Network models: dynamic growth and small world

Leonid E. Zhukov

School of Data Analysis and Artificial Intelligence
Department of Computer Science
National Research University Higher School of Economics

Network Science

Network models

Empirical network features:

- Power-law (heavy-tailed) degree distribution
- 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)

Motivation

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

- Citation networks
- Collaboration networks
- Web
- Social networks

Preferential attachment model

Barabasi and Albert, 1999

Dynamic growth model

Start at $t = 0$ with n_0 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 netwrok $k_i(i) = m$

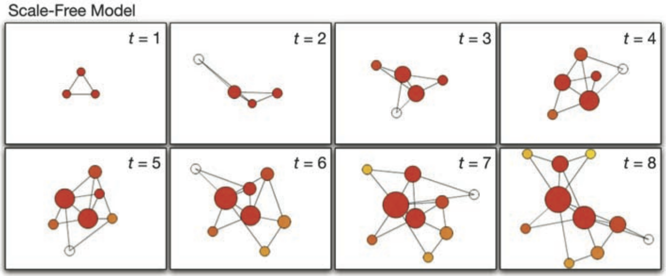
2. **Preferential attachment**

   The probability of linking to existing node i is proportional to the node degree k_i

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

after t timesteps: $t + n_0$ nodes, $mt + m_0$ edges

Preferential attachment model



Barabasi, 1999

Preferential attachment

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:

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

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

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

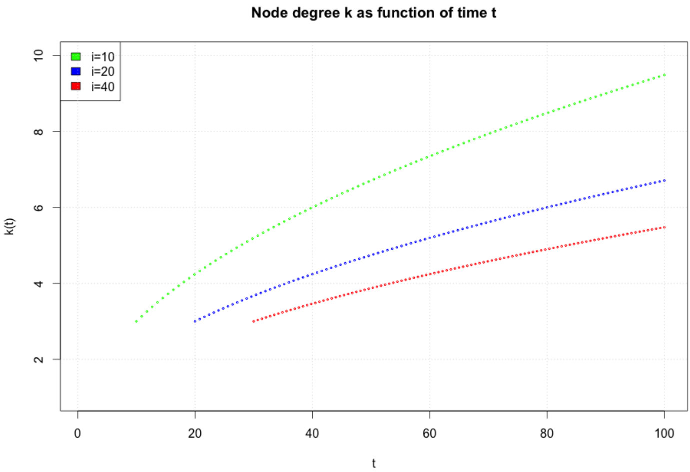
initial conditions: $\quad k_i(t=i) = m$

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

Solution:

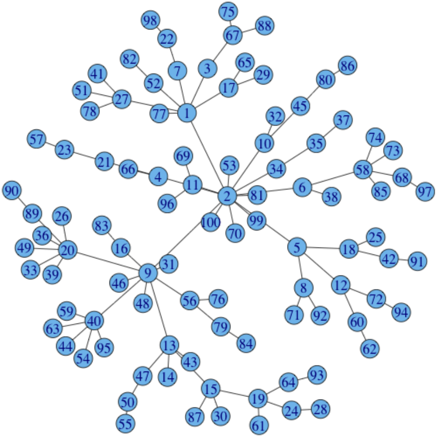
$$k_i(t) = m\left(\frac{t}{i}\right)^{1/2}$$

Preferential attachement



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

Preferential attachement



Preferential attachment

Time evolution of a node degree

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

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

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

$$i \geq \frac{m^2}{k^2} t$$

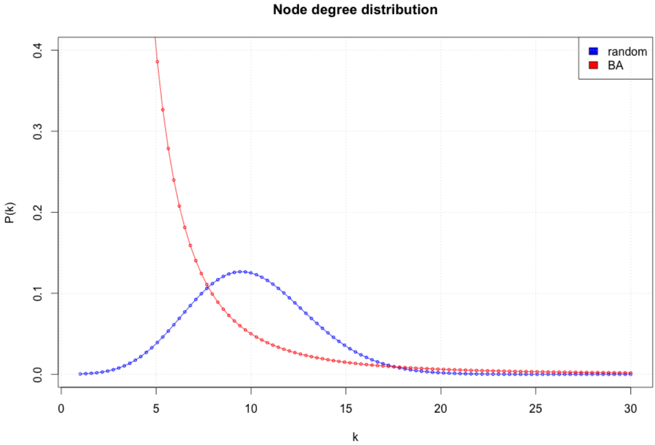
Probability of randomly selected node to have $k' \leq k$ (fraction of nodes with $k' \leq k$)

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

Distribution function:

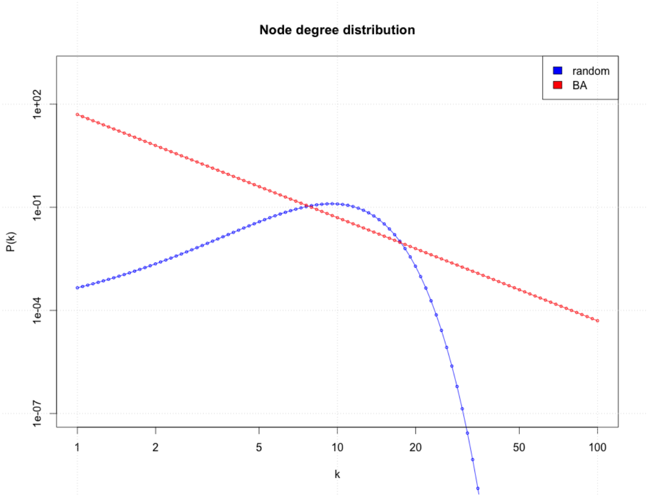
$$P(k) = \frac{d}{dk} F(k) = \frac{2m^2}{k^3}$$

Preferential attachment vs random graph



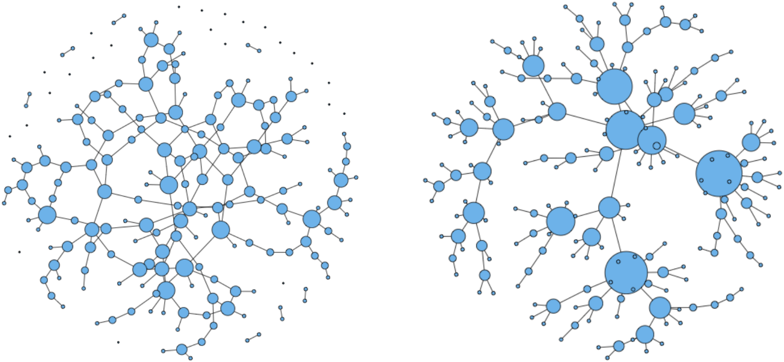
$$BA : P(k) = \frac{2m^2}{k^3}, \qquad ER : P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad \langle k \rangle = pn$$

Preferential attachment vs random graph

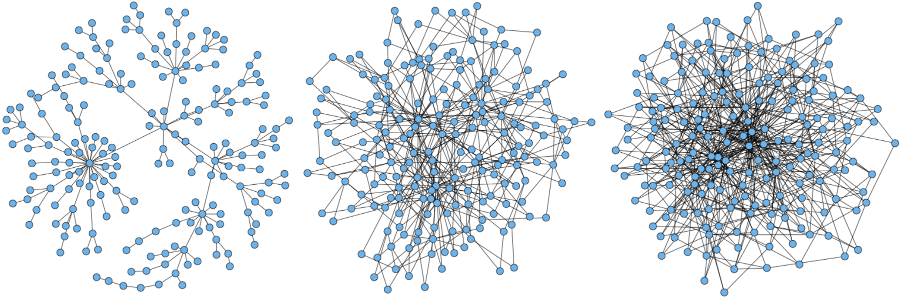


$$BA : P(k) = \frac{2m^2}{k^3}, \qquad ER : P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad \langle k \rangle = pn$$

Preferential attachment vs random graph



Preferential attachment model



$m = 1$ $\qquad$ $m = 2$ $\qquad$ $m = 3$

Growing random graph

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**~~ **Uniformly at random**
   The probability of linking to existing node i is

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

Node degree growth:

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

Node degree distribution function:

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

Preferential attachment

- 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 coefficient (numerical result):

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

Many more models

Some other models that produce scale-free networks:

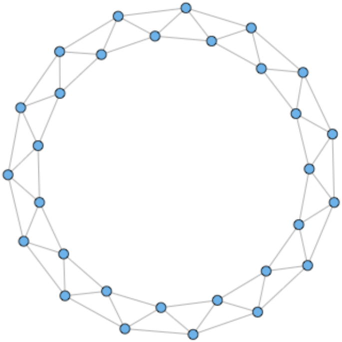
- Non-linear preferential attachment
- Link selection model
- Copying model
- Cost-optimization model
- ...

Historical note

- 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

Small world

Motivation: keep high clustering, get small diameter



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

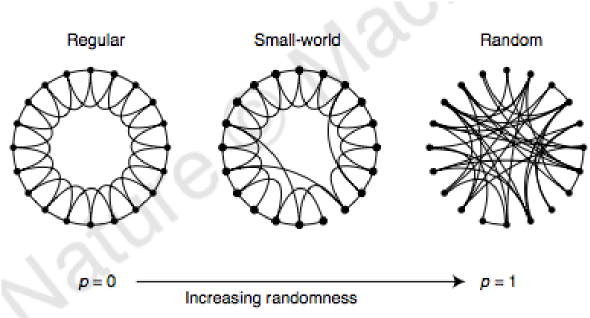
Small world

Watts and Strogatz, 1998

Single parameter model, interpolation between regular lattice and random graph

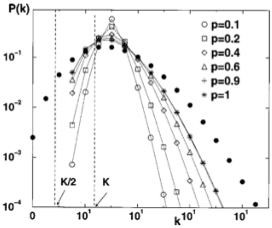
- start with regular lattice with n nodes, k edges per vertex (node degree), $k << n$
- randomly connect with other nodes with probability p, forms $pnk/2$ "long distance" connections from total of $nk/2$ edges
- $p = 0$ regular lattice, $p = 1$ random graph

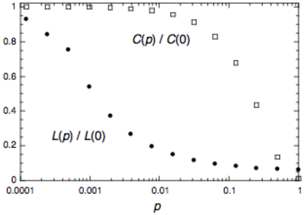
Small world



Watts, 1998

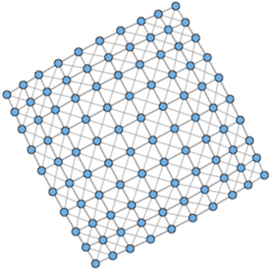
- Node degree distribution:
  Poisson like
- Ave. path length $\langle L(p) \rangle$:
  $p \to 0$, ring lattice, $\langle L(0) \rangle = 2n/k$
  $p \to 1$, random graph, $\langle L(1) \rangle = \log(n)/\log(k)$
- Clustering coefficient $C(p)$:
  $p \to 0$, ring lattice, $C(0) = 3/4 = const$
  $p \to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

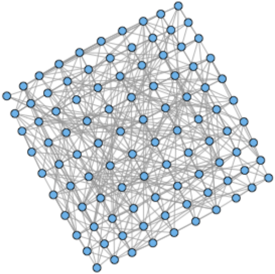




Watts, 1998

Small world model





20% rewiring:

ave. path length $= 3.58 \quad \rightarrow \quad$ ave. path length $= 2.32$

clust. coeff $= 0.49 \quad \rightarrow \quad$ clust. coeff $= 0.19$

Model comparison

	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

References

- 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