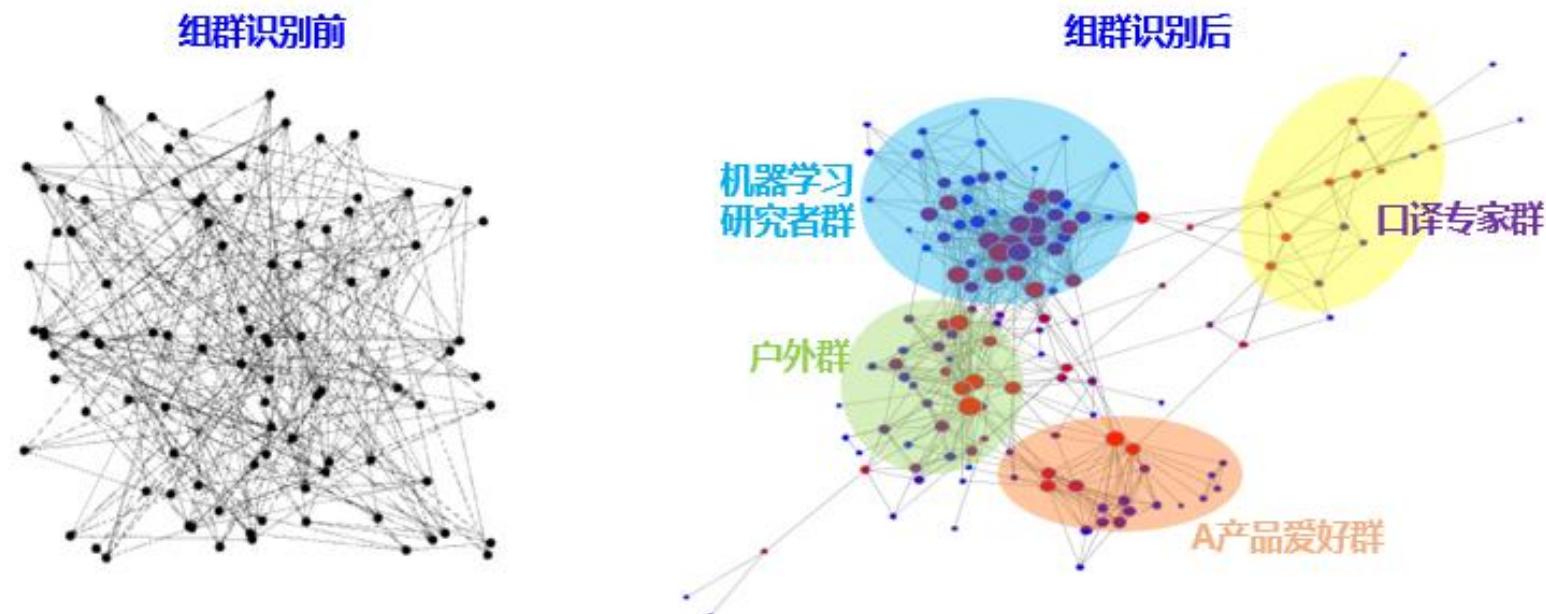
- 应用的角度：信息网络中，**组群**通常对应着功能相同、性质相近或者关系密切的节点集合；
- 组织结构的角度：组群结构是大规模网络中普遍存在的基本结构，即：一个网络是由大量**内部连接“紧密”、外部连接“稀疏”的子团**组成，这些具有统计显著性的子团结构成为一个网络的**组群**结构。





组群识别的意义

- 可以使我们更好的理解网络拓扑结构特点以及节点间的连接关系；揭示复杂系统内在功能特性；
- 为信息搜索、信息推荐、信息传播控制等应用提供支撑；





显性组群和隐性组群

显性存在的组群



大量隐性存在的组群



研究主题群：引用关系

研究合作群：合作关系



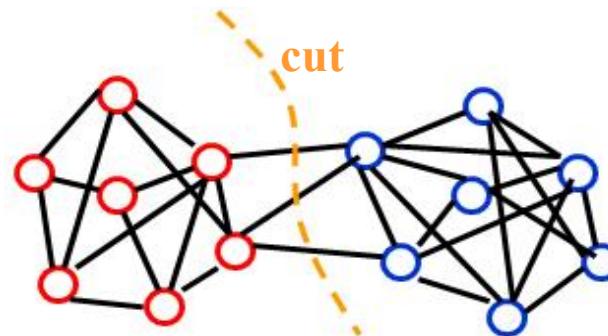


如何发现组群？

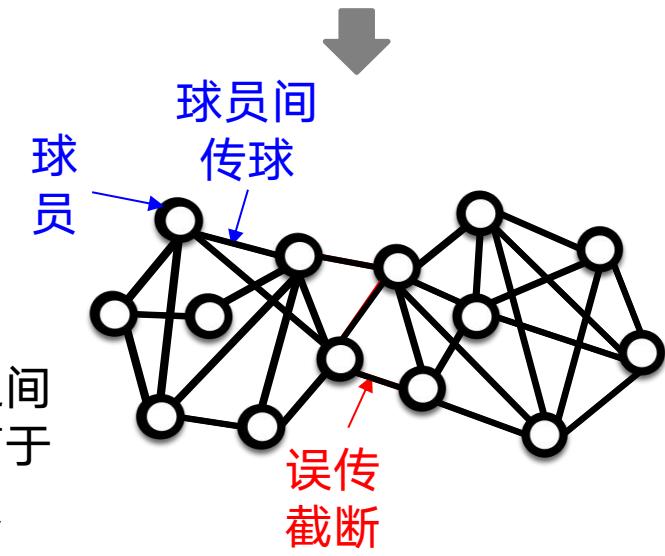
球服区分场上球员的分组



如何区分场上球员的分组？



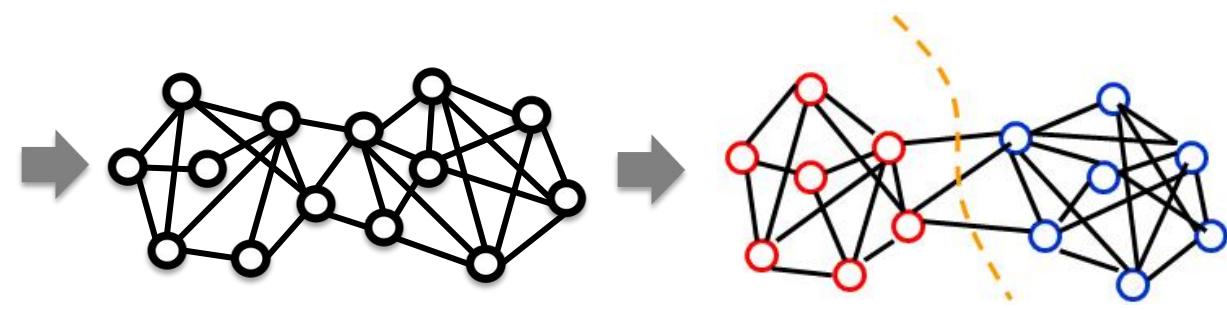
组内球员之间的
传球远高于
组外球员





如何发现组群？

目标转化：给定一组节点，将“组群发现”的任务看作是将一组节点分成两个（或多个）不同的子群



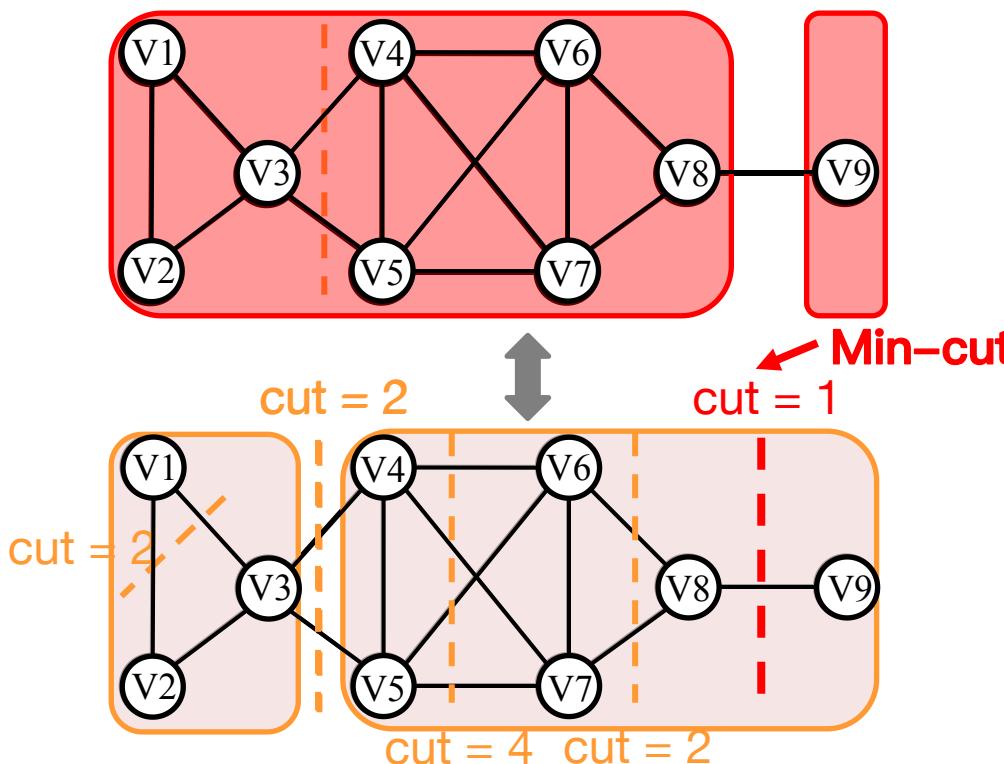
Minimum Cut (Min-Cut) Problem: 给定一个节点图，找一个划分 (cut) 将这个图分成2个子图，使得2个子图之间的连边最少。



Min-cut组群发现

Minimum Cut (Min-Cut) Problem: 给定一个节点图，找一个划分 (cut) 将这个图分成2个子图，使得2个子图之间的连边最少。

如何求解？



存在的问题？

Algorithm:

```
Input : A graph G = (V, E)
Output: Two sets S and T
1  $S = \text{Split}(G)$ 
2  $\text{min\_cut} = \infty$ 
3 for  $s \in S$  do
4   edges =  $\text{edges}(s)$ ,
5   if  $\text{edges} < \text{min\_cut}$  then
6      $\text{min\_cut} = \text{edges}$  ;
7      $s_{\min} = s$  ;
8   end
9 end
10  $\langle g_s, g_t \rangle = \text{Optimal-Split}(S_{\min})$ 
11 return  $\langle g_s, g_t \rangle$ ;
```

Min-Cut

组群成员数目不均衡
经常出现只有一个节点的组群



Ratio/Normalized Cut组群发现

- 目标转化：为了尽量平衡切分后两个组群的成员数量，将其引入目标函数，使得2个子图之间的连边尽量少，并且成员数量的差别尽量小

$$\text{Ratio Cut}(P) = \frac{1}{k} \sum_{i=1}^k \frac{\text{cut}(P_i, \bar{P}_i)}{|P_i|}$$

$$\text{Normalized Cut}(P) = \frac{1}{k} \sum_{i=1}^k \frac{\text{cut}(P_i, \bar{P}_i)}{\text{vol}(P_i)}$$

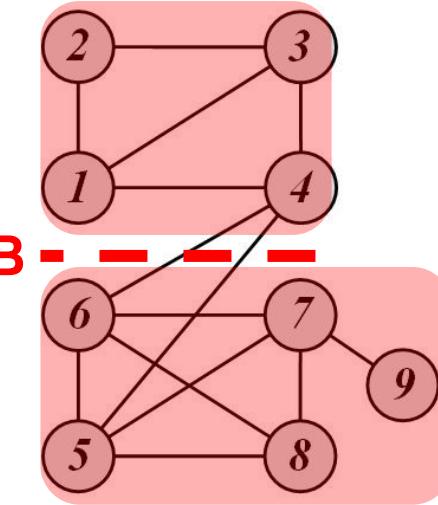
- $V = (P_i, \bar{P}_i)$, V 表示所有节点, \bar{P}_i 和 P_i 是 V 切分后的两个组群
- $\text{cut}(P_i, \bar{P}_i)$ 表示 \bar{P}_i 和 P_i 之间边的数目
- $|P_i|$ 表示子群 P_i 中的节点数
- $\text{vol}(P_i) = \sum_{v \in P_i} d_v$, 即组群 P_i 中所有节点的度之和



Ratio/Normalized Cut算法实现

$$\text{Ratio Cut}(P) = \frac{1}{k} \sum_{i=1}^k \frac{\text{cut}(P_i, \bar{P}_i)}{|P_i|}$$

$$\text{Normalized Cut}(P) = \frac{1}{k} \sum_{i=1}^k \frac{\text{cut}(P_i, \bar{P}_i)}{\text{vol}(P_i)}$$



Algorithm 1: Community Detection with Min-Cut

Input : A Graph G ;

Output: Two subgraphs g_s and \tilde{g}_s .

```

1  $S = \text{Split}(G);$ 
2  $min\_cut = \infty;$ 
3 for  $s \in S$  do
4    $edges = \text{cut}(s);$ 
5   if  $edges < min\_cut$  then
6      $min\_cut = edges;$ 
7      $s_{min} = s;$ 
8   end
9 end
10  $\langle g_s, \tilde{g}_s \rangle = \text{Comm}(G, s_{min});$ 
11 return  $\langle g_s, \tilde{g}_s \rangle;$ 

```

Algorithm 2: Community Detection with Ratio Cut

Input : A Graph G ;

Output: Two subgraphs g_s and \tilde{g}_s .

```

1  $S = \text{Split}(G);$ 
2  $min\_rc = \infty;$ 
3 for  $s \in S$  do
4    $rc = \text{ratio\_cut}(s);$ 
5   if  $rc < min\_rc$  then
6      $min\_rc = rc;$ 
7      $s_{min} = s;$ 
8   end
9 end
10  $\langle g_s, \tilde{g}_s \rangle = \text{Comm}(G, s_{min});$ 
11 return  $\langle g_s, \tilde{g}_s \rangle;$ 

```



小结

■ 组群的**定义**:

- 组内节点连接紧密，组与组之间连接稀疏；
- 组内节点功能、性质相近；

■ 组群识别的**意义**:

- 更好掌握网络的内在结构和功能；
- 更好地服务应用；

■ 组群识别的**方法**:

- Min-cut
- Ratio-cut
- Normalized-cut



讲授提纲

- 01 概述**
- 02 图论基础**
- 03 社群识别**
- 04 影响力度量**
- 05 商务案例-社会网络分析**



社交网络中的大V网红

	顾村新鲜事 本地资讯博主（上海） 8.3万 粉丝	✓
	魔都财观 知名房产博主 63.4万 粉丝	✓
	国匠城 国匠城官方微博 1.7万 粉丝	✓
	老马自奋蹄 用友集团CIO 副总裁 互联网科技... 84.9万 粉丝	✓
	雄大评装修 《装修情报》总编辑 家居博主 微... 1百万 粉丝	✓
	格隆 格隆汇科技发展（深圳）有限公... 8.6万 粉丝	✓





影响力度量主要内容

谁是有影响力的人？

如何度量网络中节点的影响力？





影响力维度

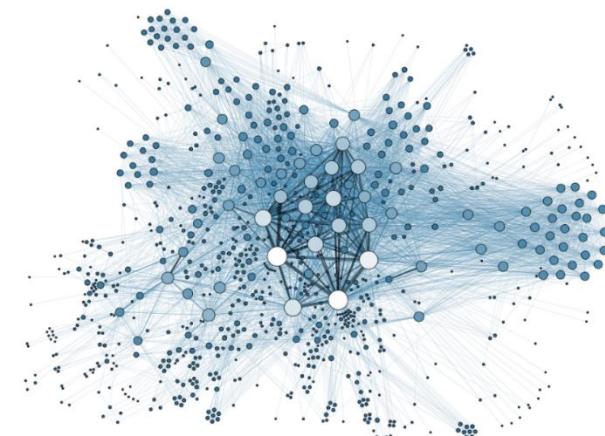
Barack Obama
This account is run by #Obama2012 campaign staff. Tweets from the President are signed -bo.
Influences 2M others
Influential about 20 topics
Government, Politics, Media

Justin Bieber
#BELIEVE is on ITUNES and in STORES WORLDWIDE! - SO MUCH LOVE FOR THE FANS...you are always there for me and I will always be there for you. MUCH LOVE. thanks
Influences 10M others
Influential about 20 topics
Celebrities, Somebody to Love, Justin Bieber

Klout Summary for Warren Buffett
Warren Buffett
Investor, Philanthropist
Omaha, Nebraska
Score Analysis
36 klout score

谁的影响力更大？如何度量

- 范围：影响很多人的行为
- 影响信息扩散的重要性：
没了他信息还能扩散出去么？
- 速度：能否快速将信息传至全网





影响力度量指标

- 度中心性 (Degree Centrality)
- 特征向量中心性 (Eigenvector Centrality)
- 中间中心性 (Betweenness Centrality)
- 接近中心性 (Closeness Centrality)



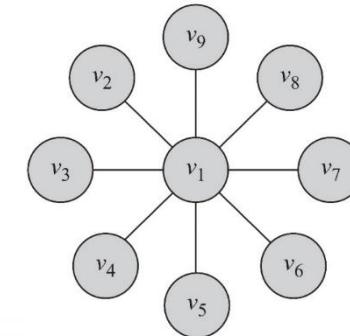
度中心性

- 人脉广、朋友多的人更重要，是网络的中心；
- 一个节点具有越多的连接关系，其**度中心性**越高；
 - 在无向图中，节点 v 的度数中心性即节点的度（邻边的数量）
 - 在有向图中，分别用入度、出度及入度出度的和作为度中心性

$$C_d(v_i) = d_i^{\text{in}} \quad (\text{声望})$$

$$C_d(v_i) = d_i^{\text{out}} \quad (\text{合群性})$$

$$C_d(v_i) = d_i^{\text{in}} + d_i^{\text{out}}$$





度中心性归一化

■ 最大可能度数归一化

$$C_d^{\text{norm}}(v_i) = \frac{d_i}{n-1} \quad n: \text{节点数}$$

■ 最大度数归一化

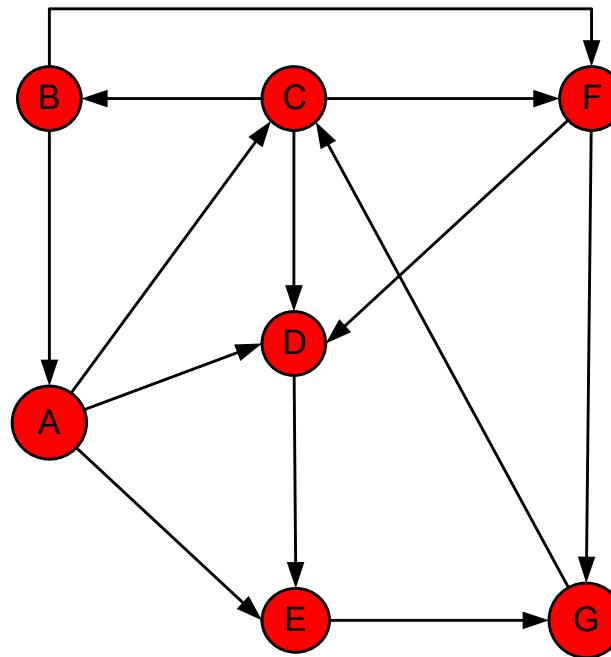
$$C_d^{\max}(v_i) = \frac{d_i}{\max_j d_j}$$

■ 度数和归一化

$$C_d^{\text{sum}}(v_i) = \frac{d_i}{\sum_j d_j} = \frac{d_i}{2|E|} = \frac{d_i}{2m} \quad m: \text{边数}$$



度中心性示例（有向图）



节点	入度	出度	入度中心性	出度中心性	度中心性
A	1	3	1/6	1/2	2/3
B	1	2	1/6	1/3	1/2
C	2	3	1/3	1/2	5/6
D	3	1	1/2	1/6	2/3
E	2	1	1/3	1/6	1/2
F	2	2	1/3	1/3	2/3
G	2	1	1/3	1/6	1/2

最大可能度数归一化：

$$C_d^{\text{norm}}(v_i) = \frac{d_i}{n-1}$$



度中心性的局限

- 拥有很多朋友不一定说明某一个人就一定很重要；
- 拥有相同度中心性的用户不一定具有相同的影响力；



特征向量中心性：拥有很多朋友不一定说明某个人很重要，而拥有更多重要的朋友提供更有力的信号。



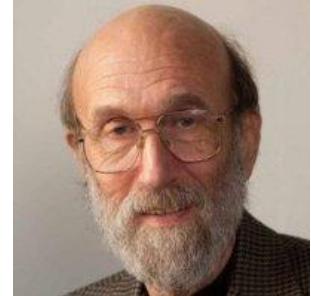
特征向量中心性

- 特征向量中心性：通过结合无向图中的邻居节点（或者有向图中输入/输出邻居节点）的重要性重新定义度中心性。
- 如果有很多重要的邻居指向 v_i ，则特征向量中心性 $c_e(v_i)$ 将较大；
- 假设 $c_e(v_i)$ 表示为所有邻居的中心性之和：

$$c_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^n A_{j,i} c_e(v_j)$$

- A 是图的邻接矩阵，如果 v_i 和 v_j 相连， $A_{j,i} = 1$

- λ 是一个常数



Phillip Bonacich

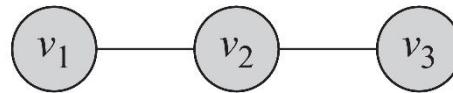


特征向量中心性的计算

- 假设 $\mathbf{C}_e = (C_e(v_1), C_e(v_2), \dots, C_e(v_n))^T$
- $c_e(v_i) = \frac{1}{\lambda} \sum_{j=1}^n A_{j,i} c_e(v_j) \rightarrow \lambda \mathbf{C}_e = A^T \mathbf{C}_e$
- \mathbf{C}_e 是邻接矩阵的特征向量， λ 是对应的特征值
- 一个矩阵可能有多个特征值，对应多个特征向量。我们应该选取哪对特征值-特征向量？
 - 所有节点的中心性应该为正数
 - 佩龙-弗罗宾尼斯定理：选择最大的特征值



特征向量中心性（计算示例）



$$\lambda \mathbf{C}_e = A \mathbf{C}_e \quad (A - \lambda I) \mathbf{C}_e = 0 \quad \mathbf{C}_e = [u_1 \ u_2 \ u_3]^T$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 - \lambda & 1 & 0 \\ 1 & 0 - \lambda & 1 \\ 0 & 1 & 0 - \lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\det(A - \lambda I) = \begin{vmatrix} 0 - \lambda & 1 & 0 \\ 1 & 0 - \lambda & 1 \\ 0 & 1 & 0 - \lambda \end{vmatrix} = 0$$

$$(-\lambda)(\lambda^2 - 1) - 1(-\lambda) = 2\lambda - \lambda^3 = \lambda(2 - \lambda^2) = 0 \quad \text{特征值的解为: } (-\sqrt{2}, 0, +\sqrt{2})$$

最大特征值

根据特征根，假设的 \mathbf{C}_e 范数为 1

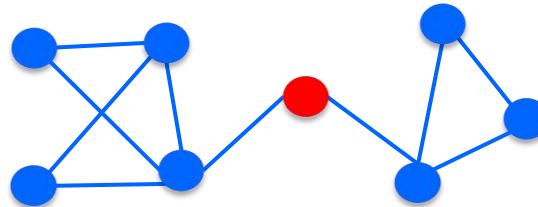
$$\begin{bmatrix} 0 - \sqrt{2} & 1 & 0 \\ 1 & 0 - \sqrt{2} & 1 \\ 0 & 1 & 0 - \sqrt{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{C}_e = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 1/2 \\ \sqrt{2}/2 \\ 1/2 \end{bmatrix}$$



中间 (Betweenness) 中心性



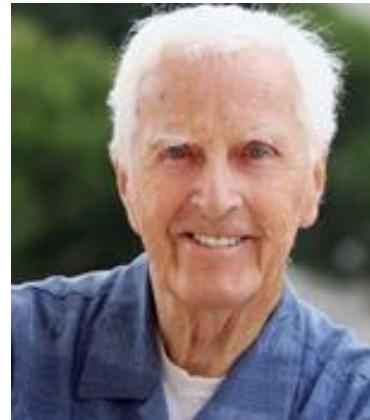
Linton Freeman



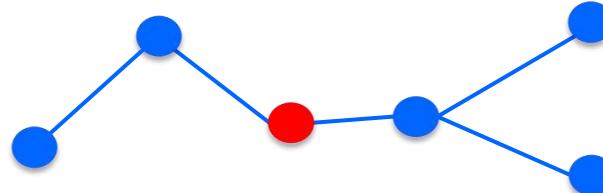
- 中间（介数）中心性用来评价节点的流量承载能力。
- 某个节点的介数越大，说明它在信息传播过程中通过该节点的信息量就越多。
- 在社交网络中介数常用于评价某一个用户对于网络中所有用户对之间传递消息的重要程度。



接近 (Closeness) 中心性



Linton Freeman



- 接近中心性计算节点与其它节点之间的平均最短路径；
- 平均最短路径越小，节点的中心性越高。
- 接近中心性的思想是：节点越趋于网络的中心，它们越能更快速地到达其他节点。

小结

■ 影响范围:

- 度数中心性
- 特征向量中心性

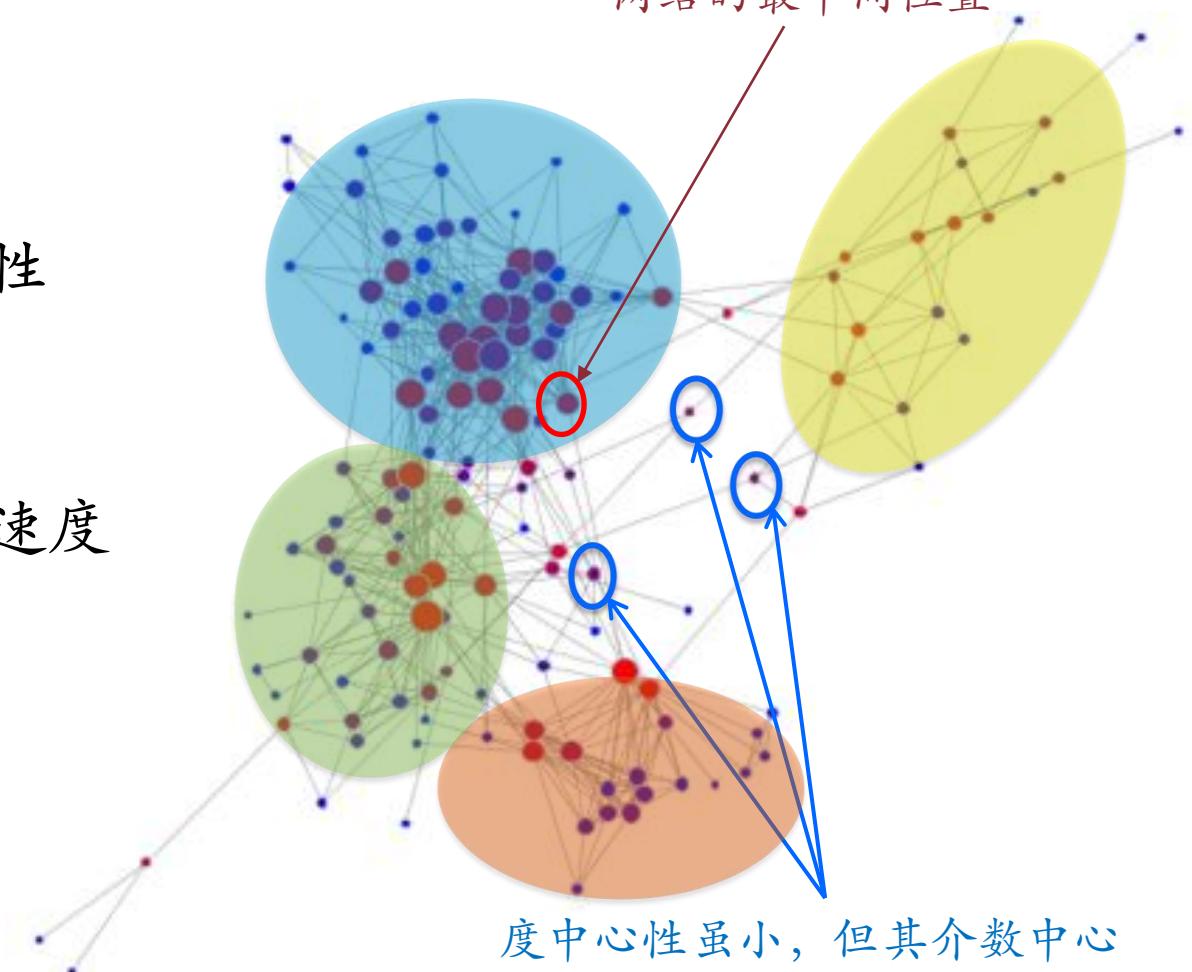
■ 信息流传递重要性

- 中介中心性

■ 信息传播全网的速度

- 亲近中心性

亲近中心性高，在网络的最中间位置



度中心性虽小，但其介数中心性大，控制网络的信息流；



讲授提纲

- 01 概述**
- 02 图论基础**
- 03 社群识别**
- 04 影响力度量**
- 05 商务案例-社会网络分析**