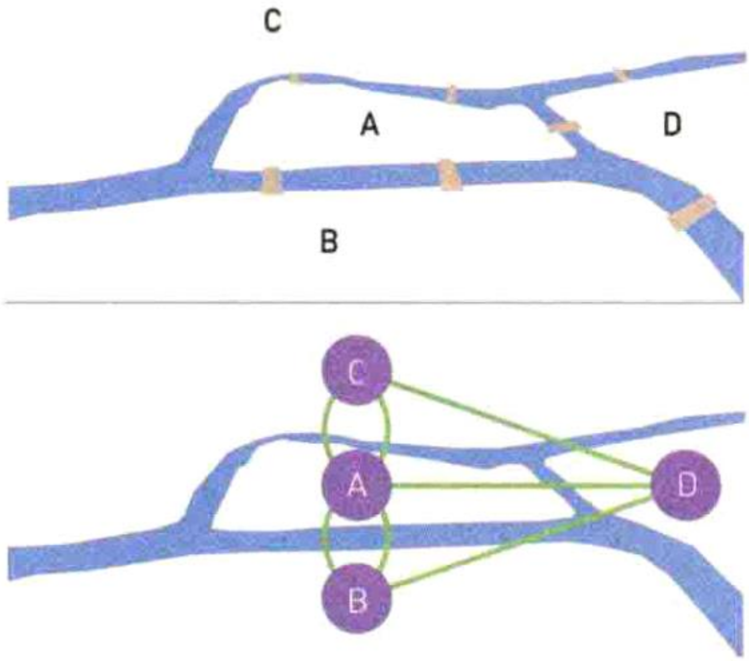
第一节 哥尼斯堡得到七桥问题

2022年12月12日　　8:53

1753年 著名的哥尼斯堡七桥问题被**欧拉**做出了严格的证明

七桥问题：有4块区域通过7座桥相连，那么能否存在一个路径使得一个人不重复的走过每一座桥。



欧拉使用字母来表示每一块区域，线来表示连接每一不同区域的桥梁，于是将上图的现实问题抽象成了字母、线段的图像问题。这极大地简化了我们求解的过程，屏蔽掉了一些复杂的干扰信息。

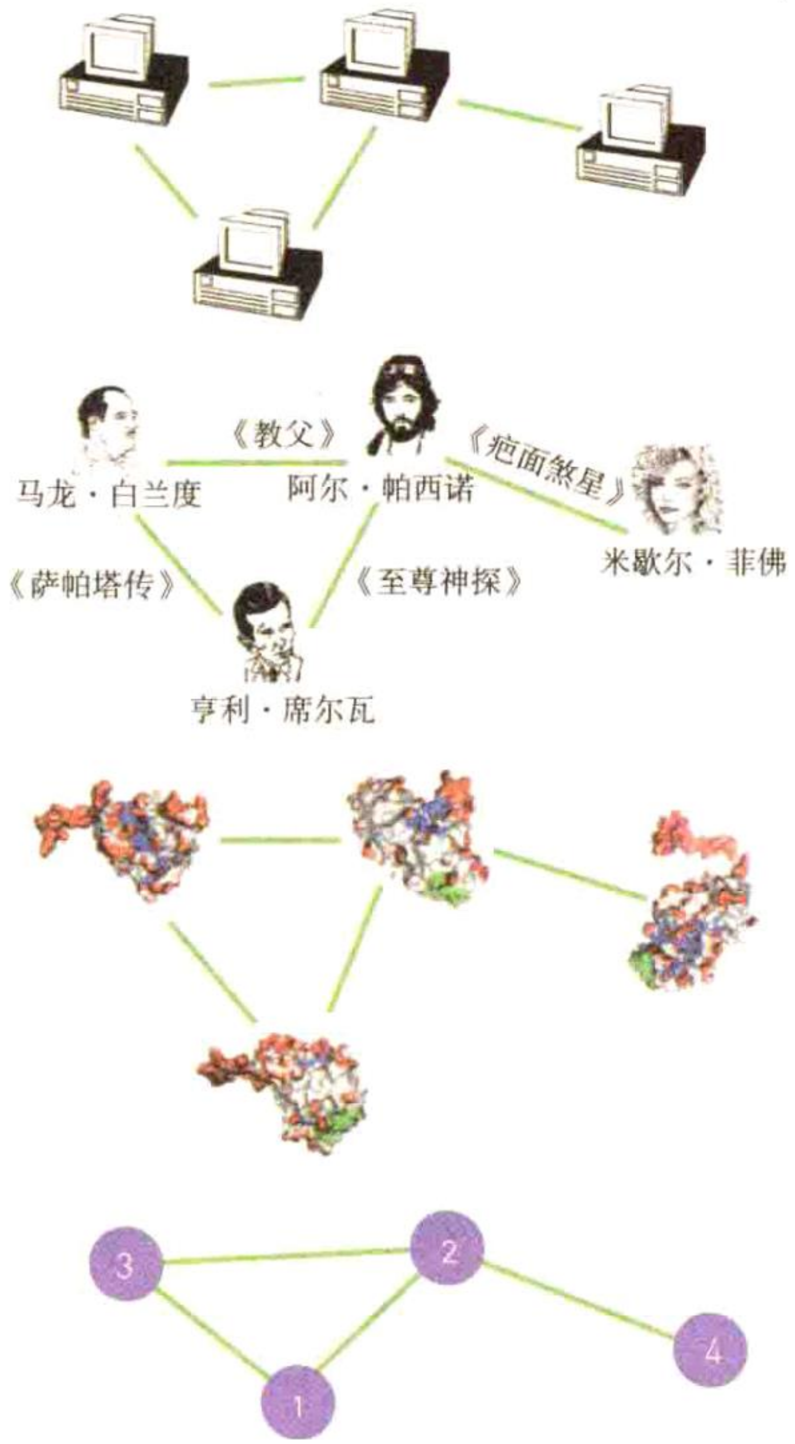
欧拉对于七桥问题给出了一个总结，如果一幅图画中*奇数连接点数超过2个*，那么这幅图就不存在一个路径能够不重复的走完每一个点，这也被我们称为**“一笔画”**问题。

同时，欧拉使用“图”的思想解决现实问题也给我们了一些启示：**图是一种能够简化现实问题，并且具备一定性质和内涵的一类工具**。这也是“图论”的起源。

第二节 网络和图

2022年12月12日　　8:53

要理解复杂的系统，我们就需要去理解复杂系统之间的交互方式。网络作为复杂系统的一种表示方式，为我们研究复杂系统提供了公共语言。



那么我们首先从网络的基本概念入手：

1.**节点i**，表示网络中实体的数量，一般为数字、字母与圆圈的组合表示，例如最后一个图中节点1，节点2，节点3，节点4，当然也可以直接称呼节点名称。

2.**节点数N**，指的是一个网络中存在的节点数量，一般用来表示网络的大小。例如最后一个图，N=4。

3.**链接l**，表示网络实体之间的交互关系，一般用线将两个或多个实体相连来表示。
注：链接可以是有向的，也可是无向的。若一个网络中所有的连接都是有向的，则我们称其为有向网络或者有向图，同理，若都是无向链接，则称为无向图或无向网络。

4.**链接数L**：表示节点间交互关系的数量，例如最后一个图中L=4，也就是网络中线段的数量。

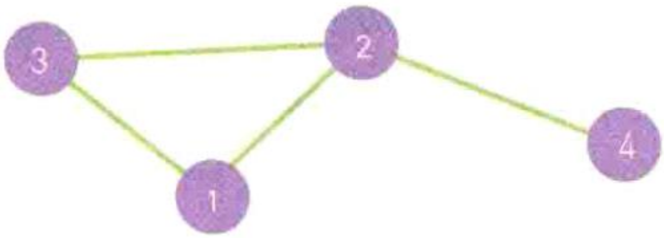
网络与图的细微差异：一般来说，我们对于一个现实社会中真实存在的网络结构称网络，而将数学中的抽象网络结构称为图，但是本质上并没有做细致的区分。图与网络在论文中也经常会当作同义词使用。

第三节 度、平均度和度分布

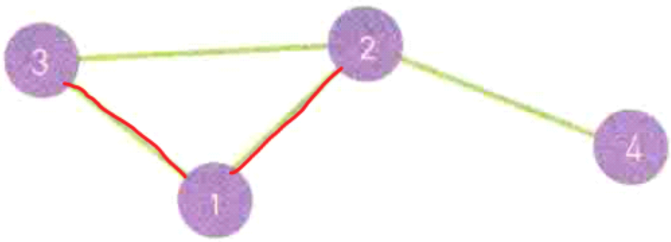
2022年12月12日　　9:32

度是节点的关键属性，表示该节点和其他节点之间的链接数。

我们使用k_i网络中第i个节点的度。



我们可以看到的是$k_1 = 2$，因为节点1和其他节点有两个链接。
因此我们可以同样写出$k_2 = 3$,$k_3 = 2$,$k_4 = 1$



需要注意的是**节点的度和网络的链接数之间存在着关系**：我们可以看到在上图这样的无向图中，每一个节点相当于占有了一半的链接，例如：节点1和节点2,共用了一条链接，在没有权重比例的情况下，节点1和节点2平分这一条链接，因此我们可以总结，*在无权无向图中，节点度之和的一半就是链接数L。*

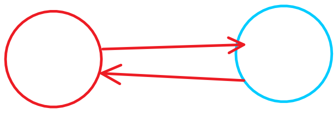
$$L = \frac{1}{2}\sum_{i=1}^{N} k_i$$

平均度也是网络中的一个重要属性，在无向无权图中：

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i = \frac{2L}{N}$$

这是用来衡量网络中的每个节点平均占有多少的度。

有向网络中，度被分成出度和入度。在下图中我们能看到，红色节点的出度为1，入度也为1，蓝色节点出度入度也为1。



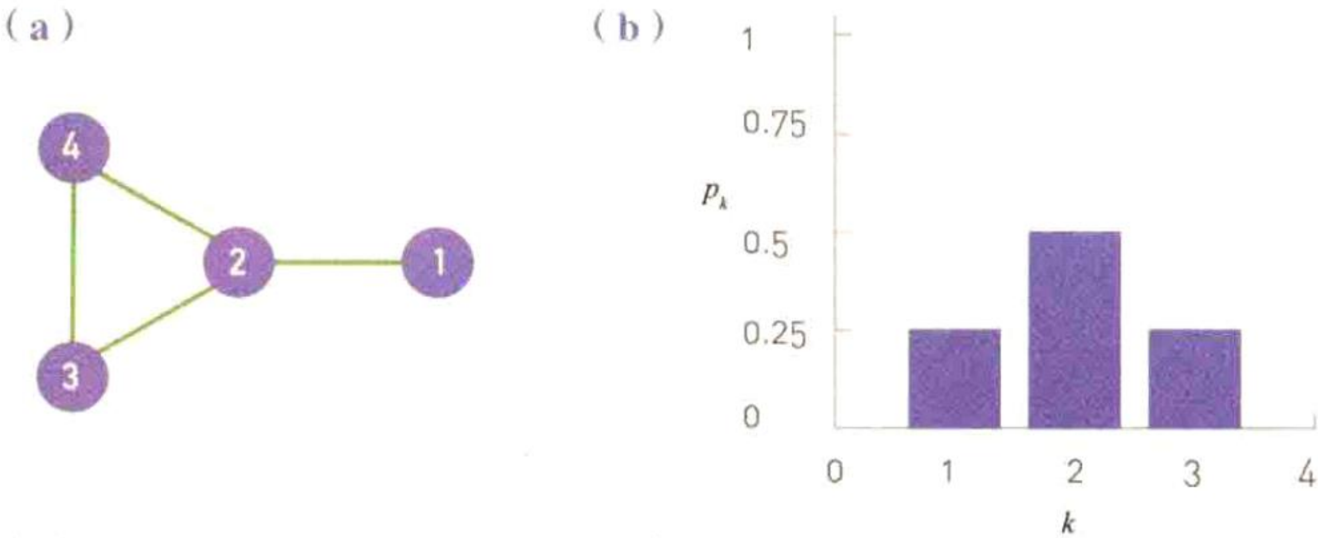
在这里，有向无权图的链接数与度数的关系为：

$$L = \sum_{i=1}^{N} k_i^{\text{in}} = \sum_{i=1}^{N} k_i^{\text{out}}$$

同样，平均度的表示为：

$$\langle k^{\text{in}} \rangle = \frac{1}{N}\sum_{i=1}^{N} k^{\text{in}} = \langle k^{\text{out}} \rangle = \frac{1}{N}\sum_{i=1}^{N} k^{\text{out}} = \frac{L}{N}$$

度分布用来表示网络中随机选择一个节点，其度为k的概率，用p_k表示：
我们可以看这样一个图例：

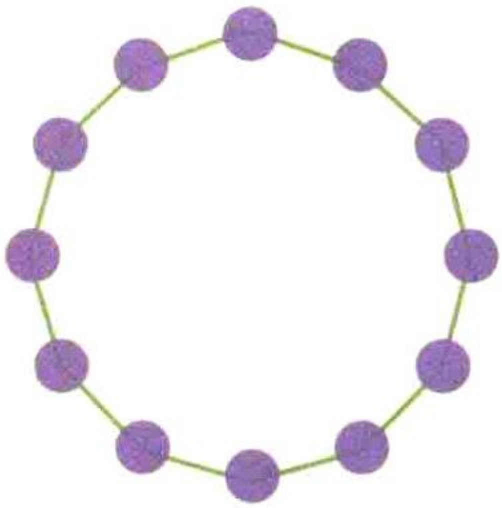


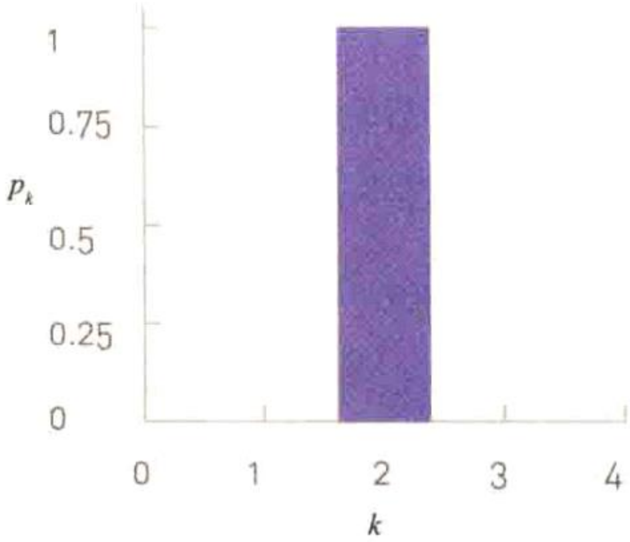
度分布的计算方式为 度为k的节点数/总的节点数：

$$p_k = \frac{N_k}{N}$$

度分布在网络中具有至关重要的作用，这主要是因为，**度分布可以找到不同网络存在的相似性，并且多数相同度分布的网络往往具备相同的结构或者相同的性质**，这是我们对网络研究想要了解的最重要内容。

而对于一些特殊的度分布网络，我们往往有特殊的名称来描述，比如：

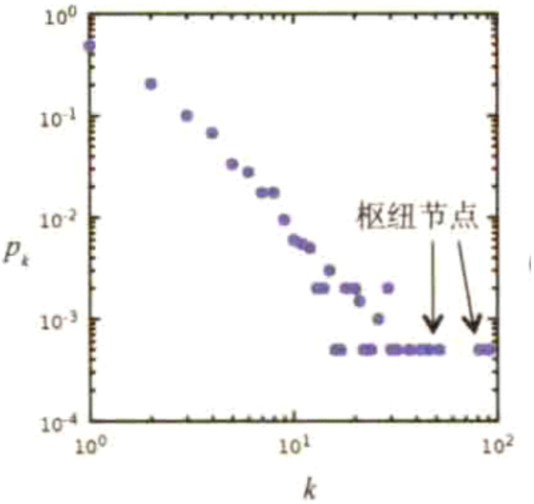




这样的网络，我们称为：***克罗内克的德尔塔函数度分布***

$$p_k = \delta(k-2)$$

同时，在真实网络中，网络中各个节点的度往往差异很大，有的可能就是单独的孤立节点，度为0，而有的则可能是枢纽节点，度非常大，而这个时候，我们**往往要通过度分布的方式去将各个节点分类画在直方图中**，从而能够让我们清楚看到整个网络大概有几个枢纽，几个孤立节点，从而能够更好的解决网络复杂性的问题。（如下图）



当然，**我们除了直接使用k和pk来衡量，我们还可以使用lnk和lnpk的双对数形式来衡量**。

第四节 邻接矩阵

2022年12月12日 10:16

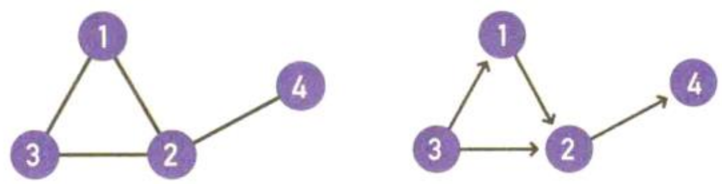
邻接矩阵可以作为网络的一种描述方式，一个有N个节点的网络，其应该存在一个N*N的邻接方阵

在邻接矩阵中，若节点i和节点j之间有链接，则在Aij位置写1,否则为0
我们的邻接矩阵的状态是这样的：

$$A_{ij} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

举一个具体的例子：

（b）无向网络 （c）有向网络



$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad A_{ij} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

无向图的邻接矩阵是对称的，但是有向图就不一定。

第五节 真实网络的稀疏性

2022年12月26日　　8:52

现实生活中的网络往往是稀疏的，例如我们举一个非方阵的例子，以用户和淘宝商品之间的关系来举例

如果你购买过某商品，那么在这一矩阵位置的值为1，否则为0，假设用户有800万，淘宝中的商品有100万中，
仔细思考，以个人的购买能力即使再强，也不会把这100万件商品都购买一遍，可能每个人也就浏览或者看过1000件商品，仅仅占1/1000的位置会被标记为1，其他都是0，那么这个关系或者邻接矩阵将是一个极大比例为0的矩阵，这是一个很稀疏的矩阵。

网络是巨大的，但是对于我们个人来说，我们远远用不完整个网络中的资源。

第六节 加权网络

2022年12月26日 8:58

现在我们主要讨论的都是没有加权的网络，但是现实生活中，很多网络是有权重的，比如社交关系网，你与同学A的关系和你与同学B的关系，以及你与父母的关系的密切程度或者重要程度是不一样的，所以，我们可以用不同的权重来表示不同的重要性。

对于加权网络而言，邻接矩阵的元素表示链接的权重，即

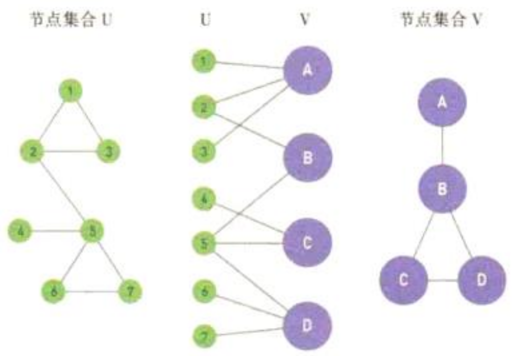
$$A_{ij} = w_{ij} \tag{1.13}$$

梅特卡夫定律：网络的价值与网络的节点个数的平方成正比（N^2）
但是在现实生活中，网络的实际价值往往不是平方级的变化，因为正如之前说的，网络是稀疏的，那么很多的资源是没有办法利用的。**一般情况下，网络的价值增加是随节点数N的线性变化。**
同时我们也要考虑网络权重对于网络价值的影响。

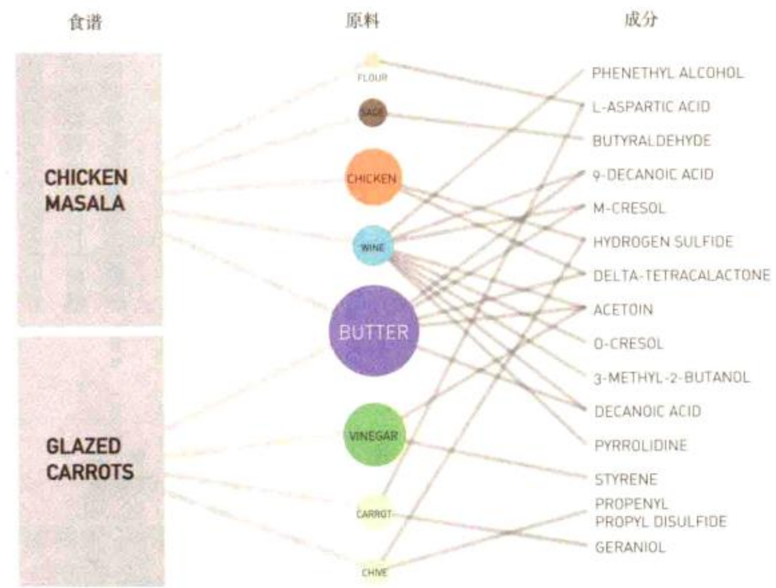
第七节 二分网络

2022年12月26日 9:08

二分网络：其节点可以分为两个节点集U和V。使得每一个链接中都有一个U中的节点和一个V中的节点。



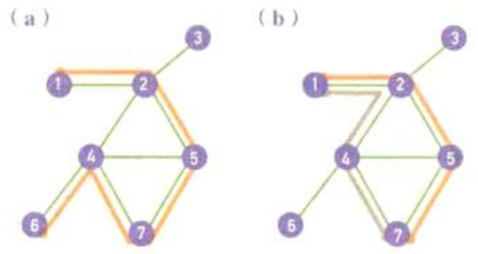
除了二分网络外，还可能存在三分网络或者多分网络。定义可以根据二分网络引申。



第八节 路径与距离

2022年12月26日　　9:12

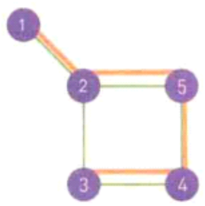
路径：表示为网络中链接行走的路线
路径的长度：路线中包含的链接数量



例如（a）中，橙色的路径为1->2->5->7->4->6，路径长度是5，因为包含了5段链接。

最短路径：从节点i到达节点j的包含连接数最少的路径
直径：最短路径的最大长度，也就是所有最短路径中最长的那一个路径的长度
欧拉路径：每条链接恰好经过一次的路径
汉密尔顿路径：每个节点恰好经过一次的路径

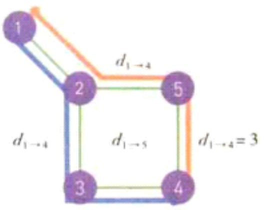
(a)



路径

一个节点序列，序列中每个节点有链接指向其后面紧邻的节点。每个路径包含 $n+1$ 个节点和 n 个链接。路径的长度是其包含的链接数，多次出现的链接重复计数。例如，橙色路径 $1\rightarrow2\rightarrow5\rightarrow4\rightarrow3$ 的路径长度为 4。

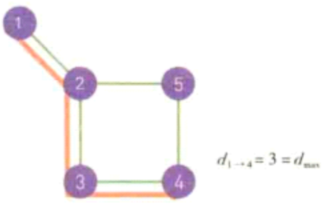
(b)



最短路径

两个节点间长度最短的路径。我们将两个节点之间最短路径的长度称为测地距离 d。注意，最短路径不一定是唯一的：节点 1 和节点 4 之间有两条最短路径，分别是 $1\rightarrow2\rightarrow3\rightarrow4$（蓝色）和 $1\rightarrow2\rightarrow5\rightarrow4$（橙色），它们的路径长度都为 $d_{1,4}=3$。

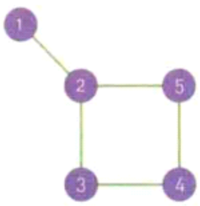
(c)



直径（d_{max}）

最短路径的最大长度，也是距离最远的节点对之间的距离。如图所示，节点 1 和节点 4 之间的距离最远，因此网络直径为 $d_{max}=3$。

(d)

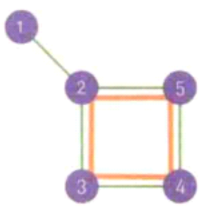


$$\langle d\rangle=(d_{1\to2}+d_{1\to3}+d_{1\to4}+d_{1\to5}+d_{2\to3}+d_{2\to4}+d_{2\to5}+d_{3\to4}+d_{3\to5}+d_{4\to5})/10=1.6$$

平均路径长度（$\langle d\rangle$）

所有节点对之间最短路径的平均长度。如图所示，该网络的平均路径长度为 $\langle d\rangle=1.6$，计算过程列在了图的旁边。

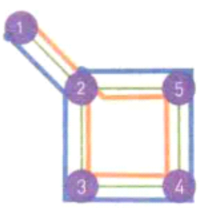
(e)



环

起始节点和终止节点相同的路径。如图所示，橙色路径是一个环。

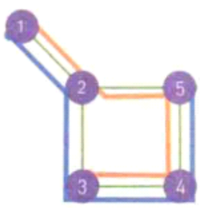
(f)



欧拉路径

经过每条链接恰好一次的路径。图中给出了两条欧拉路径，分别为橙色和蓝色的。

(g)



汉密尔顿路径

经过每个节点恰好一次的路径。图中给出了两条汉密尔顿路径，分别为橙色和蓝色的。

广度优先算法：

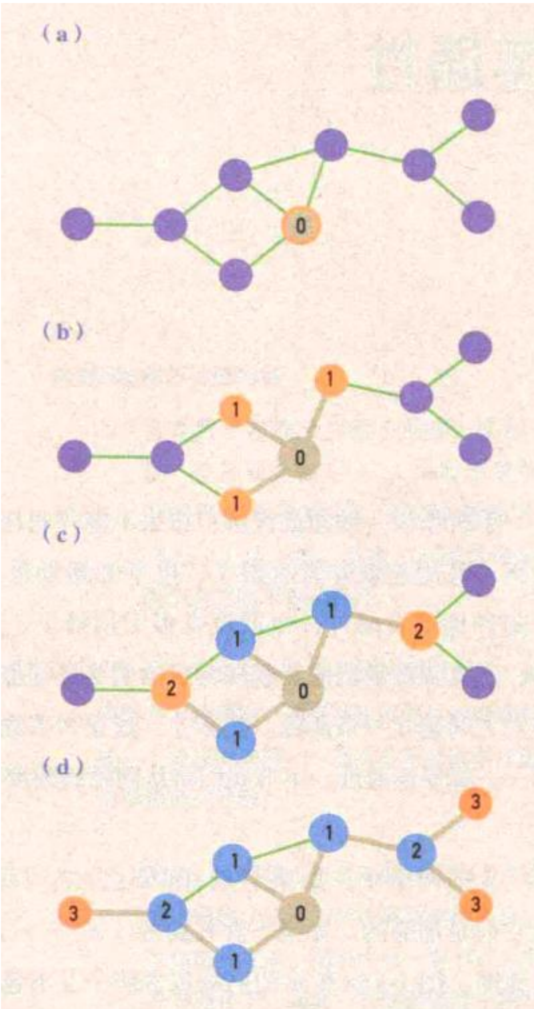
（1）从节点 i 出发，将其标记为 0；

（2）找出与节点 i 直接相连的节点，标记为距离 1，然后将它们放到一个队列中；

（3）从队列中取出最前面的节点（其距离标记假设为 n），找出和该节点相连且尚未被标记的节点，将它们标记为 $n+1$，然后把它们放在队列的后面；

（4）重复步骤 3，直到碰到目标节点 j 或者队列中没有节点了；

（5）节点 i 和节点 j 之间的距离为节点 j 的标记。如果节点 j 没有标记，则 $d_{ij}=\infty$



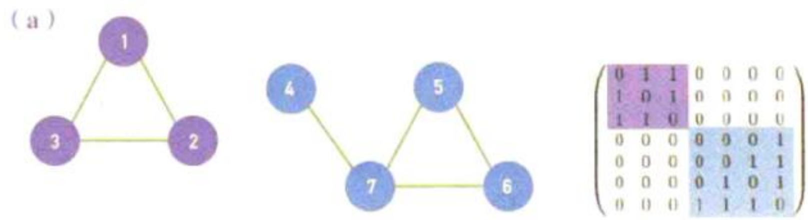
复杂度是(N+L),N为节点数，L为链接数。

第九节 连通性

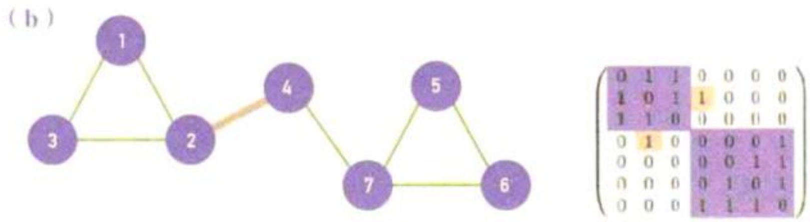
2022年12月26日　　9:30

在无向图中，如果节点i与节点j之间存在路径，则我们称之为连通，若不存在路径则不连通。

对于整个网络而言，如果网络中所有节点都是连通的，则称这个网络是连通的，否则有至少一个不连通的节点，则网络不是连通的。对于不连通的网络来说，将网络的邻接矩阵重新排序后，我们可以使用线性代数中的分块对角矩阵来判别。



而想让两个不连通的网络重新连通，我们需要找到关键节点，将两个节点架桥就可以实现。



寻找连通分支：

寻找网络的连通分支

1. 从随机选择的一个节点 i 出发，执行边栏 1.5 描述的 BFS 算法。将从节点 i 出发可到达的所有节点标记为 $n=1$。

2. 如果已标记节点的个数等于 N，则网络是连通的。如果已标记节点的个数小于 N，则网络包含多个连通分支。找出这些连通分支需要继续执行步骤 3。

3. 标号增加 1，即 $n \to n+1$。选择一个未标记的节点 j，将其标记为 n。使用 BFS 算法找出从节点 j 出发可到达的所有节点，将它们都标记为 n。重复步骤 3，直到标记完所有的节点为止。

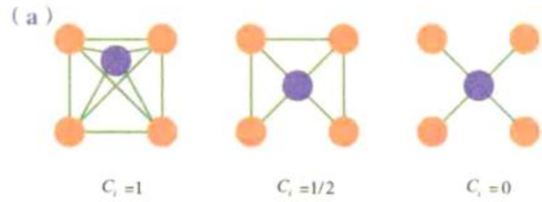
第十节 集聚系数

2022年12月26日 9:38

局部集聚系数：刻画了相邻节点之间连接的稠密程度
对于一个度为K的节点i的集聚系数计算为：

$$C_i = \frac{2L_i}{k_i\left(k_i - 1\right)}$$

Li为与该节点连接的节点之间的连接数

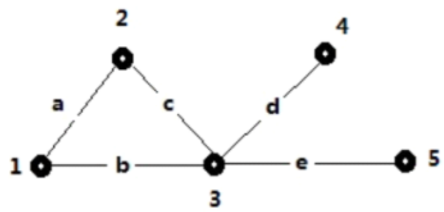


例如第二个，与紫球连接的是四个黄球，四个黄球之间的链接数是3，则L=3
集聚系数是一个0-1的值，若集聚系数为0，则表明该节点不连通，若为1，则表明这个节点与其余所有节点都连通。

全局集聚系数：主要用于计算网络中闭合三角形的数量

$$C_\Delta = \frac{3 \times 三角形个数}{连通三元组的个数}$$

例如下图{1，(2,3)}构成的triplet是封闭的，{3，(4,5）}构成的triplet是开放的



全局的Clustering coefficient比较简单，公式如下：Clustering coefficient(global) = number of closed triplet / number of triplet(closed+open)
以上图为例:
closed triplet ={1，(2,3)}，{2，(1,3)}，{3，(1,2)}
all triplet = {1，(2,3)}，{2，(1,3)}，{3，(1,2)}，{3，（2,4）}，{3，（4,5）}，{3，（1,5）}，{3，（2,5）}，{3，（1,4）}
number of closed triplet = 3
number of triplet = 8
number of triplet / number of triplet = 3/8=**0.375**

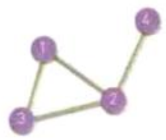
平均局部集聚系数：所有局部集聚系数的平均值，一般用来衡量网络的平均集聚情况。

第一章概念补充和总结

2022年12月26日 9:47

（a）

无向



$$A_{ij}=\begin{pmatrix}0&1&1&0\\1&0&1&1\\1&1&0&0\\0&1&0&0\end{pmatrix}$$

$A_{ii}=0$ $A_{ij}=A_{ji}$

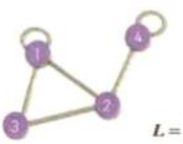
$L=\frac{1}{2}\sum_{i,j=1}^{N}A_{ij}$ $<k>=\frac{2L}{N}$

无向网络

无向网络中链接没有方向。例如，互联网、电网、科学合作网络。

（b）

自环



$$A_{ij}=\begin{pmatrix}1&1&1&0\\1&0&1&1\\1&1&0&0\\0&1&0&1\end{pmatrix}$$

$\exists i, A_{ii}\neq 0$ $A_{ij}=A_{ji}$

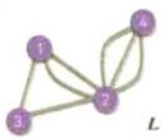
$L=\frac{1}{2}\sum_{i,j=1,i\neq j}^{N}A_{ij}+\sum_{i=1}^{N}A_{ii}$?

自环

在许多网络中，节点不和自己进行交互，因此其邻接矩阵的对角线元素是0，即 $A_{ii}=0, i=1,\cdots,N$。在有些系统中，自我交互是允许的，在这样的网络中，自环表示节点的自我交互。例如，万维网、蛋白质相互作用网络。

（c）

多重
无向



$$A_{ij}=\begin{pmatrix}0&2&1&0\\2&0&1&3\\1&1&0&0\\0&3&0&0\end{pmatrix}$$

$A_{ii}=0$ $A_{ij}=A_{ji}$

$L=\frac{1}{2}\sum_{i,j=1}^{N}A_{ij}$ $<k>=\frac{2L}{N}$

多重网络 / 简单网络

在多重网络中，两个节点之间可以有多个平行的链接。因此，A_{ij} 可以是任意正整数。不允许多重链接存在的网络被称为简单网络。多重网络的一个例子是：区分朋友关系、家庭关系和同事关系的社会网络。

（d）

有向

$$A_{ij}=\begin{pmatrix}0&1&0&0\\0&0&1&1\\1&0&0&0\\0&0&0&0\end{pmatrix}$$

$A_{ij}\neq A_{ji}$

$L=\sum_{i,j=1}^{N}A_{ij}$ $<k>=\frac{L}{N}$

有向网络

有向网络中，链接具有方向。例如，万维网、手机通话网络、引文网络。

（e）

加权
无向

$$A_{ij}=\begin{pmatrix}0&2&0.5&0\\2&0&1&4\\0.5&1&0&0\\0&4&0&0\end{pmatrix}$$

$A_{ii}=0$ $A_{ij}=A_{ji}$

$<k>=\frac{2L}{N}$

加权网络

加权网络中，链接具有权重、强度或流量参数。对于权重为 w_{ij} 的链接，其邻接矩阵中对应的元素为 $A_{ij}=w_{ij}$。对于无权网络，邻接矩阵的元素仅表示链接是否存在（$A_{ij}=1$ 或 $A_{ij}=0$）。加权网络的例子有：手机通话网络、电子邮件网络。

（f）

完全图
无向

$$A_{ij}=\begin{pmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix}$$

$A_{ii}=0$ $A_{i\neq j}=1$

$L=L_{max}=\frac{N(N-1)}{2}$ $<k>=N-1$

完全图（团）

完全图或团中，任意两个节点彼此相连。例子包括出演同一部电影的演员网络，其是每两个演员都彼此相连。

第一节 定义随机网络

2022年12月12日　　8:55

真实世界的网络往往具有随机性，它的结构性和观赏性可能并不强。
在当前我们主要用来定义一个随机网络的方法有两种：
一种是G（N,L）方法：N个节点通过L个随机路径链接
另一种是G（N，p）方法：N个节点中彼此相连的概率为p

我们一般使用更具备操作性的方法2，也就是G（N，p）方法，那么我们建立一个随机网络的步骤是：

（1）从N个孤立节点开始。

（2）选择一对节点，产生一个0到1之间的随机数。如果该随机数小于p，在这对节点之间放置一条链接；否则，该节点对保持不连接。

（3）对所有$N(N-1)/2$个节点对，重复步骤（2）。

这样我们就建立了随机网络。
随机网络又叫埃尔德什-雷尼网络

第二节 链接数

2022年12月27日 10:05

对于产生的随机网络，我们有如下数学表示链接数：

（1）L 个点对之间存在链接的概率，即 p^L。

（2）剩余 $N(N-1)/2-L$ 个点对之间没有链接的概率，即 $(1-p)^{N(N-1)/2-L}$。

（3）在所有 $N(N-1)/2$ 个点对中选择 L 个点对放置链接，所有可能的选择方式数为：

$$\begin{pmatrix} \frac{N(N-1)}{2} \\ L \end{pmatrix} \quad (2.0)$$

因此，随机网络恰好有 L 条链接的概率为：

$$p_L = \begin{pmatrix} \frac{N(N-1)}{2} \\ L \end{pmatrix} p^L (1-p)^{\frac{N(N-1)}{2}-L} \quad (2.1)$$

因此我们可以根据上式来计算期望：随机网络的期望连接数：

$$\langle L \rangle = \sum_{L=0}^{\frac{N(N-1)}{2}} L p_L = p \frac{N(N-1)}{2} \quad (2.2)$$

随机网络平均度是：

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1)$$

我们可以看到随机网络的相关性质满足二项分布:
我们会常用到二项分布的式子:

二项分布的形式为:

$$p_x = \binom{N}{x} p^x (1-p)^{N-x}$$

分布的均值（一阶矩）为:

$$x = \sum_{x=0}^{N} x p_x = Np \tag{2.4}$$

其二阶矩为:

$$x^2 = \sum_{x=0}^{N} x^2 p_x = p(1-p)N + p^2 N^2 \tag{2.5}$$

因此，二项分布的标准差为:

$$\sigma_x = \left(x^2 - x^2\right)^{\frac{1}{2}} = \left[p(1-p)N\right]^{\frac{1}{2}} \tag{2.6}$$

一阶矩：指的是期望，
二阶矩：指的是变量的平方的期望
二阶中心距：方差
三阶中心矩指的是随机变量的偏度
四阶中心矩指的是随机变量的峰度

第三节 随机网络的度分布

2022年12月28日 9:15

在随机网络中，有一些节点很重要，他有很多的链接数，但是有一些节点可能存在于网络的边缘，链接数极少。
所以我们可以通过分析一个随机网络的度分布来查看，节点链接分布情况。

第一个常见的是二项分布：

（1）k 个链接出现的概率，即 p^k。

（2）剩下 $(N-1-k)$ 个链接不出现的概率，
即 $(1-p)^{N-1-k}$。

（3）节点 i 的 $N-1$ 个可能存在的链接中选出 k 个，
选择方式的总数为：

$$\binom{N-1}{k}$$

此，随机网络的度分布服从二项分布：

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} \qquad (2.7)$$

其次大部分真实网络都是极其稀疏的，因此**网络平均度的大小远远小于节点数N**，此时会形成泊松分布的情况：

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \qquad (2.8)$$

公式 2.8 和公式 2.7 通常被称为随机网络的度分布。

注意，我们所说的满足泊松分布的条件是平均度<k><<N，所以小网络更可能是二项分布的，

大网络更可能是泊松分布的。

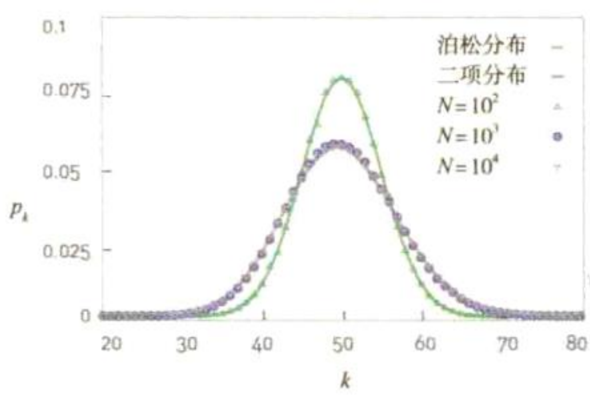


图 2-5

度分布与网络大小无关

平均度 $\langle k \rangle = 50$，大小分别为 $N=10^2$、$N=10^3$、$N=10^4$ 的三个随机网络的度分布。

小网络：二项分布

对于小网络（$N=10^2$），由于不满足泊松近似的条件 $N \gg \langle k \rangle$，该网络的度分布明显偏离泊松分布（公式 2.8）。因此，小网络的度分布需要使用精确的二项分布形式（公式 2.7）（绿线）。

大网络：泊松分布

对于大网络（$N=10^3$，$N=10^4$），其度分布与灰线所示的泊松分布（公式 2.8）相差无几。因此，当网络大小 N 很大时，度分布和网络大小无关。为了避免随机性带来的噪声，图中所示的结果是在 1 000 个独立生成的随机网络上平均得到的。

第四节 真实网络不是泊松分布的

2022年12月28日 9:26

我们通过书中的社会网络的假设可以得到一个结论：
在大的随机网络中，大多数的度节点分布在平均度的极小且狭窄的范围内。

为什么没有度很大的节点？
根据斯特林近似：

有：

$$k! \sim \left[\sqrt{2\pi k}\right]\left(\frac{k}{e}\right)^{k}$$

因此，公式 2.8 可以重写为：

$$p_k = \frac{e^{-\langle k\rangle}}{\sqrt{2\pi k}}\left(\frac{e\langle k\rangle}{k}\right)^{k} \qquad (2.9)$$

我们发现，随着k的增大，p会急速缩小，这个速度比指数级更快，因此我们机会没有足够大的概率观测到一个k极大的节点。

但是在现实生活的网络中，我们确实会观测到一些度很大的节点，这也说明真实网络并不完全是随机网络。

第五节 随机网络的演化

2022年12月28日　　9:37

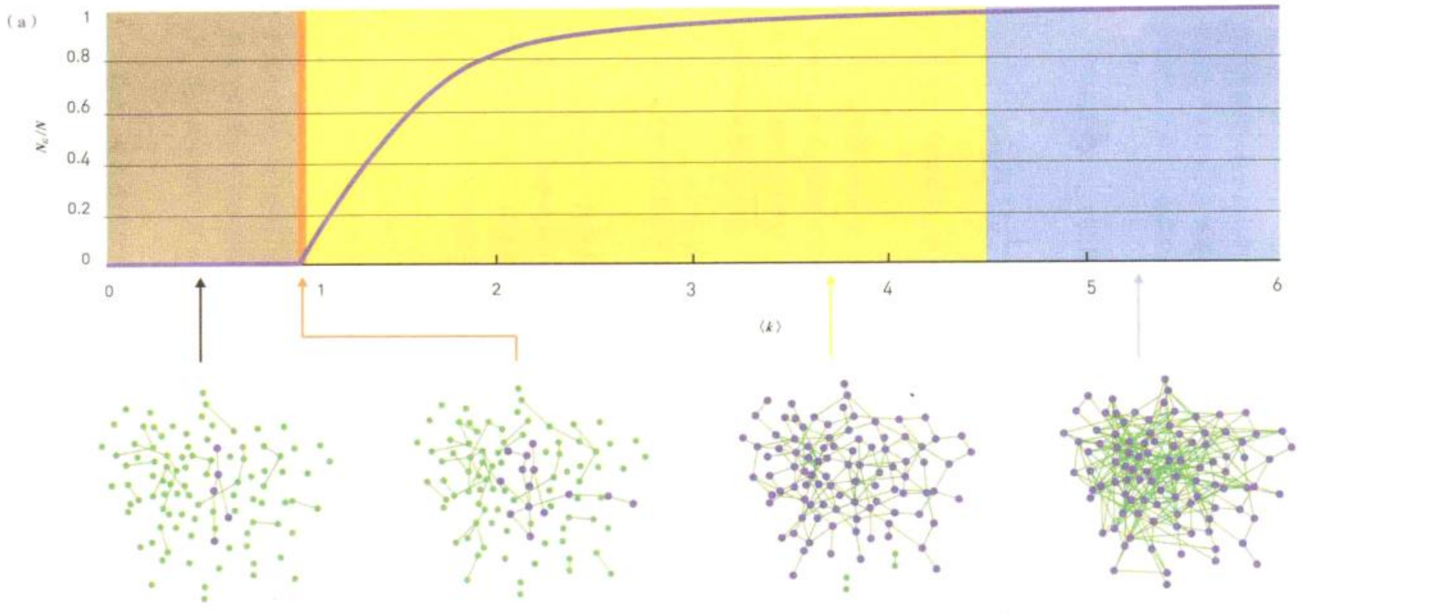
我们通过量化连通分支大小和平均度的大小变化来设想随机网络的演化是如何进行的。

- 当 $p=0$ 时，$\langle k\rangle=0$，所有节点都是孤立的。因此，最大连通分支的大小为 $N_G=1$。对于大的 N，有 $N_G/N\to 0$。

- 当 $p=1$ 时，$\langle k\rangle=N-1$，网络是完全连通图，所有节点属于同一个连通分支。因此，$N_G=N$，$N_G/N=1$。

一般人可能会认为，平均度和连通分支的大小可能是逐渐变化的，但是其实并不是这样，
在<k>较小时，NG/N一直为0，但是当平均度到达阈值之后，NG/N开始迅速增长，这意味着一个巨大的连通分支出现了，我们称为巨分支。
根据计算预测，巨分支出现的平均度阈值是<k>=1.

网络的演化过程：



(b) 亚临界状态

- 无巨连通分支
- 连通分支大小的分布：$p_s\sim s^{-3/2}e^{-\alpha s}$
- 最大连通分支的大小：$N_G\sim \ln N$
- 连通分支是树

(c) 临界点

- 无巨连通分支
- 连通分支大小的分布：$p_s\sim s^{-3/2}$
- 最大连通分支的大小：$N_G\sim N^{2/3}$
- 连通分支中可能包含环

(d)

- 有一个巨连通分支
- 小连通分支大小的分布：$p_s\sim s^{-3/2}e^{-\alpha s}$
- 巨连通分支的大小：$N_G\sim(p-p_c)N$
- 小连通分支是树
- 巨连通分支中有环

(e) 连通状态

- 仅有一个巨连通分支构成
- 没有孤立节点和小连通分支
- 巨连通分支的大小：$N_G=N$
- 巨连通分支中有环

图 2-7
随机网络的演化

(a) 埃尔德什－雷尼随机网络模型中，巨连通分支的相对大小关于平均度 $\langle k\rangle$ 的函数。从图中可以看出，该函数在 $\langle k\rangle=1$ 处有一个相变，该相变标志着具有非零相对大小的巨连通分支出现了。

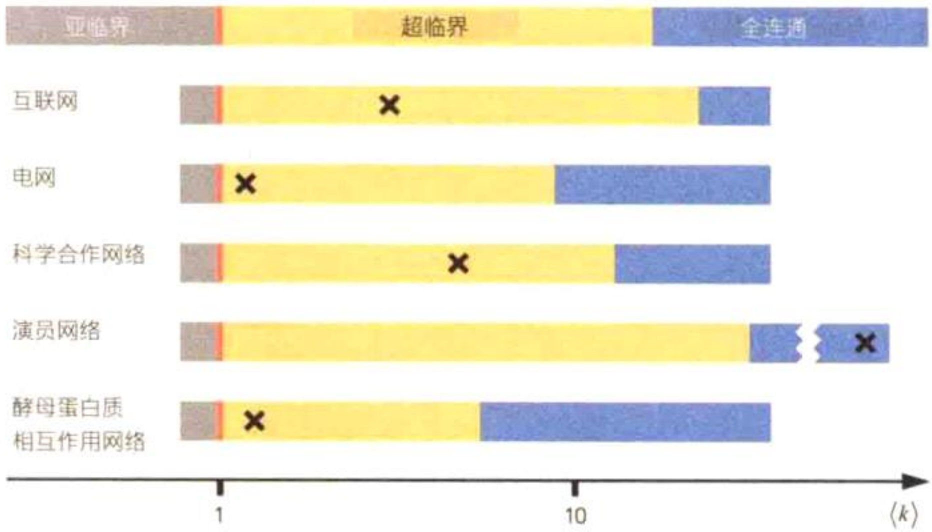
真实网络一般是超临界状态的。

第六节 真实网络是超临界的

2022年12月30日　　10:03

大部分的真实网络都是超临界的，
但是超临界就意味着，有一些小的连通分支并不能与大的连通分支连通，这是一个令人不安的结论。
这意味着，如果互联网网络是超临界的，那么意味着一些小的网络路由没有办法连接到万维网之中。

我们常见的网络如下：



判断一个真实网络出于什么状态，我们可以计算网络的平均度，1< <k> <lnN,则这个网络是超临界的。

第七节 小世界网络

2022年12月31日　9:14

小世界现象又称为六度分隔，是指地球上的两个人，他们之间最多间隔6个相识关系。
这意味着，在真实网络中，任意两个节点的间隔距离其实很短。
这可以通过平均度的理论来回答：

$\langle k\rangle$ 个距离为 1 的节点（$d=1$）；
$\langle k\rangle^2$ 个距离为 2 的节点（$d=2$）；
$\langle k\rangle^3$ 个距离为 3 的节点（$d=3$）；
……
$\langle k\rangle^d$ 个距离为 d 的节点。

从而引出网络直径的概念：网络直径指的是网络中的最大距离
网络直径的求法为：

$$d_{\max} \approx \frac{\ln N}{\ln\langle k\rangle}$$

第八节 随机网络集聚系数

2022年12月31日　9:36

我们之前说了网络的局部集聚系数，现在我们说一下随机网络的集聚系数

对于随机网络中的一个节点 i，要计算其集聚系数 C_i，我们需要估计出该节点的 k_i 个邻居之间的链接数 L_i。随机网络中，节点 i 的两个邻居之间的链接概率是 p。由于节点 i 的 k_i 个邻居之间最多有 $k_i(k_i-1)/2$ 条链接，L_i 的期望值为：

$$\langle L_i \rangle = p\frac{k_i(k_i-1)}{2} \tag{2.20}$$

因此，随机网络的局部集聚系数为：

$$C_i = \frac{2\langle L_i \rangle}{k_i(k_i-1)} = p = \frac{\langle k \rangle}{N} \tag{2.21}$$

那么随机网络就具备以下两种性质：

（1）给定 $\langle k \rangle$ 的情况下，网络越大，节点的集聚系数越小。因此，节点的局部集聚系数和 $1/N$ 成正比。注意，网络的平均集聚系数 $\langle C \rangle$ 也服从公式 2.21。

（2）随机网络中，节点的局部集聚系数和节点的度相互独立。

但是真实网络并不符合随机网络的性质，真实网络的集聚系数要比随机网络高得多。