Network models: dynamic growth and small world

Yury Dvorkin

New York University

Date of Presentation

Empirical network features:

- Power-law (heavy-tailed) degree distribution (see the previous slide-deck)
- Small average distance (graph diameter)
- Large clustering coefficient (transitivity)
- Giant connected component, hierarchical structure,etc

Generative models:

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

Most of the networks we study are evolving over time, they expand by adding new nodes:

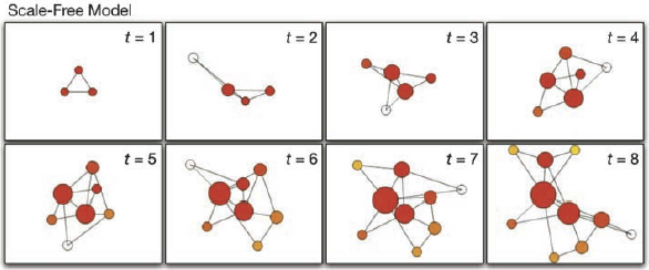
- Citation networks
- Collaboration networks
- Web
- Social networks

Barabasi and Albert, 1999
Dynamic growth model
Start at $t = 0$ with $n0$ nodes and some edges $m_0 \geq n_0$

- 1. **Growth**
  At each time step add a new node with m edges ($m \leq n_0$), connecting to m nodes already in network $k_i(i) = m$
- 2. **Preferential attachment**
  The probability of linking to existing node i is proportional to the node degree k_i

$$\prod(k_i) = \frac{k_i}{\sum_i k_i}$$

after t time-steps: $t + n_0$ nodes, $mt + m_0$ edges



Barabasi, 1999

Continues approximation: continues time, real variable node degree
$\langle k_i(t)\rangle$- expected value over multiple realizations
Time-dependent degree of a single node:

$$k_i(t+\delta t) = k_i(t) + m\prod(k_i)\delta t$$

$$\frac{dk_i(t)}{dt} = m\prod(k_i) = m\frac{k_i}{\sum_i k_i} = \frac{mk_i}{2mt}$$

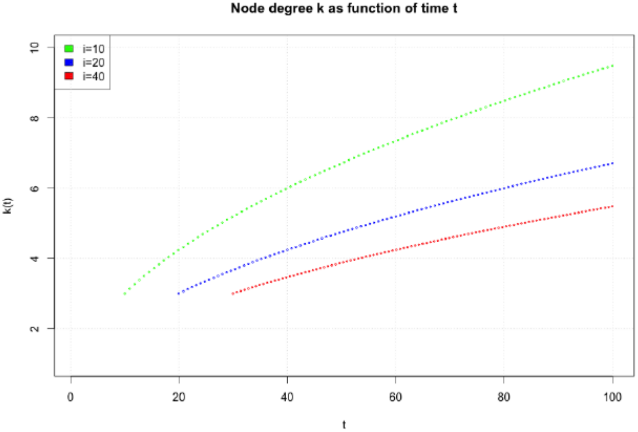
$$\frac{dk_i(t)}{dt} = \frac{k_i(t)}{2t}$$

initial conditions: $k_i(t=i) = m$

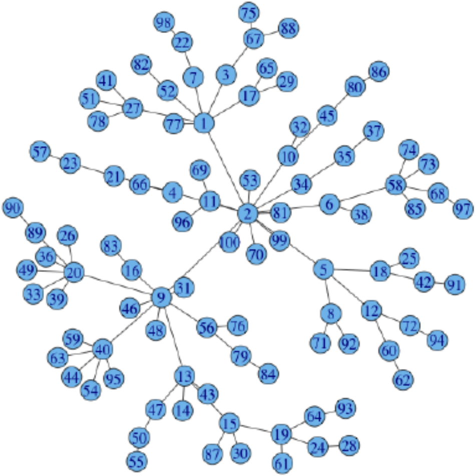
$$\int_m^{k_i(t)} \frac{dk_i}{k_i} = \int_i^t \frac{dt}{2t}$$

Solution:

$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$



$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Time evolution of a node degree

$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$

Nodes with $k_i(t) \leq k$:

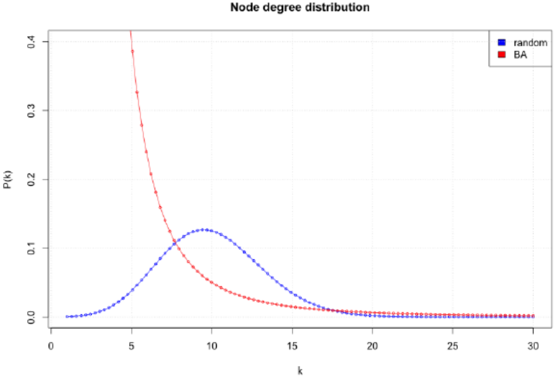
$$m\left(\frac{t}{i}\right)^{1/2} \leq k \qquad i \geq \frac{m^2}{k^2}t$$

Probability of randomly selected node to have $k' \leq k$ (fraction of nodes with $k' \leq k$)

$$F(k) = P(k' \leq k) = \frac{n_0 + t - m^2 t/k^2}{n_0 + t} \approx 1 - \frac{m^2}{k^2}$$

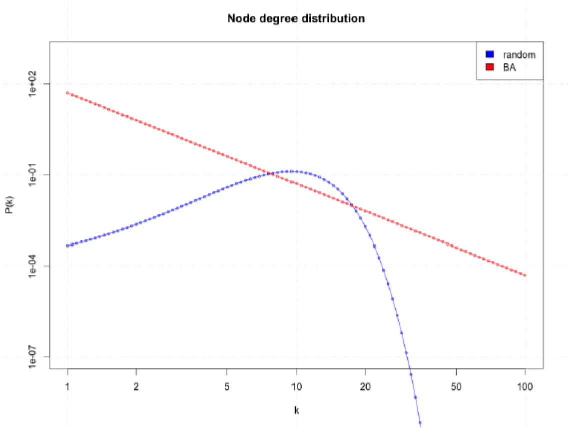
Distribution function:

$$P(k) = \frac{d}{dk}F(k) = \frac{2m^2}{k^3}$$



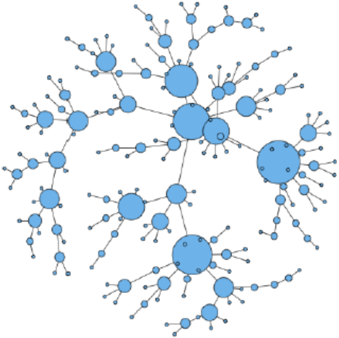
BA:

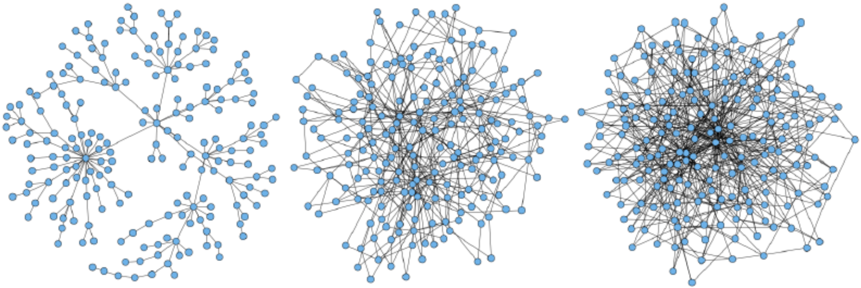
$$P(k) = \frac{2m^2}{k^3}, \quad ER : P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \langle k \rangle = pn$$



BA:

$$P(k) = \frac{2m^2}{k^3}, \quad ER: P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \langle k \rangle = pn$$





$m = 1$ $m = 2$ $m = 3$

- **Growth**
  At each time step add a new node with m edges ($m \leq n_0$), connecting to m nodes already in network $k_i(i) = m$
- **Uniformly at random**
  The probability of linking to existing node i is

$$\prod(k_i) = \frac{1}{n_0 + t - 1}$$

Node degree growth:

$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right)$$

Node degree distribution function:

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

- Power law distribution function:

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

- Average path length (analytical result) :

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

- Clustering coeficient (numerical result):

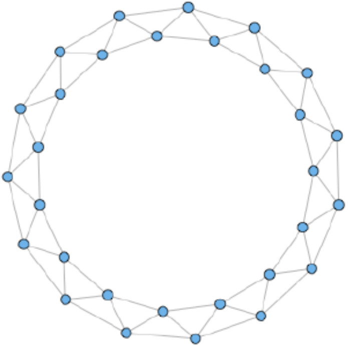
$$C \sim N^{-0.75}$$

Some other models that produce scale-free networks:

- Non-linear preferential attachment
- Link selection model
- Cost-optimization model
- ...

- Polya urn model, George Polya, 1923
- Yule process, Udny Yule, 1925
- Distribution of wealth, Herbert Simon,1955
- Evolution of citation networks, cumulative advantage, Derek de Solla Price, 1976
- Preferential attachment network model, Barabasi and Albert, 1999

Motivation: keep high clustering, get small diameter



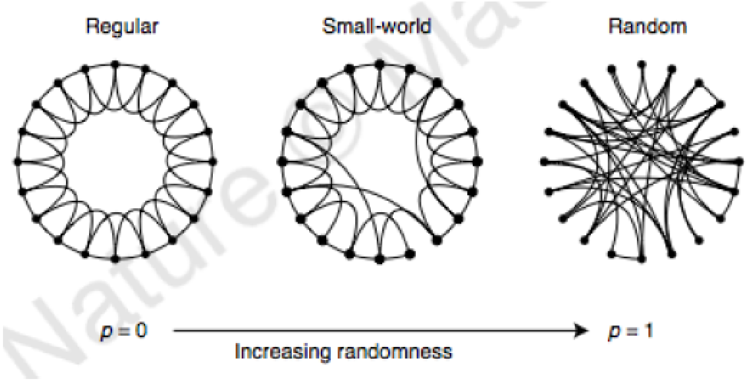
Clustering coefficient $C = 1/2$
Graph diameter $d = 8$

Watts and Strogatz, 1998

Single parameter model, interpolation between regular lattice and random graph

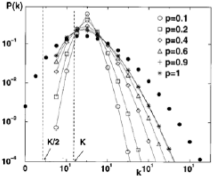
- start with regular lattice with n nodes, k edges per vertex (node degree), $k << n$
- randomly connect with other nodes with probability p, forms $pnk/2$ "long distance" connections from total of $nk/2$ edges
- $p = 0$ regular lattice, $p = 1$ random graph

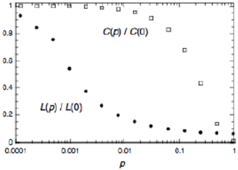
Complementary cumulative distribution function cCDF



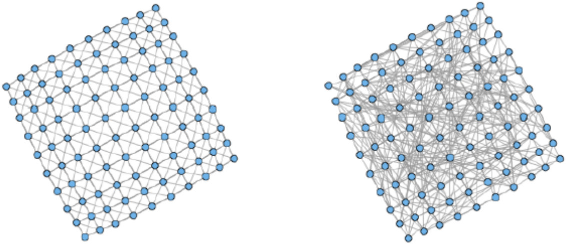
Watts, 1998

- Node degree distribution:
  Poisson like
- Ave. path length $\langle L(p) \rangle$:
  $p \to 0$, ring lattice, $\langle L(0) \rangle = 2n/k$
  $p \to 1$, random graph, $\langle L(1) \rangle = \log(n)/\log(k)$
- Clustering coefficient $C(p)$:
  $p \to 0$, ring lattice, $C(0) = 3/4 = const$
  $p \to 1$, random graph, $C(1) = k/n$





Watts, 1998



20% rewiring:
ave. path length $= 3.58 \rightarrow$ ave. path length $= 2.32$
clust. coeff $= 0.49 \rightarrow$ clust. coeff $= 0.19$

	Random	BA model	WS model	Empirical networks
$P(k)$	$\frac{\lambda^k e^{-\lambda}}{k!}$	k^{-3}	poisson like	power law
C	$\langle k \rangle / N$	$N^{-0.75}$	const	large
$\langle L \rangle$	$\frac{\log(N)}{\log(\langle k \rangle)}$	$\frac{\log(N)}{\log \log(N)}$	$\log(N)$	small

- Emergence of Scaling in Random Networks, A.L. Barabasi and R. Albert, Science 286, 509-512, 1999
- Collective dynamics of small-world networks. Duncan J. Watts and Steven H. Strogatz. Nature 393 (6684): 440-442, 1998