Ch5 复杂网络机制及演化模型

复杂网络

Network	Reference
WWW, site level, undir.	Adamic 1999
Internet, domain level	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	Watts, Strogatz 1998
LANL coauthorship	Newman 2001a,b
MEDLINE coauthorship	Newman 2001a,b
SPIRES coauthorship	Newman 2001a,b,c
NCSTRL coauthorship	Newman 2001a,b
Math coauthorship	Barabási *et al.* 2001
Neurosci. coauthorship	Barabási *et al.* 2001
E. coli, substrate graph	Wagner, Fell 2000
E. coli, reaction graph	Wagner, Fell 2000
Ythan estuary food web	Montoya, Solé 2000
Silwood park food web	Montoya, Solé 2000
Words, cooccurence	Cancho, Solé 2001
Words, synonyms	Yook *et al.* 2001
Power grid	Watts, Strogatz 1998
C. Elegans	Watts, Strogatz 1998

复杂网络出现在自然系统或人造系统的各个领域

复杂网络

- Internet(路由器，线路)
- WWW (web页面，超连接)
- 文章合作网络(作者，合作关系)
- 食物链(动物，捕食关系)
- 化学反应(分子，发生化学反应)

……

复杂网络的形成机制

- 自组织

  个体

  偏好

  生长

- 涌现

几个重要的复杂网络模型

- 随机网络模型
- 小世界网络模型
- 无标度网络模型

随机网络模型

- Input: *(n,p)*

  n 是节点个数，*p*是边出现的概率。
- 二项式算法

  从*n*个孤立点开始

  对任意一对顶点，以概率*p*连接。

随机网络的特征

(1) 聚集系数较小；

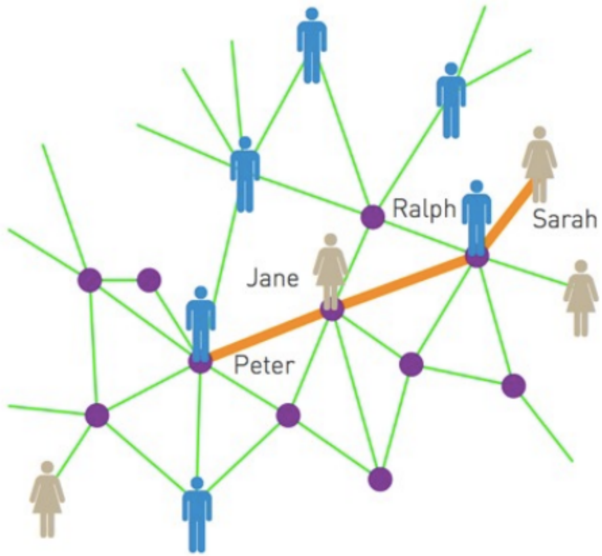
(2) 网络平均距离小；

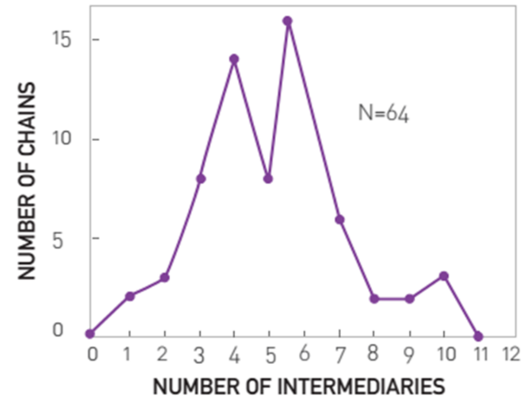
(3) 节点度服从Poisson分布。

随着概率p 从0到1逐渐增加，网络的某些性质会突然出现。

小世界特性

Milgram小世界实验





Milgram实验: 信件传递.

1967年，哈佛大学的社会心理学家米尔格兰姆(Stanley Milgram)设计了一个连锁信件实验。他将一套连锁信件随机发送给居住在内布拉斯加州奥马哈的160个人，信中放了一个波士顿股票经纪人的名字，信中要求每个收信人将这套信寄给自己认为是比较接近那个股票经纪人的朋友。朋友收信后照此办理。最终，大部分信件在经过五、六个步骤后都抵达了该股票经纪人。

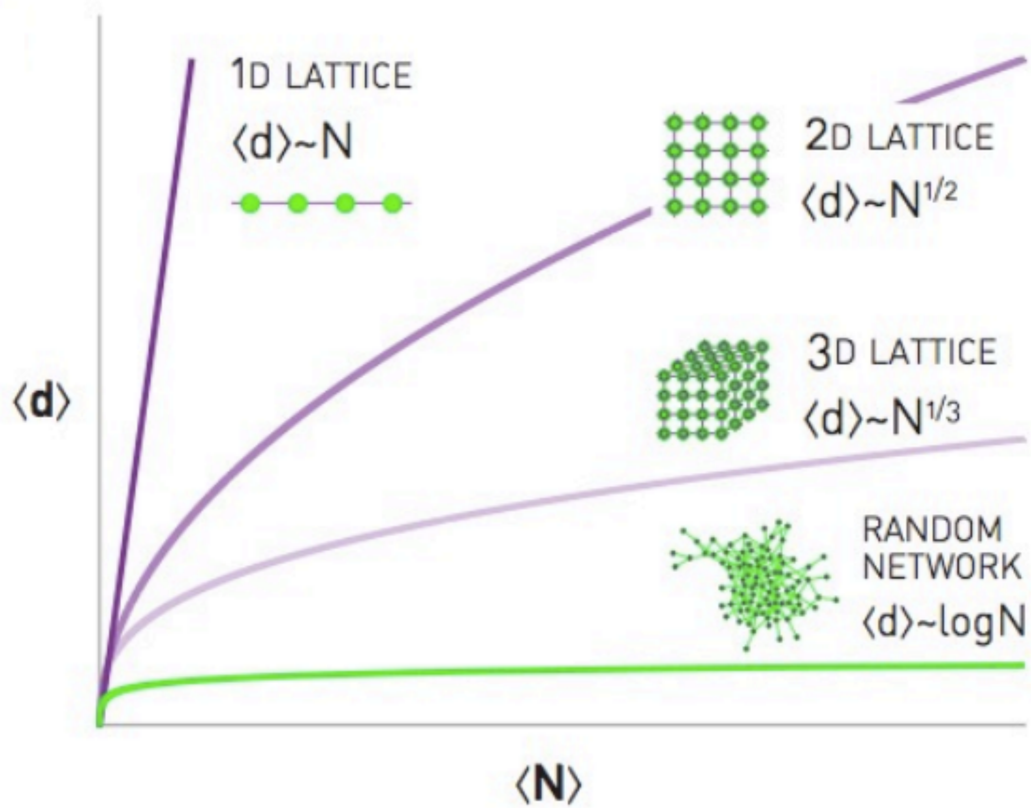
六度空间

- 上世纪60年代哈佛大学社会心理学家Stanley Milgram社会调查后推断出: 世界上任意两个人平均距离是6.

小世界中“小”的含义

平均距离

- 平均距离与网络规模比起来是小的
- 与网络规模的对数成正比
- ER网络就具有这个性质 $\langle d \rangle \sim \frac{\log N}{\log \langle k \rangle}$


1D LATTICE
⟨d⟩~N
2D LATTICE
⟨d⟩~N1/2
3D LATTICE
⟨d⟩~N1/3
RANDOM NETWORK
⟨d⟩~logN
⟨d⟩
⟨N⟩

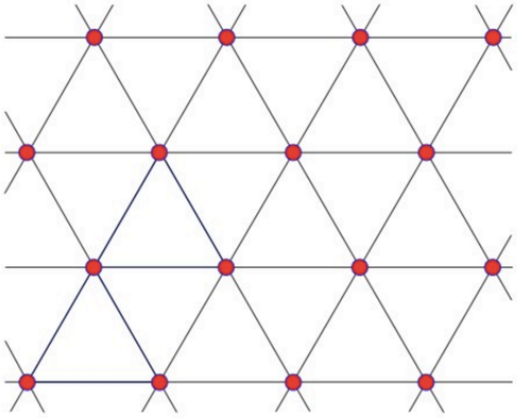
集聚系数

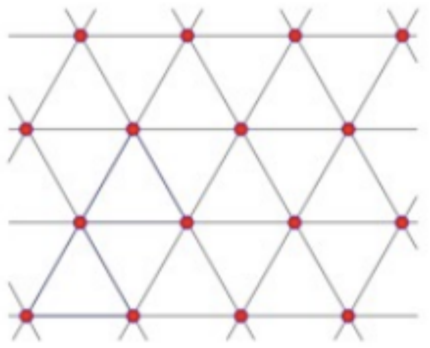
- ER网络的集聚系数

$$C = \frac{\langle k \rangle}{N}$$

- 实际网络的集聚系数

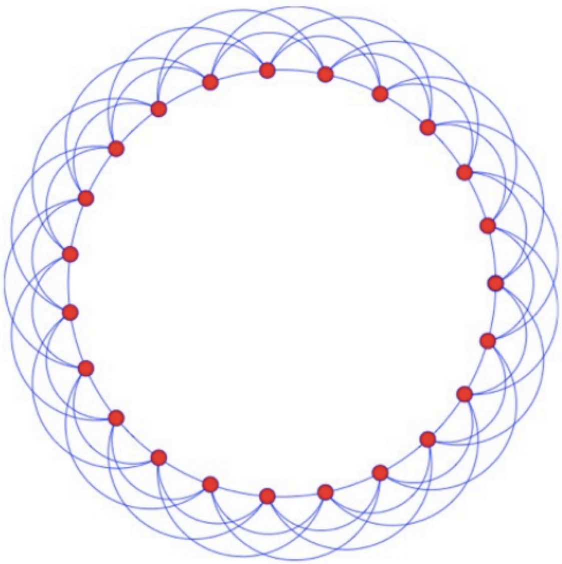
- 晶格的集聚系数





实际网络

最近邻居连接网络



N， m=3

$$C_i = \frac{3m(m-1)/2}{\binom{2m}{2}} = \frac{3(m-1)}{2(2m-1)}$$

$$d^{circle} = \frac{N+2m-1}{4m}$$

$$d^{circle} = \frac{N}{4m}$$

Watts-Strogatz model

Duncan Watts

NATURE | VOL 393 | 4 JUNE 19

Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Networks of coupled dynamical systems have been used to m biological oscillators[1–4], Josephson junction arrays[5,6], excit media[7], neural networks[8–10], spatial games[11], genetic co networks[12] and many other self-organizing systems. Ordina the connection topology is assumed to be either compl regular or completely random. But many biological, technolo and social networks lie somewhere between these two extre

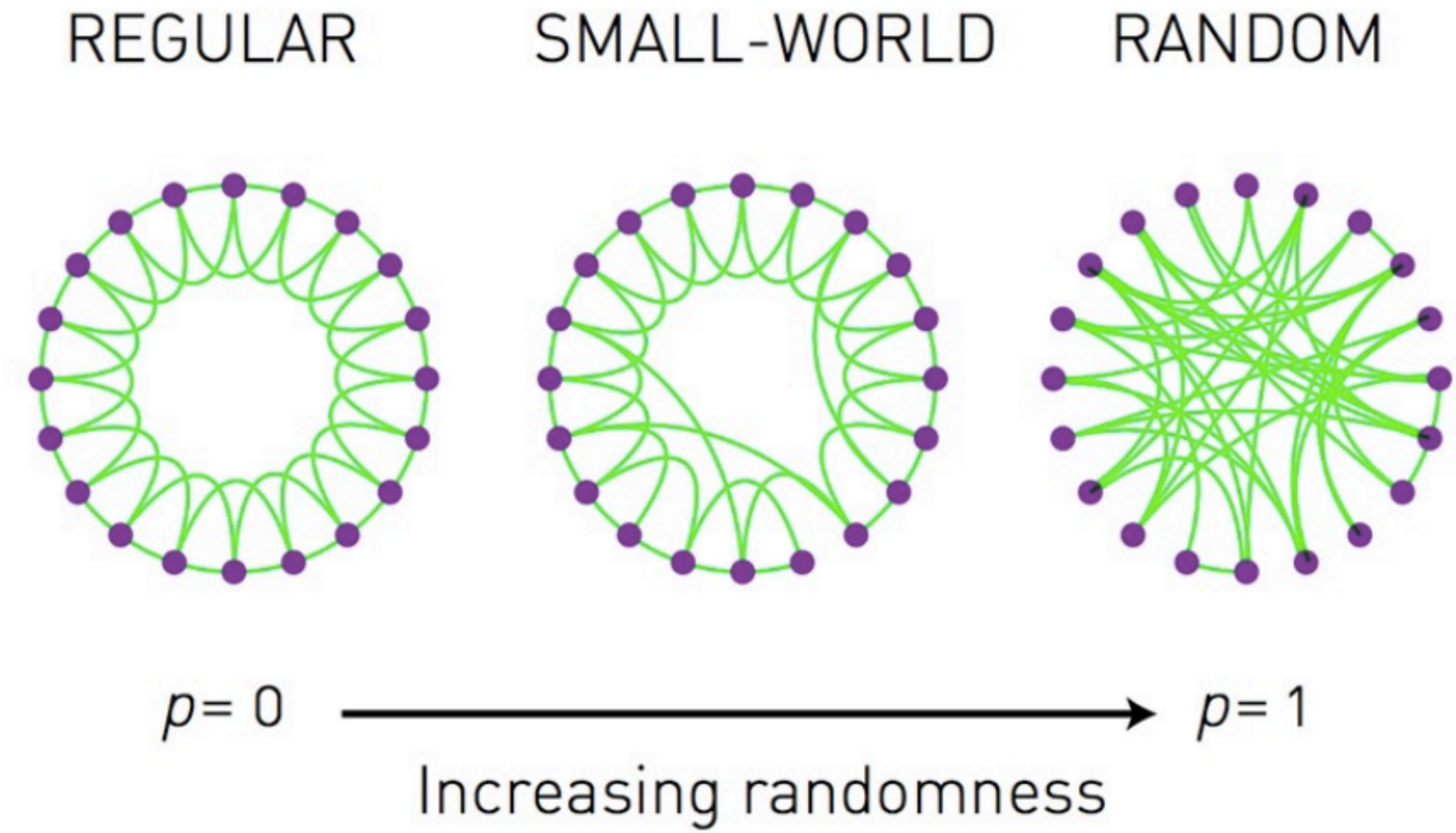
小世界模型算法

Start with a circle network with N nodes, each with 2m neighbors. A WS network is constructed by considering each of the links of the circle network and independently rewiring each of them with probability $p (0 \leq p \leq 1)$ as follows:

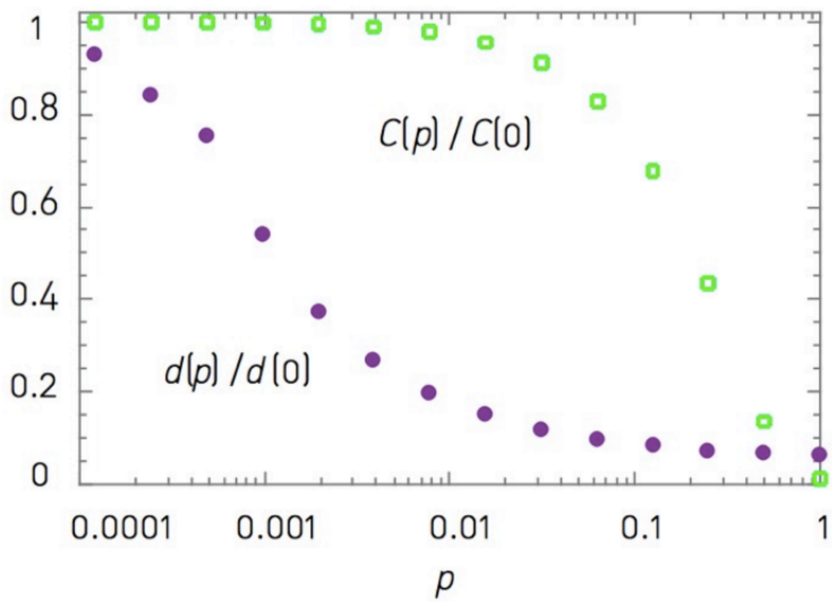
1. Visit each node along the ring one after the other, clockwise.

2. Let i be the current node. Each edge connecting node i to one of its m neighbors clockwise, is considered and rewired with a probability p, or left in place with a probability $(1 - p)$.

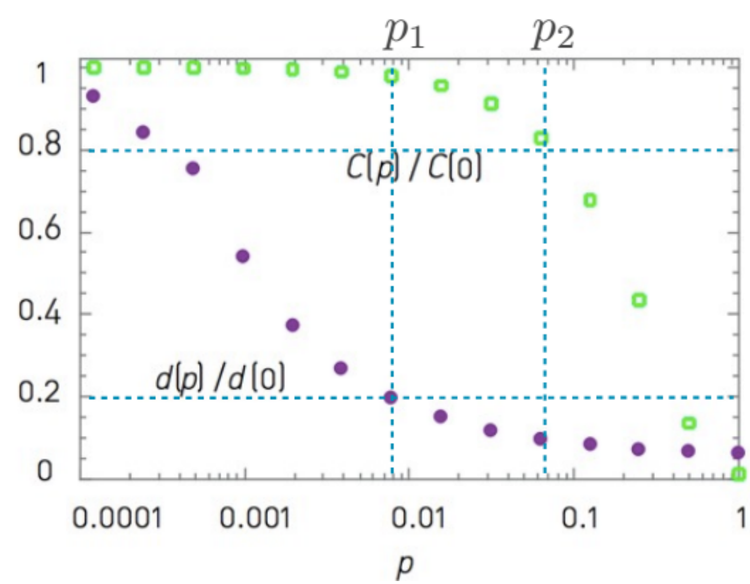
3. Rewiring means shifting the end of the edge, other than that in node i, to a new vertex chosen uniformly at random from the whole lattice, with the constraint than no two vertices can have more than one edge running between them, and no vertices can be connected by an edge to itself.


REGULAR
SMALL-WORLD
RANDOM
p= 0
p= 1
Increasing randomness

小世界性质

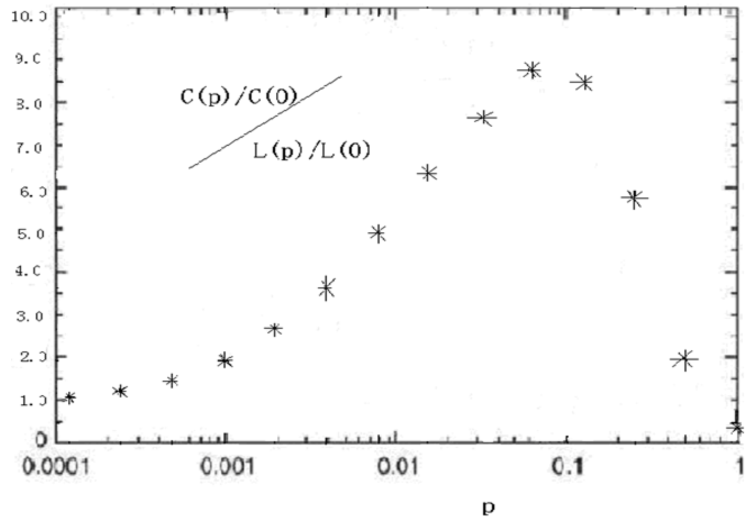


1
0.8
0.6
0.4
0.2
0
C(p) / C(0)
d(p) / d(0)
0.0001
0.001
0.01
0.1
1
p



p_1
p_2
1
0.8
0.6
0.4
0.2
0
C(p) / C(0)
d(p) / d(0)
0.0001
0.001
0.01
0.1
1
p

小世界性质



许多实际网络具有小世界特性

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}
WWW, site level, undir.	153 127	35.21	3.1	3.35	0.1078	0.00023
Internet, domain level	3015–6209	3.52–4.11	3.7–3.76	6.36–6.18	0.18–0.3	0.001
Movie actors	225 226	61	3.65	2.99	0.79	0.00027
LANL co-authorship	52 909	9.7	5.9	4.79	0.43	1.8×10^{-4}
MEDLINE co-authorship	1 520 251	18.1	4.6	4.91	0.066	1.1×10^{-5}
SPIRES co-authorship	56 627	173	4.0	2.12	0.726	0.003
NCSTRL co-authorship	11 994	3.59	9.7	7.34	0.496	3×10^{-4}
Math. co-authorship	70 975	3.9	9.5	8.2	0.59	5.4×10^{-5}
Neurosci. co-authorship	209 293	11.5	6	5.01	0.76	5.5×10^{-5}
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06
Silwood Park food web	154	4.75	3.40	3.23	0.15	0.03
Words, co-occurrence	460.902	70.13	2.67	3.03	0.437	0.0001
Words, synonyms	22 311	13.48	4.5	3.84	0.7	0.0006
Power grid	4941	2.67	18.7	12.4	0.08	0.005
C. Elegans	282	14	2.65	2.25	0.28	0.05

平均距离、集聚系数

作业

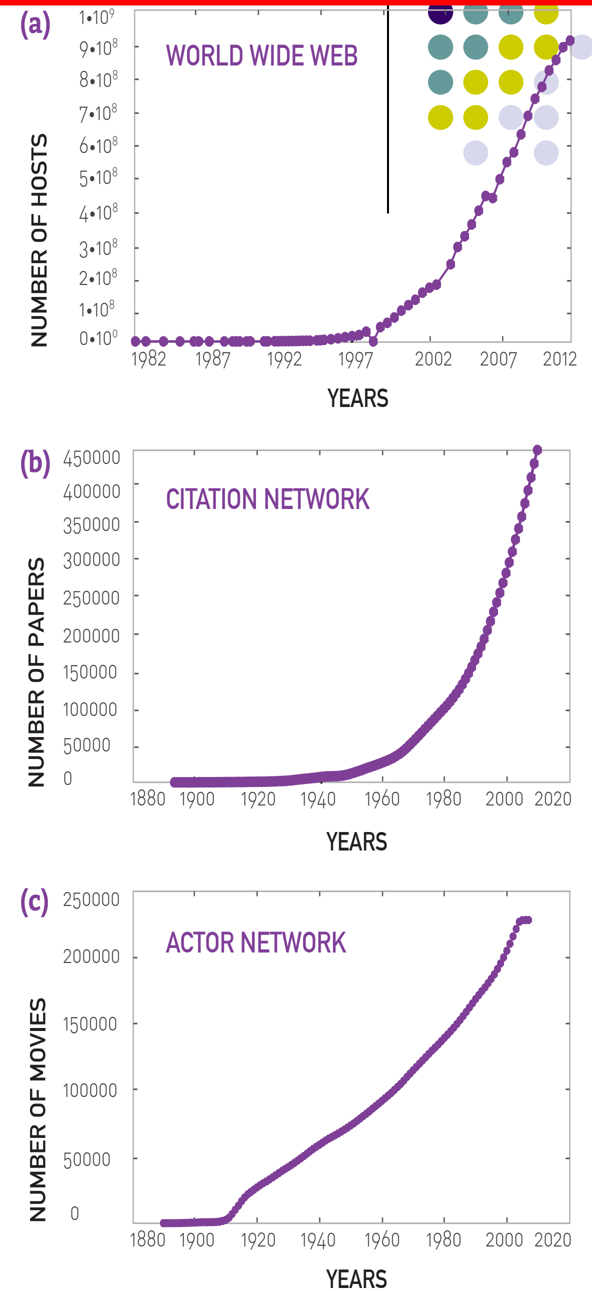
- 阅读WS模型的文献
- 用程序生成WS模型，并且对其小世界特性进行计算。（选做）

无标度网络的经典机制模型—BA模型

结点数目的增长

ER model:
结点数目固定

实际网络的演化是有新**结**点的加入



Barabási & Albert, *Science* **286,** 509 (1999)

偏好连接

ER model:随机加边

新**结**点更**倾**向于和**连**接多的**结**点**进**行**连**接

Barabási & Albert, *Science* **286,** 509 (1999)

Scale-free model

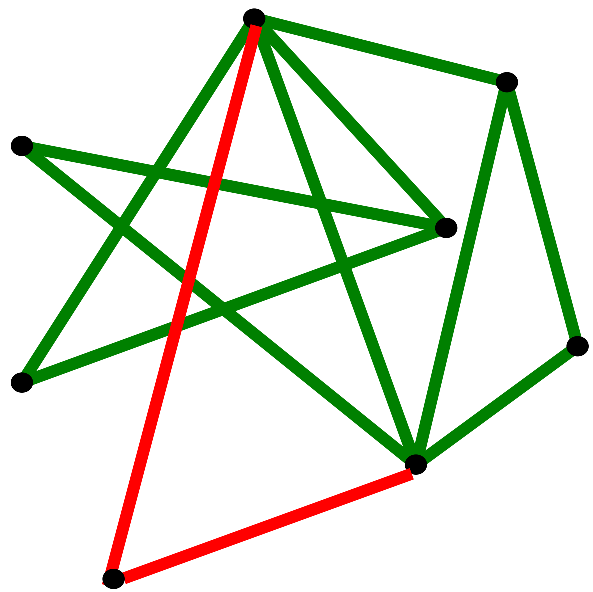
(1) GROWTH :

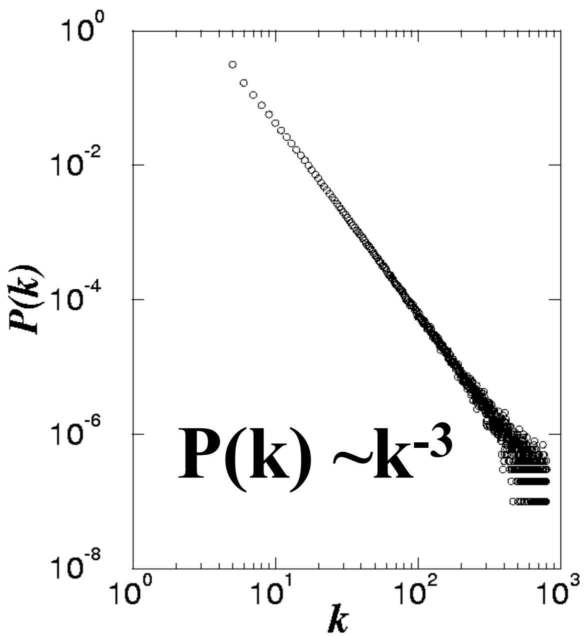
At every timestep we add a new node with *m* edges (connected to the nodes already present in the system).

(2) PREFERENTIAL ATTACHMENT :

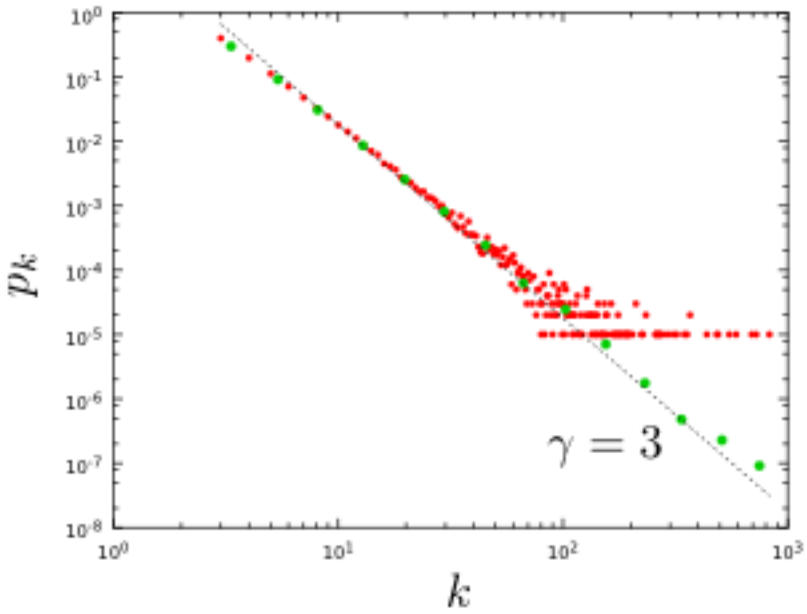
The probability **Π** that a new node will be connected to node *i* depends on the connectivity k_i of that node

$$\Pi(k_i) = \frac{k_i}{\Sigma_j k_j}$$





A.-L.Barabási, R. Albert, Science **286,** 509 (1999)



size N=100,000 and m=3
linearly-binned (red)
log-binned (green).

- 初始m0个节点的连接方式?
- m条边如何加入？ （一起加入、还是依次加入）

所有的结点服从一样的增长规则

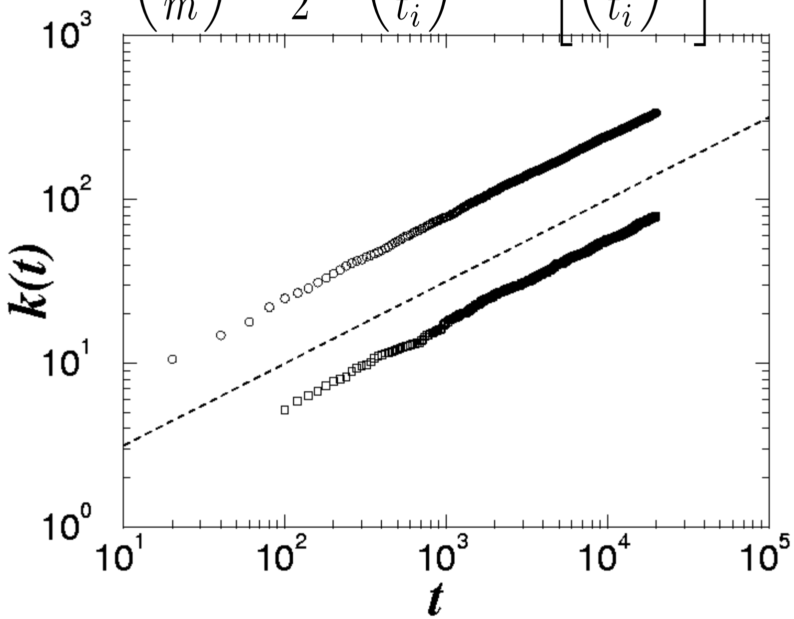
$$\frac{\partial k_i}{\partial t} \propto \Pi(k_i) = A\frac{k_i}{\sum_j k_j}$$

Use: $\sum_j k_j = 2mt$ During a unit time (time step): *Δk=m* → *A=m*

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t} \qquad \frac{\partial k_i}{k_i} = \frac{\partial t}{2t} \qquad \int_m^k \frac{\partial k_i}{k_i} = \int_{t_i}^t \frac{\partial t}{2t}$$

$$\ln\left(\frac{k}{m}\right) = \frac{1}{2}\ln\left(\frac{t}{t_i}\right) = \ln\left[\left(\frac{t}{t_i}\right)^{\frac{1}{2}}\right]$$

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \qquad \beta = \frac{1}{2}$$



β: dynamical exponent

A.-L.Barabási, R. Albert and H. Jeong, *Physica* A **272,** 173 (1999)

度分布（平均场理论）

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \qquad \beta = \frac{1}{2}$$

A node *i* can come with equal probability any time between $t_i=m_0$ and *t*, hence:

$$P(t_i) = \frac{1}{m_0 + t} \qquad P(t_i < \tau) = \frac{1}{m_0 + t}\int_0^{\tau} dt_i = \frac{\tau}{m_0 + t}$$

$$P(k) = P\left(t_i \le \frac{m^{1/\beta} t}{k^{1/\beta}}\right) = 1 - \frac{m^{1/\beta} t}{k^{1/\beta}(t + m_0)}$$

$$\therefore P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^2 t}{m_o + t}\frac{1}{k^3} \sim k^{-\gamma}$$

$$\gamma = 3$$

A.-L.Barabási, R. Albert and H. Jeong, *Physica* A **272,** 173 (1999)

速率方程

*t*时刻度值为*k*的顶点组成大小为$N_k(t)$的集团

新顶点进入时，变化率为：

$$\frac{dN_k(t)}{dt} = mW^{get}(k-1) - mW^{get}(k) + \delta_{km}$$

其中：$W^{get}(k) = \frac{kN_k(t)}{\sum_d dN_d(t)}$ 且有：$\sum_k kN_k(t) = 2mt$

度分布：$P(k) = \lim_{t\to\infty} \frac{N_k(t)}{\sum_d N_d(t)}$

Master方程

- 分布函数：$p(k;t_i,t)$　初始条件：$p(k;t_i,t)=\delta(k-m)$

$$p(k;t_i,t+1)=p(k;t_i,t)+w^{get}(k-1)p(k-1;t_i,t)$$
$$-w^{get}(k)p(k;t_i,t)$$

其中：　$w^{get}(k)=k/2t$

度分布：$P(k)=\lim_{t\to\infty}\frac{\sum_{t_i}p(k;t_i,t)}{t}=\frac{2m(m+1)}{k(k+1)(k+2)}\sim k^{-3}$

增长和偏好连接缺一不可

Model A

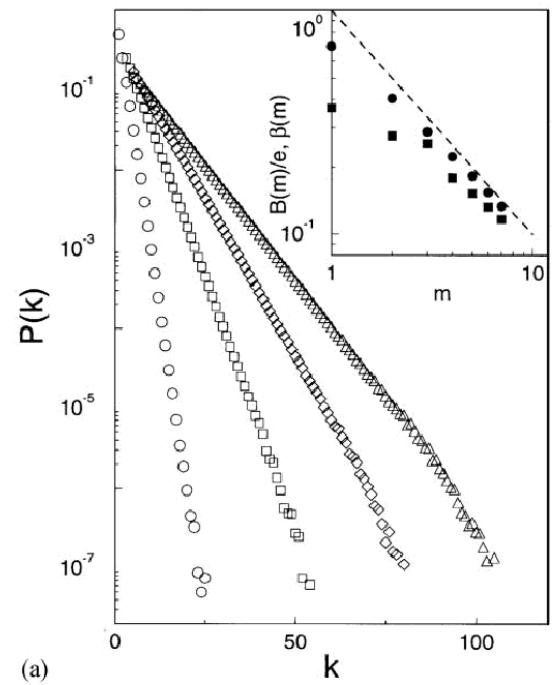
growth ~~preferential attachment~~

$$\Pi(k_i) : \text{uniform}$$

$$\frac{\partial k_i}{\partial t} = A\Pi(k_i) = \frac{m}{m_0 + t - 1}$$

$$k_i(t) = m\left(\ln(\frac{m_0 + t - 1}{m + t_i - 1}) + 1 \right)$$

$$P(k) = \frac{e}{m}\exp(-\frac{k}{m}) \sim e^{-k}$$



Model B

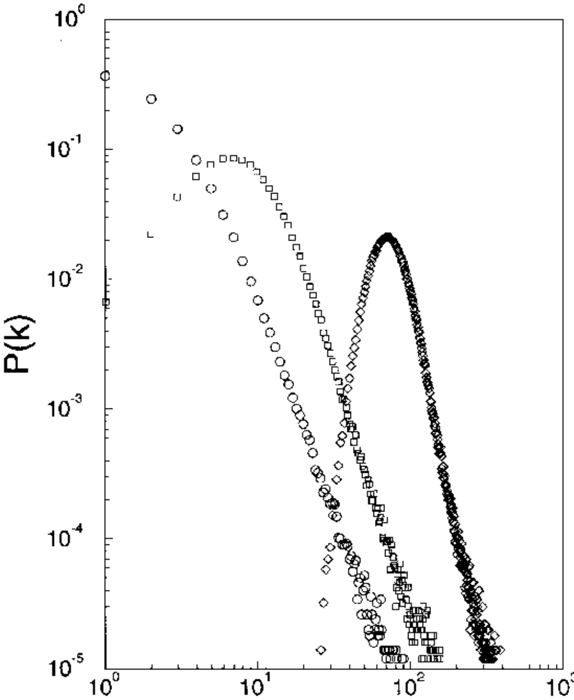
~~growth~~	preferential attachment

$$\frac{\partial k_i}{\partial t} = A\Pi(k_i) + \frac{1}{N} = \frac{N}{N-1}\frac{k_i}{2t} + \frac{1}{N}$$

$$k_i(t) = \frac{2(N-1)}{N(N-2)}t + Ct^{\frac{N}{2(N-1)}} \sim \frac{2}{N}t$$

P(k) : power law (initially)

$\Rightarrow$ Gaussian



(a)

作业

- 阅读BA模型的文献
- 程序生成BA模型，拟合度分布。（选做）

问题与改进

- 指数γ＝3且与参数m，m_0无关

  Krapivsky等2000：考虑顶点历史

  Dorogovtsev等2000：m条边不从新顶点出发

  A-B2000:加点，加边，重联

  Cooper等2001：加边，加点
- 度度相关性为零——与实证相符的匹配模式
- In－Out幂律分布

加权网络的演化模型

- 边权固定模型
  - Yook-Jeong-Barabási-Tu (YJBT)， Phys. Rev. Lett. 86, 5835 (2001).
  - Zheng-Trimper-Zheng-Hui (ZTZH)， Phys. Rev. E 67, 040102 (2003)
  - Antal-Krapivsky (AK)， Phys. Rev. E 71, 026103 (2005)
- 边权演化模型
  - Barrat-Barthélemy-Vespignani (BBV)， Phys. Rev. Lett. 92, 228701 (2004)
  - Dorogovtsev-Mendes（DM）， arXiv: cond-mat/0408343
  - 中科大流驱动模型，Phys Rev Lett. 94. 188702

基于科学家合作网的加权网络演化模型

$$w_{ij} = f(T_{ij})$$

$$\prod_{n\to i} = (1-p)\frac{k_i}{\sum_j k_j} + (p-\delta)\frac{s_i}{\sum_j s_j} + \delta\frac{l_{ni}}{\sum_{j\in\partial_n^d} l_{nj}}$$

$$T_{ni^*}(t+1) = T_{ni^*}(t) + 1 \qquad w_{ni^*}(t+1) = f(T_{ni^*}(t+1))$$