

# 链接选择模型生成无标度网络

2021级萃英班 张羽翔（整理自barabasi *Network Science*）

无标度网络的产生本质有两点：网络生长和偏好链接。而偏好链接的起源有多种解释，换言之，我们有多种方法产生网络的偏好链接。

将产生偏好链接的方法归为两类：

- 局部机制

局部机制相当于在微观层面解释系统线性偏好链接的起源，我们关注网络生长中加入一个新节点的情况，比较成熟的方法有：

1. 链接选择模型

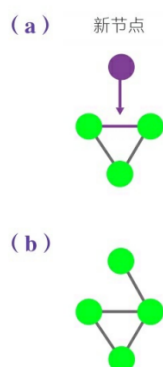


图4-13 链接选择模型

(a) 网络通过添加新节点而生长，每个新节点在加入时会随机选择一条链接（紫色）。

(b) 新节点以相同的概率从该链接的两个端点中选择一个与之相连。在这个例子中，新节点选择了链接的右侧端点。

2. 复制模型

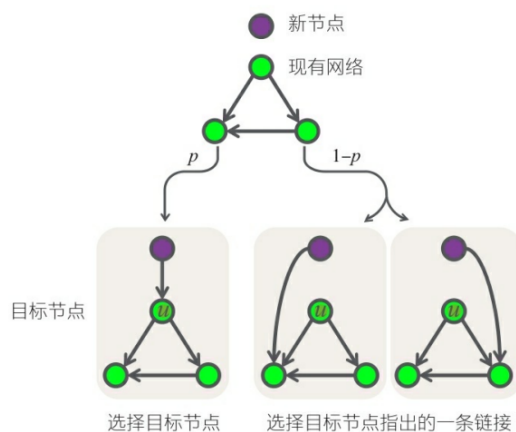


图4-14 复制模型

复制模型的主要步骤。新节点以概率 $p$ 连接到一个随机选择的目标节点 $u$ ，或者以概率 $1-p$ 连接到节点 $u$ 指向的某个节点。换言之，新节点以概率 $1-p$ 复制了目标节点 $u$ 的一条链接。

## • 优化模型

优化模型是从整体上来看，在网络的节点间链接的形成更倾向于连接度大的节点以获得较小的效益，从而形成无标度网络。

以互联网为例，其节点是通过电缆连接在一起的路由器。建立一条新的互联网链接需要在两台路由器之间铺设电缆。由于铺设电缆的成本很高，在铺设之前需要仔细进行成本效益分析。每个新加入互联网中的路由器在选择建立链接时，必将在获得良好的网络性能（例如带宽）和电缆建设费用（例如物理距离）之间做出权衡。二者可能是相互冲突的两种诉求，因为最近的节点不一定能够提供最好的网络性能。

在优化模型中，幂律起源于两种相互竞争的机制：

1. 优化：每个节点都有一个吸引域，落在该节点的吸引域中的节点就会和它相连。每个吸引域的大小和处于域中心的节点 $j$ 的 $h_j$ 有关。
2. 随机：我们随机选择新节点的位置，落在 $N$ 个吸引域之中的某一个。已有节点中，度最大的节点具有最大的吸引域，因此会吸引最多的新节点，得到最多的新链接，最终形成偏好连接。

**本文主要计算链接选择模型，证明其具有幂律分布，从而验证其可以生成无标度网络。**

## 链接选择模型的推导

链接选择模型的推导如下：

### 1. 生长

在每个时间步，我们向网络中添加一个新节点。

### 2. 链接选择

随机选择一条链接，将新节点连接到所选链接的其中一个端点上。

该模型不需要全局网络拓扑信息，因此本质上属于局部和随机机制。与巴拉巴西 - 阿尔伯特模型不同，该模型缺少一个内置的 $\Pi(k)$ 函数。然而，接下来我们就会看到，该模型能够产生偏好连接。

首先，一条随机选择的链接，其端点的度为 $k$ 的概率为：

$$q_k = Ckp_k$$

它刻画了两个效应：

第一，节点的度越高，该节点就越容易成为随机选择的链接的一端。

第二，网络中度为 $k$ 的节点越多（也就是 $p_k$ 越高），这些节点越容易成为随机选择的链接的一端。在公式中，常数 $C$ 可以通过归一化条件 $\sum q_k = 1$ 计算得到，即 $C = \frac{1}{\langle k \rangle}$ 。因此，一条随机选择的链接其端点的度为 $k$ 的概率为：

$$q_k = \frac{kp_k}{\langle k \rangle}$$

是新节点和度为 $k$ 的节点相连的概率。公式中，概率 $q_k$ 和 $k$ 是线性关系。这表明，链接选择模型通过产生偏好连接构造出了无标度网络。

# 用线性偏好链接构造无标度网络

下面写代码用线性偏好链接构造无标度网络。

程序用python编写，主要利用networkx库进行网络生成：

```
import networkx as nx
import random
import numpy as np
import matplotlib.pyplot as plt
```

创建一个空的无向图

```
G = nx.Graph()
```

添加初始节点

```
G.add_node(0)
G.add_node(1)
G.add_edge(0, 1)
```

生成一个10000节点的网络，生成过程利用链接选择模型

```
nn = 10000
# 迭代nn步
for i in range(2, nn+1):
    # 随机选择一个原有的链接
    existing_link = random.choice(list(G.edges()))
    # 随机选择链接的一端
    existing_node = random.choice(existing_link)
    # 添加新节点
    G.add_node(i)
    # 将新节点连接到选择的一端
    G.add_edge(i, existing_node)
```

绘制度分布图

```
degree_sequence = sorted([d for n, d in G.degree()], reverse=True)
degree_sequence = np.array(degree_sequence)
unique_values, counts = np.unique(degree_sequence, return_counts=True)
```

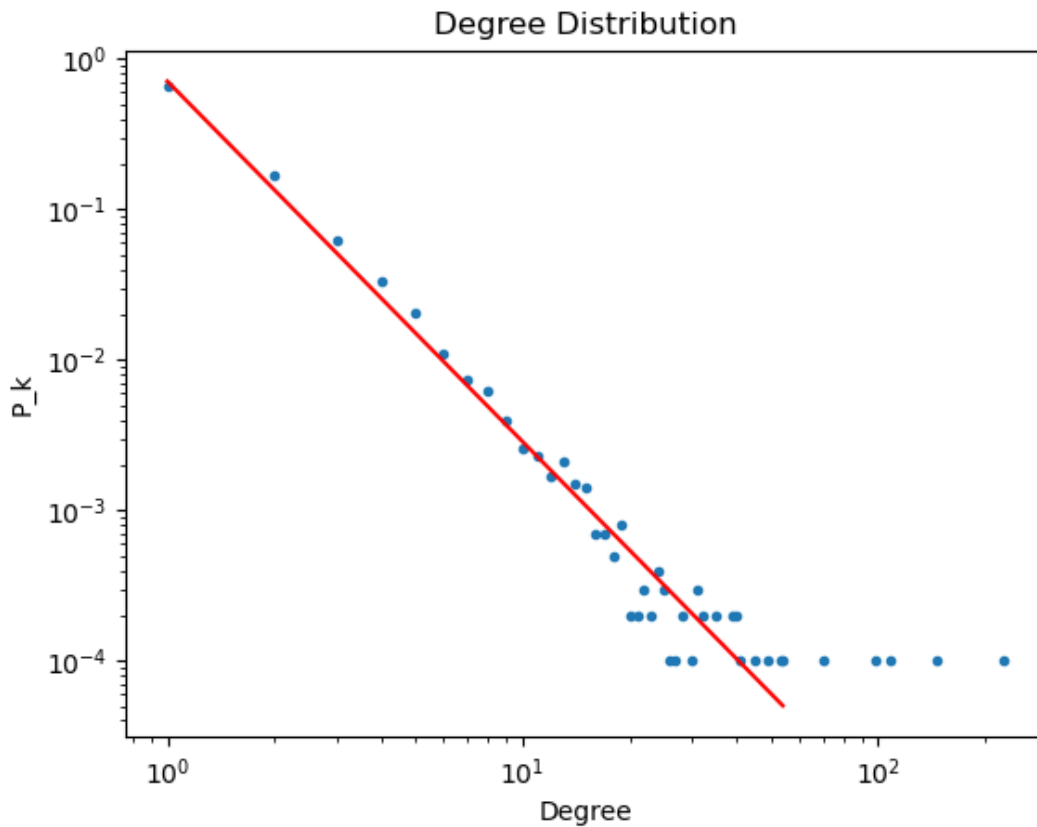
作截断并拟合度分布

```
uni = unique_values[0:-5]
cou = counts[0:-5]
x = np.log(uni)
y = np.log(cou/nn)
k, b = np.polyfit(x, y, 1)
print(k, b)
yy = np.exp(b) * uni**k
```

画出度分布图及拟合曲线

```
plt.loglog(unique_values, counts/nn, marker='.', linestyle='none')
plt.loglog(uni, yy, color='r')
plt.title("Degree Distribution")
plt.xlabel("Degree")
plt.ylabel("P_k")
plt.show()
```

结果如下：

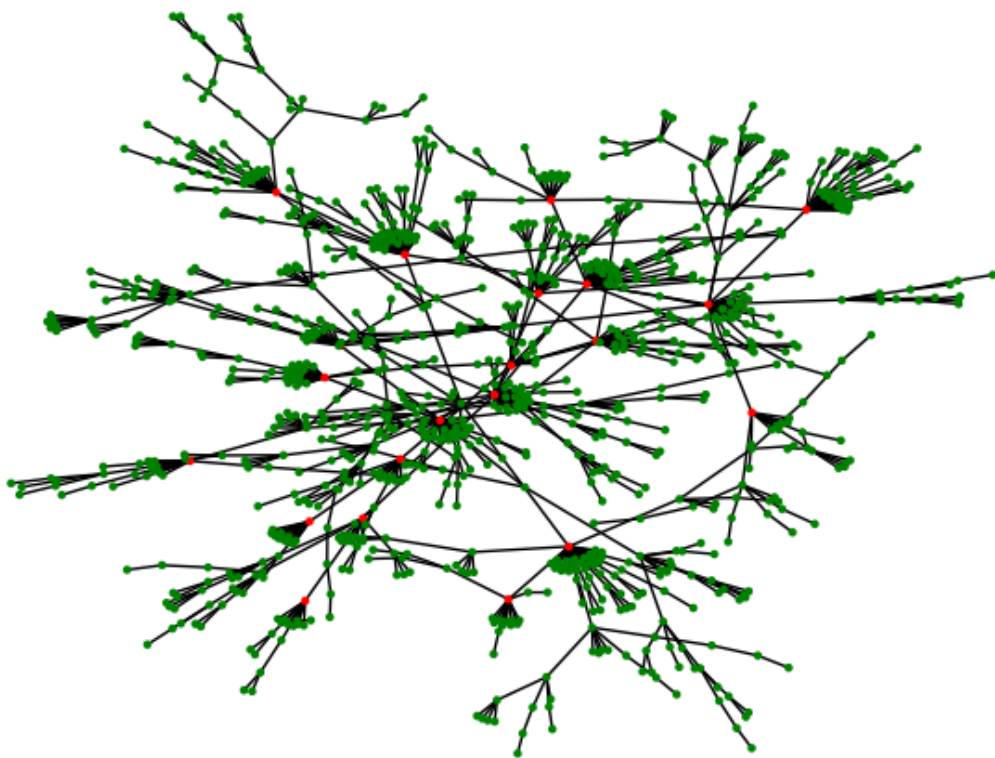


可见，度分布在指数坐标下呈线性关系，说明度分布满足幂律分布。拟合直线斜率大概为-2.5左右。

其中，图像右下角散点几乎平行分布是正常现象，这是由于在无标度网络中大度的节点往往很少，这时候在指数坐标下它们的概率 $p_k$ 就非常相近，形成一条水平直线，在做拟合时，为了避免影响结果，可以对后面几个大度节点做截断。

由此可以说明，链接选择模型可以生成线性偏好链接，由此生成无标度网络。

1000节点的网络图（度大于10的节点标为红色，小于10的为绿色）：



节点度分布的图像：



链接选择模型原始论文：

*S.N. Dorogovtsev and J.F.F. Mendes. Evolution of networks. Oxford Clarendon Press, 2002.*