Mathematical Models of Networks

CS5128701:
Practice of Social Media Analytics

NTUST CSIE
Yi-Ling Chen

Network models

Empirical network features:

- Power-law (heavy-tailed) degree distribution
- Small average distance (graph diameter)
- Large clustering coefficient (transitivity)
- Giant connected component

Generative models:

- Random graph model (Erdos & Renyi, 1959)
- Preferential attachement model (Barabasi & Albert, 1999)
- Small world model (Watts & Strogatz, 1998)

Random graph model

Graph $G\{E, V\}$, nodes $n = |V|$, edges $m = |E|$
Erdos and Renyi, 1959.

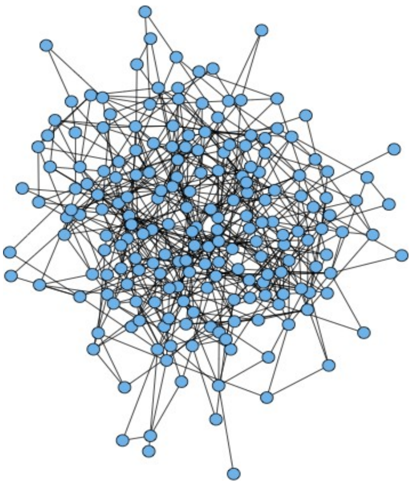
- $G_{n,p}$ - each pair out of $N = \frac{n(n-1)}{2}$ is connected with probability p, number of edges m - random number

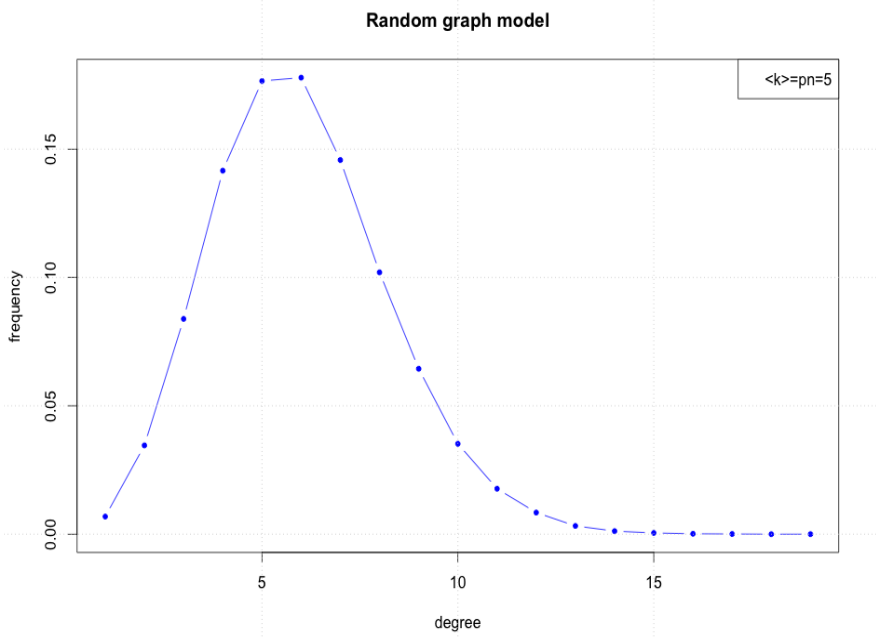
$$\langle m \rangle = p\frac{n(n-1)}{2}$$

$$\langle k \rangle = \frac{1}{n}\sum_i k_i = \frac{2\langle m \rangle}{n} = p\,(n-1) \approx pn$$

$$\rho = \frac{\langle m \rangle}{n(n-1)/2} = p$$

Random graph





Node degree distribution (Poisson distribution):

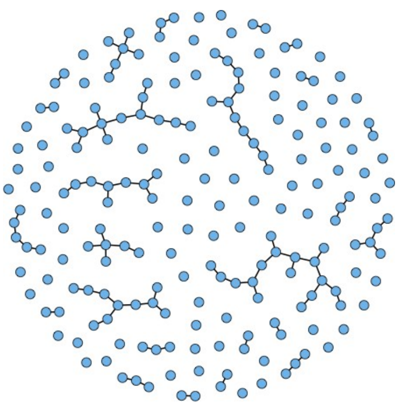
$$P(k_i = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \qquad \lambda = pn = \langle k \rangle$$

Phase transition

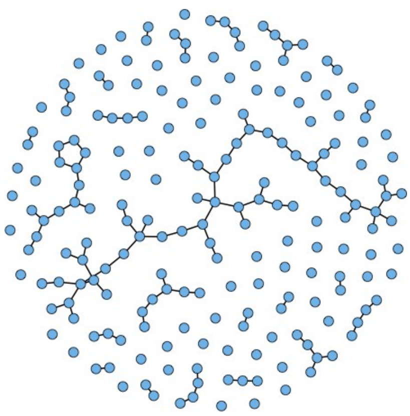
Consider $G_{n,p}$ as a function of p

- $p = 0$, empty graph
- $p = 1$, complete (full) graph
- There is a critical probability p_c, where structural changes from $p < p_c$ to $p > p_c$
- Gigantic connected component appears at $p > p_c$

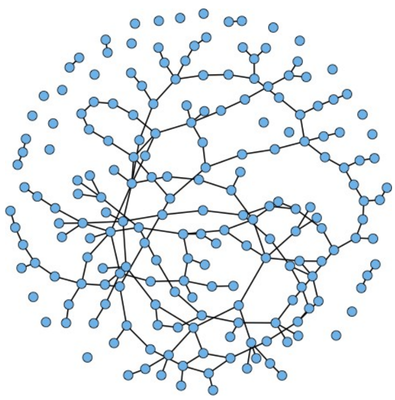
Phase transition



$p < p_c$



$p = p_c$



$p > p_c$

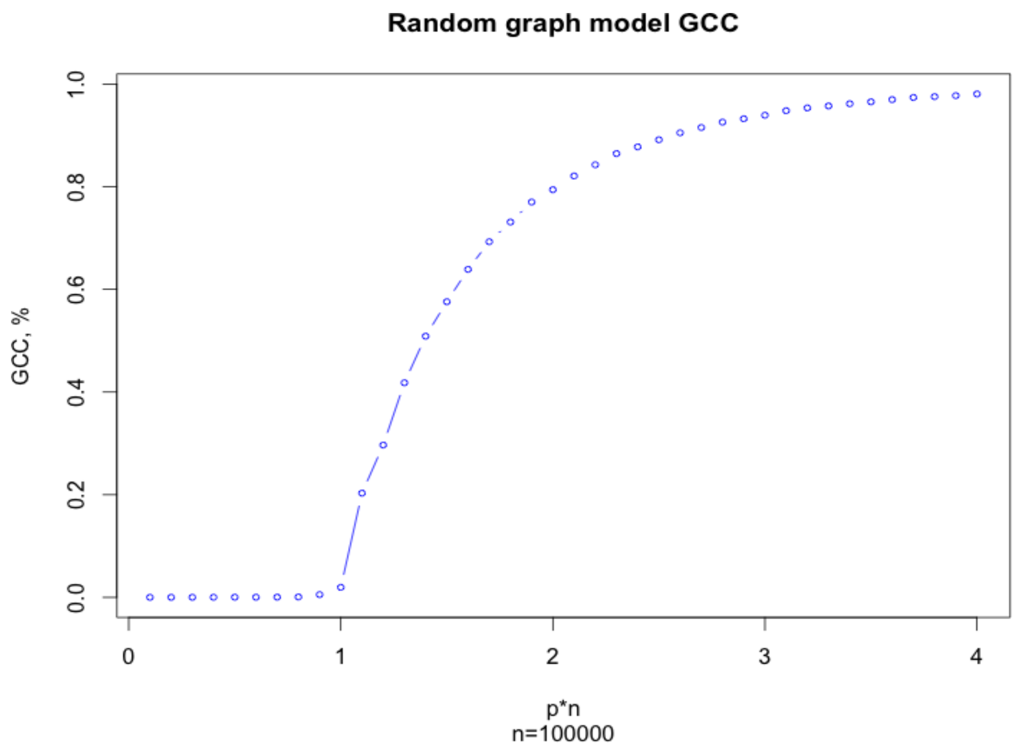
Phase transition

Graph $G(n, p)$, for $n \to \infty$, critical value $p_c = 1/n$

- when $p < p_c$, $(\langle k \rangle < 1)$ there is no components with more than $O(\ln n)$ nodes, largest component is a tree
- when $p = p_c$, $(\langle k \rangle = 1)$ the largest component has $O(n^{2/3})$ nodes
- when $p > p_c$, $(\langle k \rangle > 1)$ gigantic component has all $O(n)$ nodes

Critical value: $\langle k \rangle = p_c n = 1$- on average one neighbor for a node

Connected component



$$\langle k \rangle = pn$$

Clustering coefficient

- Clustering coefficient

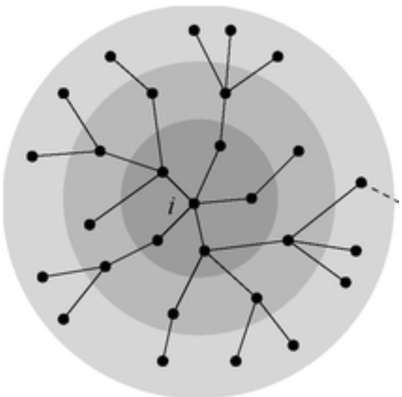
$$C(k) = \frac{\#\text{of links between NN}}{\#\text{max number of links NN}} = \frac{pk(k-1)/2}{k(k-1)/2} = p$$

$$C = p = \frac{\langle k \rangle}{n}$$

- when $n \to \infty, \quad C \to 0$

Graph diameter

- $G(n, p)$ is locally tree-like (GCC) (no loops; low clustering coefficient)



- on average, the number of nodes d steps away from a node $\langle k \rangle^d$
- in GCC, around p_c , $\langle k \rangle^d \sim n$,

$$d \sim \frac{\ln n}{\ln \langle k \rangle}$$

Random graph model

- Node degree distribution function - Poisson:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad \lambda = pn = \langle k \rangle$$

- Average path length:

$$\langle L \rangle \sim \log(N)/\log\langle k \rangle$$

- Clustering coefficient:

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

Preferential attachment model

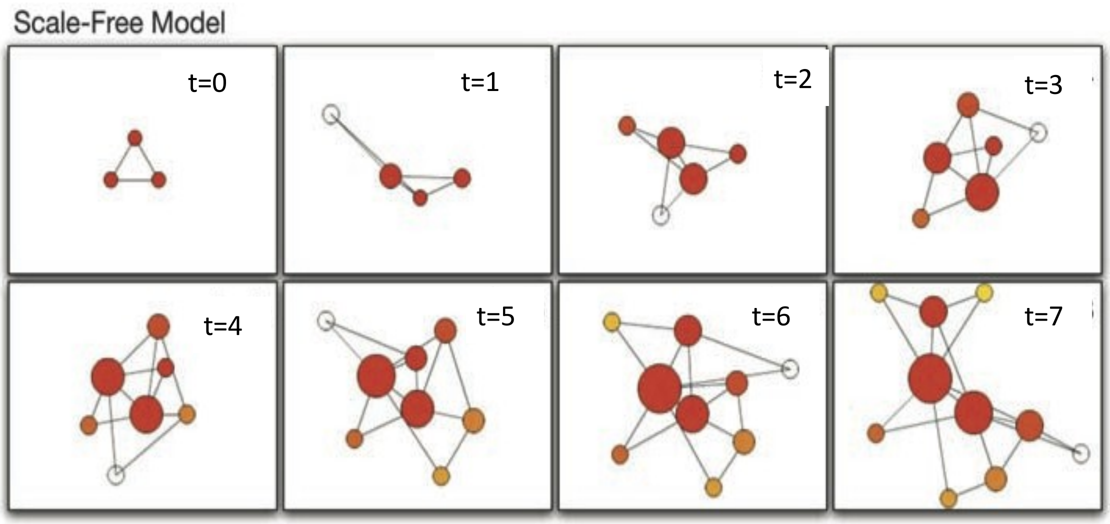
Barabasi and Albert, 1999
Dynamical growth model

- $t = 0$, n_0 nodes
- growth: on every step add a node with m_0 edges ($m_0 \leq n_0$), $k_i(i) = m_0$
- Preferential attachment: probability of linking to existing node is proportional to the node degree k_i

$$\Pi(k_i) = \frac{k_i}{\sum_i k_i} = \frac{k_i}{2m_0 t}$$

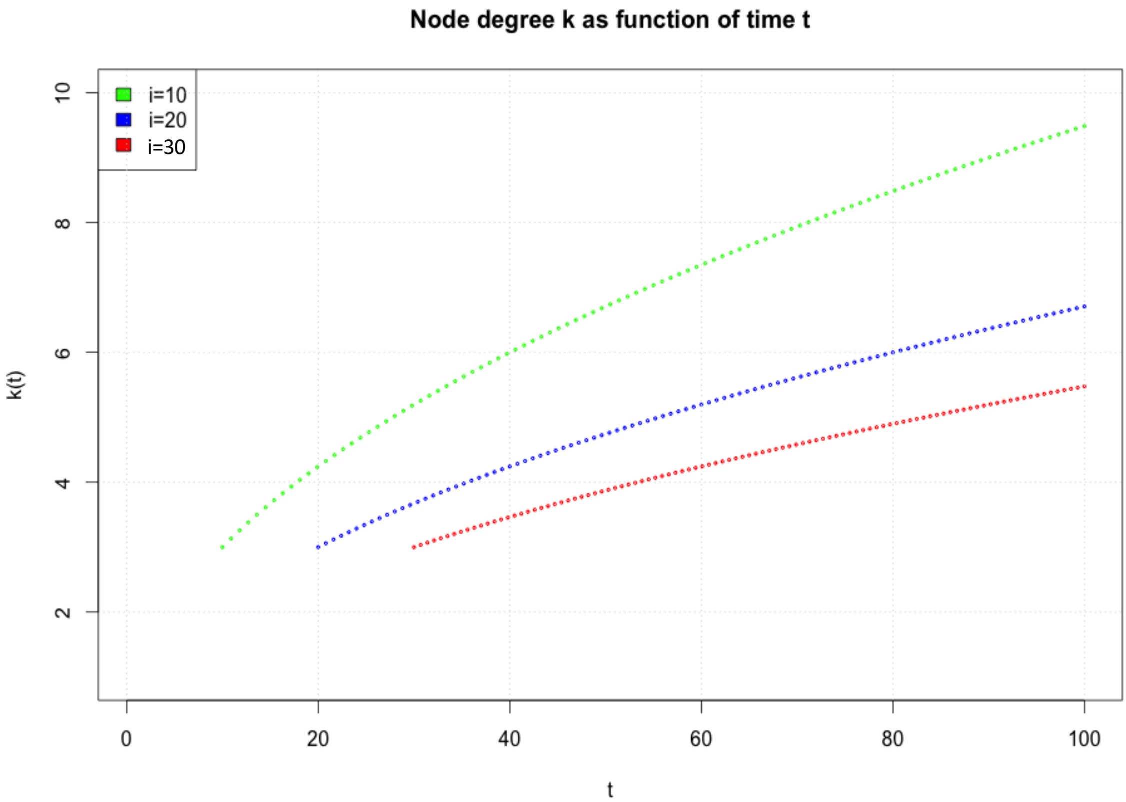
after t steps: $n_0 + t$ nodes, $m_0 t$ edges

Preferential attachment model

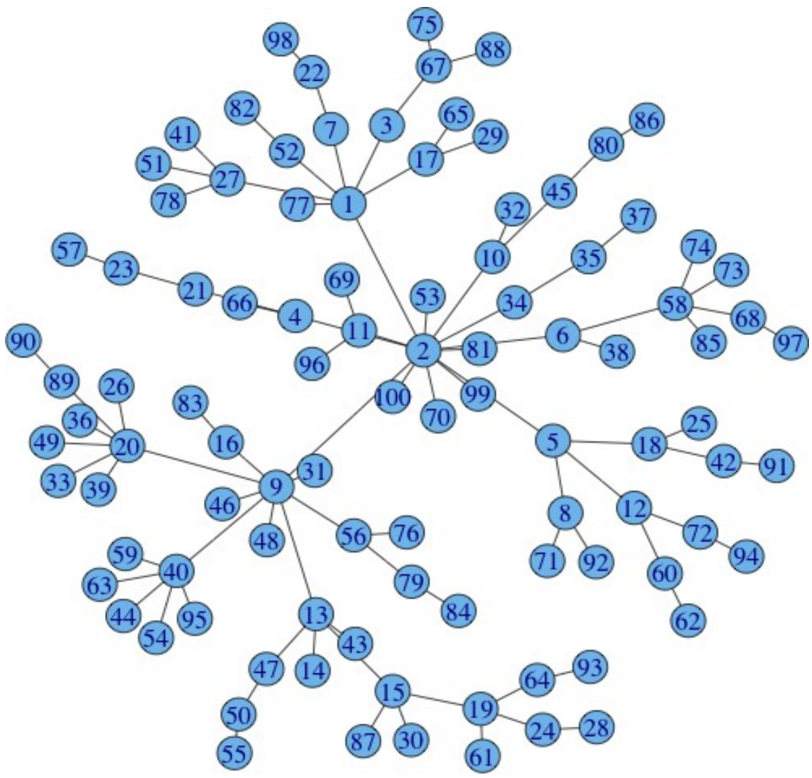


Barabasi, 1999

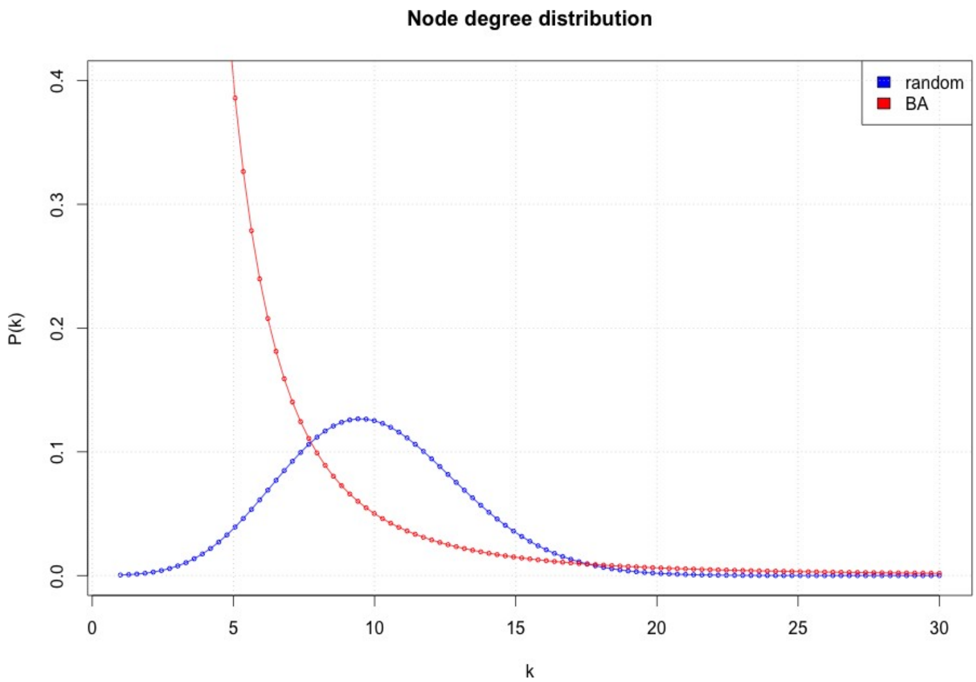
Preferential attachment



Preferential attachment



Preferential attachment

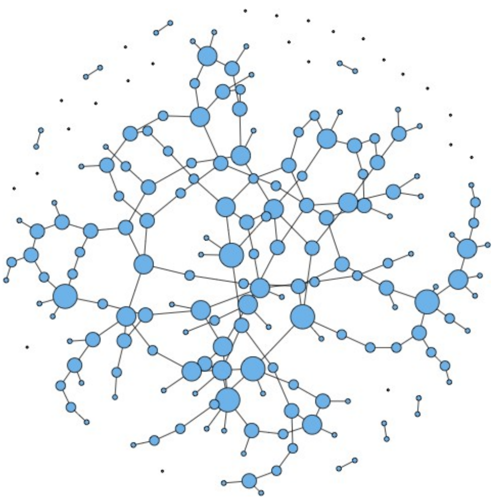


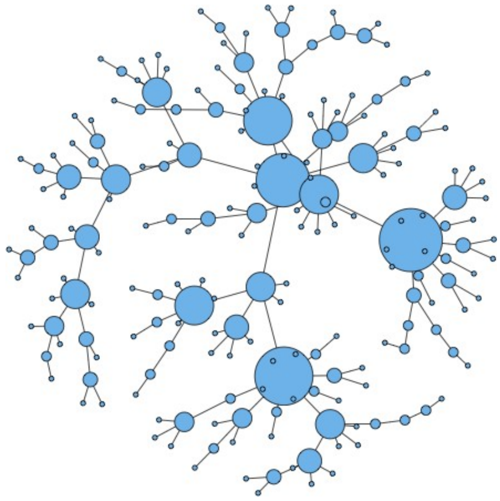
Node degree distribution:

$$P(k_i = k) = \frac{2m_0^2}{k^3}$$

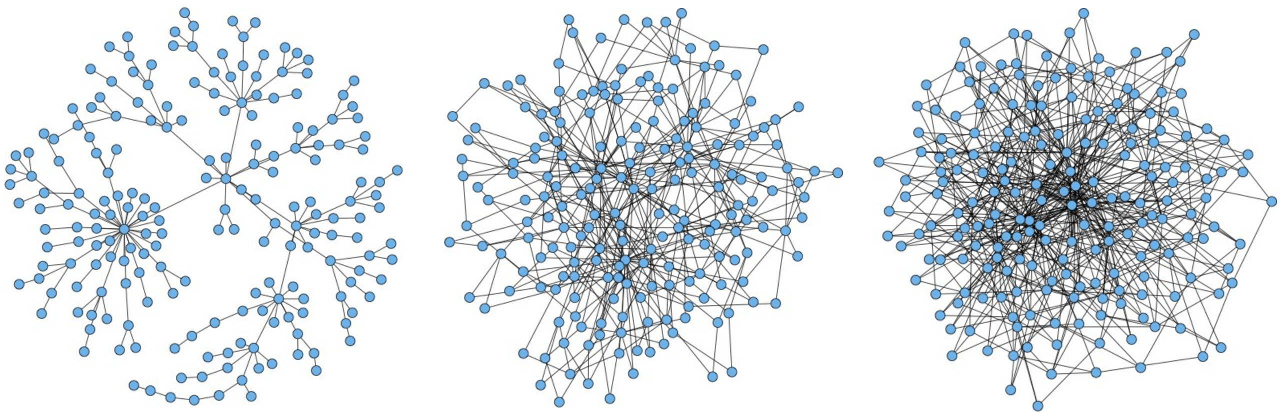
Preferential attachment

Preferential attachment vs. random graph





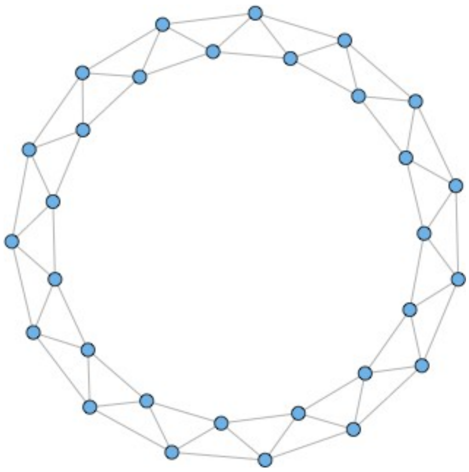
Preferential attachment model



$m_0 = 1$ $m_0 = 2$ $m_0 = 3$

Small world

Motivation: keep high clustering, get small diameter



(O) Clustering coefficient $C = 1/2$

(X) Graph diameter $d = 8$

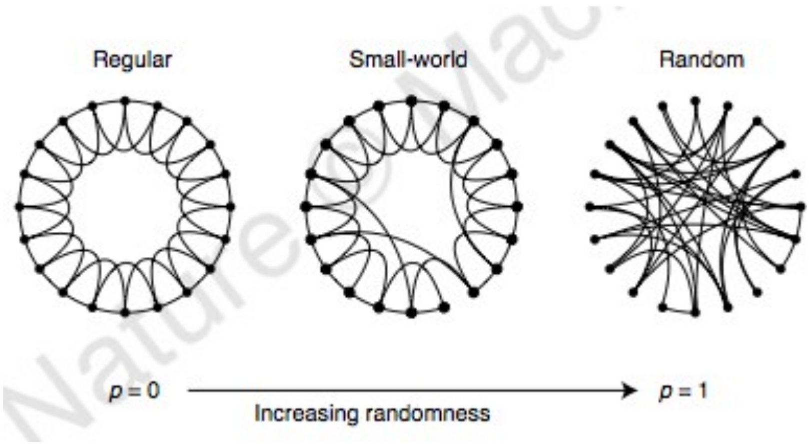
Small world

Watts and Strogatz, 1998

Single parameter model, interpolation between regular lattice and random graph

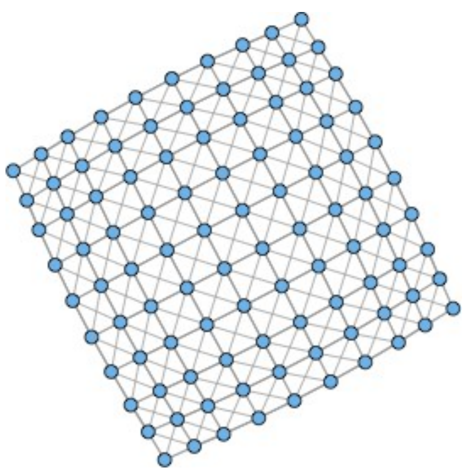
- start with regular lattice with n nodes, k edges per vertex (node degree), $k << n$
- randomly connect with other nodes with probability p, forms $pnk/2$ "long distance" connections from total of $nk/2$ edges
- $p = 0$ regular lattice, $p = 1$ random graph

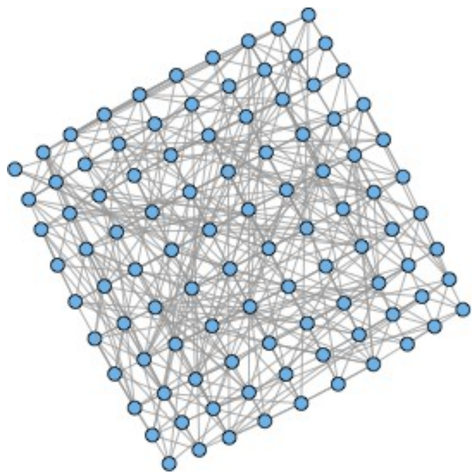
Small world



Watts, 1998

Small world model





20% rewiring:

ave. path length = 3.58 → ave. path length = 2.32

clust. coeff = 0.49 → clust. coeff = 0.19