# 1. Introduction

## 1.1 复杂网络无处不在

互联网的时间变化：

1969 ARPAnet(Advanced Research Projects Agency Network美国高等研究计划署网络)

1986 NSFNET(The National Science Foundation Network 国家科学基金会网络)

1995 INTERNET 互联网

网络分层：IP层，路由层，域名层

【移动通信网络】

人类通讯动态的呼叫模式 (幂律分布 [基于机器人的呼叫者，电信欺诈，电话销售])

(韦博分布 [普通用户])

网络科学：1.构建网络 2.研究细节 3.研究动态 4.关联动态与突发事件

【全球航空公司网络】

【交通网络】(公路，铁路，水路)

从最近的城市开始地面旅游超过48小时的不超过10%的土地。

！！！网络让世界更加互联：互联网，移动通信网，交通网

【悖论】交通博弈

【金融网络】

搜索引擎，级联故障，电网，金融危机

复杂网络在金融上的应用

主要课题：复杂金融网络对经济稳定性的挑战

网络科学家，经济学家和监管机构首次合作讨论

！！！太中心不行，太大不行。关注节点位置而不是大小。

【生物网络】

网络生物学，网络医学，生态网络

人的大脑有100~1000亿个神经元

【专利网络】

【软件系统网络】为大型软件系统提供探索性分析。设计，效率，可重用性，脆弱性，可控性

【语言网络】

【科学合作网络】

【恐怖网络】

【篮球队策略网络】球员是节点，球的轨迹是边

【组织结构】

【如何评估学生的表现】早期发现问题能及时纠正，可以减少辍学人数。 连续的方式

【成功研讨会】社会科学家，计算机科学家，经济学家，物理学家，数学家，讨论成功的规律。过去几年个人效绩的大量数据促使科学家探索个人成功的路径模式。

【社会动员关键时间】 复杂网络的时间演变

【网络的网络】网络相互依赖

【网络科学】什么是连接，网络的出现和演化有共同的基本规律和可再现的机制驱动的

## 小世界网络

世界上任意两个人之间的平均距离是多少？

【1967小世界实验】信件链 20%到达，距离6.5

【六度分离】美国任意2个人的距离是6

2002实验:4.0 2006msn:6.6 2009renren:5.38 2011facebook:4.74 贝肯数 埃尔德什数

## 网络科学的崛起

网络地图的出现，网络特性的普遍性

【网络传播】通讯，运输，电网，社会，生物，经济

【社会网络传播】病毒，时尚，行为，谣言，意见，信仰

【理论科学】图论，博弈论，统计物理学，计算机科学

【应用科学】通讯科学，电力工程，生命科学，社会科学

【网络建模】测试数据，分析行为，发现财产，设计性能

【目标】收集足够信息以便：正确的描述，正确的量化，数学的公式，合理的预测，强力的控制

【游戏改变者】对于每种技术，前十年发展，第二个十年市场跟进

【挑战】将复杂问题分解为小问题

【特点】跨学科，数据驱动，定量，计算性

【典型网络科学研究】发现，建模，验证

【数学风格】

【物理风格】

【控制风格】复杂网络可控性，复杂网络可观测性

【计算机科学风格】社区检测风格，链路预测，推荐算法

【实验风格】6100万人的社会影响和政治动员实验

【社会影响偏差】随机实验

【大数据时代】数据资源，数据产业，数据战略，数据科学

【数据量增长快？】1.数据产生方式改变，2.人类生活依赖数据

【大数据】海量数据或巨量数据，其规模巨大到无法通过目前主流的计算机系统在合理时间内获取、存储、管理、处理并提炼以帮助使用者的策略

【特点】Value(价值密度低)，Velocity(快速)，Volume(数据量大)，Vari ety(多样)，

Complexity(复杂度)

# 网络和图

图=事物+联系=节点(vertex,node,point)+边(连接,链接,关系,联系;edge,link,tie)

【起源】哥尼斯堡七桥问题：每架桥只经过一次，走过所有桥。

【图论之父】Leonhard Eular(莱昂哈德 欧拉) ：陆地为节点，桥为边

【证明】奇数个边的节点必须为起点或终点，七桥问题有4个奇数边的节点，所以不能遍历。

【贡献】将问题用图表示，能变得更简单易于处理。

图和网络的结构是理解复杂世界的关键

隐藏在图和网络中的属性，会限制或增强我们处理问题的能力

图论是网络的基础

【网络(图)表示现实世界】不同性质的实际系统可以具有相同的网络表示

例如：互联网，好莱坞演员网络，蛋白质作用网络

【区别】网络：通常指实际系统，术语：网络(network)，节点(node)，连接(link)

图论：网络的数学表示，术语：图(graph)，顶点(vertex)，边(edge)

【数学表示】G(V,E),V={a,b,...},E包含于{(x,y)|x,y属于V,x不等于y}

【异构】

【图的个数】=2^C(n,2) Polya定理：?

【分类】(有标号图，无标号图)(有向图，无向图)(权重图，无权重图)

【简单图】无多边，无自边

【完全图】边 = n(n-1)/2

【计算机表示】邻接矩阵，关联矩阵

【路径】path,通路：节点序列(相邻节点存在边) 长度：涉及的边条数，最长为节点数-1

【距离】两个节点之间的最短路径长度

【连通图】任何两个节点之间存在路径

【连通分量】连通子图，不包含在其他连通子图中，孤立的子图

关注路径长度而不是物理距离

【最大连通分量】

【最大分量】包含图的显著部分

【桥】特别性质的边，删除之后两个端点就不再有路(删除会增加连通分量)

【捷径】删除之后两个节点之间的距离至少为3(桥是捷径的特例)

【有向路径】节点序列，相邻节点之间有从前往后的有向边

【强连通分量】任意两个节点之间存在有向路径，不真包含在其他强连通分量中

【二部图】没有长度为奇数的圈：

从任意节点出发，在广度优先搜索过程中，同一层节点有边，则存在长度为奇数的圈

【遍历】广度优先搜索

【应用】合作图，交流图，信息链接图

【三元闭包】如果两个互不认识的人有一个共同的朋友，则将来成为朋友的可能性提高

【聚集系数】刻画网络结构中的节点属性

节点A的聚集系数

=与A相邻任意两个朋友之间也是朋友的概率

=与A相邻的朋友对的个数/总的对数

【另一种描述】两个互不认识的人的共同朋友数越多，则在未来成为朋友的可能性越大

【验证】电子邮件网络：100对节点，某时刻没边，有5个共同好友，一个月里，20对节点发生通信，80没有，则在一个月里成为朋友的概率为0.2

【共同朋友个数】 = 邻接矩阵相乘

【强关系、弱关系】

【嵌入性】边的属性，一条边两端共同的邻里数

【结构洞】一个节点，移除该节点会使网络变成多个连通分量的节点

【图划分算法】

分割法：逐步去掉“跨接边”

聚集法：“滚雪球”

近似：准确与效率的平衡

【Girvan-Newman方法】

【介数】一条边承载的一种“流量” 流量均分到所有最短路径上，

考虑所有对后，累积流量就是介数

【介数计算】从一个节点开始，做宽度优先搜索，确定最短路径条数，确定向其他节点发送流量时，经过每条边的流量。重复以上过程，累计除以2，即得每条边的介数

【总结】边的属性用强弱性、嵌入性表示，点的属性可以用聚集系数、结构洞表示

【强三元闭包】如果一个节点有两个强关系的邻居，则这两个邻居不能什么关系都没有

【捷径】A符合强三元闭包，则与A相连的捷径为弱关系。

【验证】共同朋友越多，关系强度越高

【社会归属网】描述从属关系与社会关系

【社团闭包】选择。由于参与同一件事，原本没联系的人建立联系

【会员闭包】影响。由于朋友参与这件事，另一个人也加入这件事

【同质性】两种特征。节点数n,边数e,不同节点占比p,1-p,两端节点相同的边s

【自动化】n:人物节点个数 m:社交聚点个数 A(i,j):i与j的连接关系，B(i,j):i是否加入j

A(i,\*):A的第i行向量 B(\*,j):B的第j列向量

三元闭包：若A(i,j)=0 且A(i,\*)\*A(j,\*)>=3,则A(i,j)将=1

社团闭包：若A(i,j)=0 且B(i,\*)\*B(j,\*)>=2,则A(i,j)将=1

会员闭包：若B(i,j)=0 且A(i,\*)\*B(\*,j)>=2,则B(i,j)将=1

【谢林模型】起源：芝加哥黑人居住区的改变

同质性动态：越来越多的黑人在某个区域聚集

模型:一个空间居住两类自然特征人群，邻居指紧邻的居住单元（8个）。

t<3就搬家，随机执行