Network models: dynamical 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

Structural Analysis and Visualization of Networks

NATIONAL RESEARCH
UNIVERSITY

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 attachement model (Barabasi & Albert, 1999)
- Small world model (Watts & Strogatz, 1998)

Motivation

Growing (evolving with time) networks:

- Citation networks
- Collaboration networks
- Web
- Social networks

Growing random graph

Stochastic growth model:

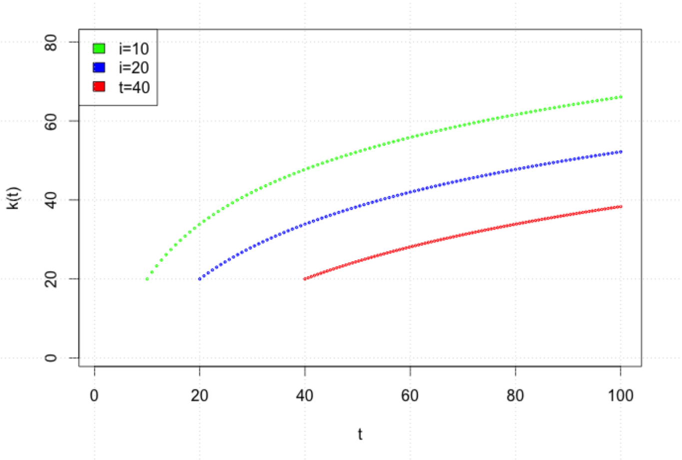
- $t = 0$, n_0 unconnected nodes
- growth: on every time step $t = \{1, 2, 3, 4...\}$ add a node with $m \le n_0$ edges $k_i(t = i) = m$
- attachment: form m edges with existing nodes uniformly at random, $\Pi(k_i) = \frac{1}{n_0+t-1}$

Expected i-node degree at t $\langle k_i(t) \rangle$:

$$k_i(t) = m + \frac{m}{n_0 + i - 1} + \frac{m}{n_0 + i} + .. \frac{m}{n_0 + t - 1}$$

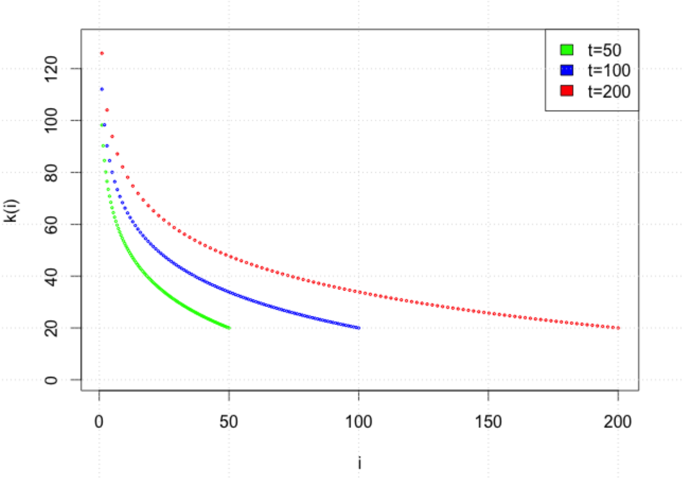
$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right)$$

Growing random graph



$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right), \quad m = 20, \ i = 10, 20, 40, \ t \geq i$$

Growing random graph



$$k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right), \quad m = 20, \ t = 50, 100, 200, \ i \leq t$$

Growing random graph

Find all nodes that at time t has degree less than k, $k_i(t) \le k$?
(for example $k_i(t) \le 40$)

$$
\begin{aligned}
k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right) &\le k \\
\log\left(\frac{t}{i}\right) &\le \frac{k}{m} - 1 \\
\frac{t}{i} &\le e^{\frac{k-m}{m}} \\
i &\ge te^{\frac{m-k}{m}}
\end{aligned}
$$

Fraction of nodes with degrees $k_i(t) \le k$ (CDF):

$$F(k) = P(k_i(t) \le k) = \frac{n_0 + t - i}{n_0 + t} = \frac{n_0 + t - te^{\frac{m-k}{m}}}{n_0 + t} \approx 1 - e^{\frac{m-k}{m}}$$

$$P(k) = \frac{d}{dk}F(k) = \frac{1}{m}e^{-\frac{k-m}{m}} = \frac{e}{m}e^{-\frac{k}{m}}, \quad k \ge m$$

Mean field approximation

Continues time approximation ($\Pi(k_i) = \frac{1}{n_0+t-1} \approx \frac{1}{t}$)

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

$$k_i(t + \delta t) = k_i(t) + \frac{m}{t}\delta t$$

Differential equation:

$$\frac{dk_i(t)}{dt} = \frac{m}{t}$$

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

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

Solution:

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

Preferential attachment model

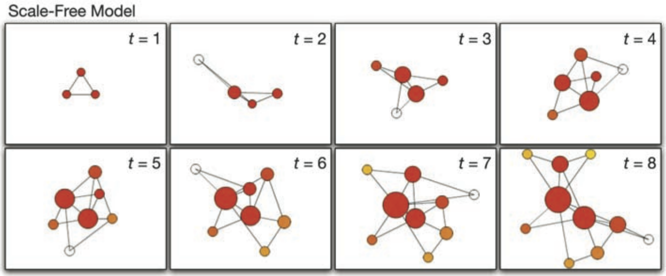
Barabasi and Albert, 1999
Dynamical growth model

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

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

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

Preferential attachment model



Barabasi, 1999

Preferential attachment

Continues time, mean field approximation:

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

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

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

initial conditions: $k_i(i) = m$

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

Solution:

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

Preferential attachment

Time evolution of a node degree

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

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

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

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

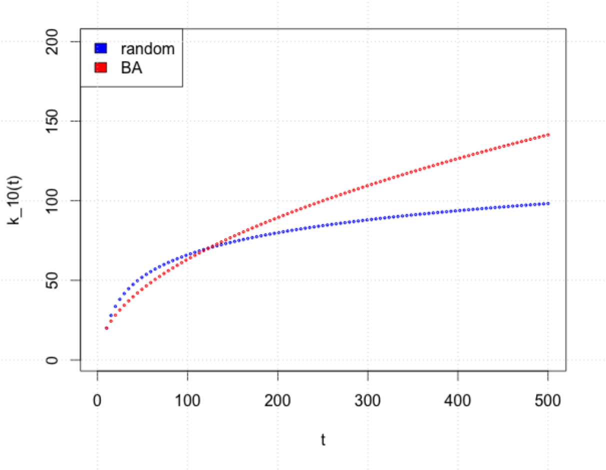
Fraction of nodes with $k_i(t) < k$ (CDF):

$$F(k) = P(k_i(t) \leq k) = \frac{n_0 + t - i}{n_0 + t} = \frac{n_0 + t - m^2 t/k^2}{n_0 + t} \approx 1 - \frac{m^2}{k^2}$$

Distribution function:

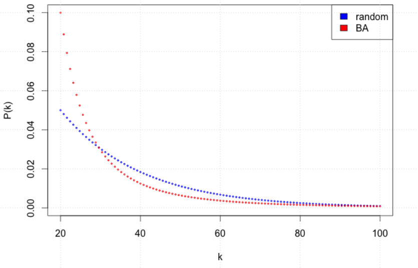
$$P(k) = \frac{d}{dk}F(k) = \frac{2m^2}{k^3}$$

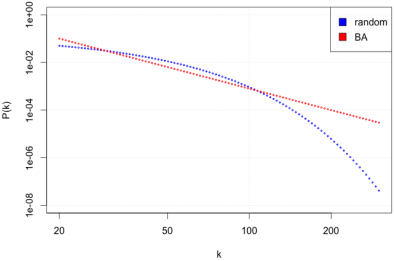
Dynamic growth



$$BA : k_i(t) = m\left(\frac{t}{i}\right)^{1/2}, \quad RG : k_i(t) = m\left(1 + \log\left(\frac{t}{i}\right)\right)$$

Dynamic growth





$$BA : P(k) = \frac{2m^2}{k^3}, \qquad RG : P(k) = \frac{e}{m} e^{-\frac{k}{m}}$$

Preferential attachment

- Power law distribution function:

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

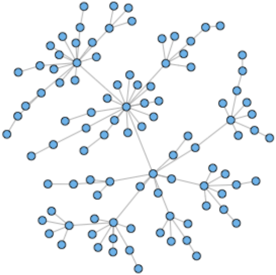
- Average path length (analytical result) :

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

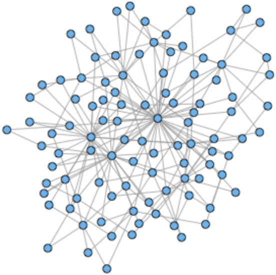
- Clustering coefficient (numerical result):

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

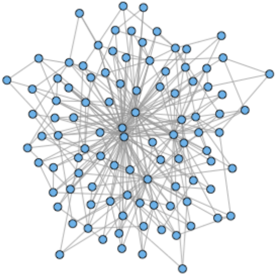
Preferential attachment model



$m = 1$



$m = 2$



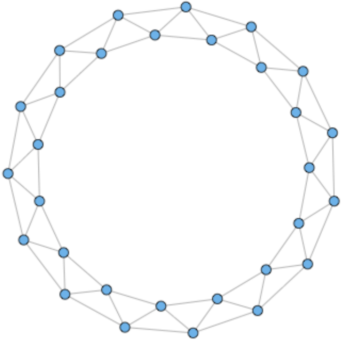
$m = 3$

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 $d = 8$

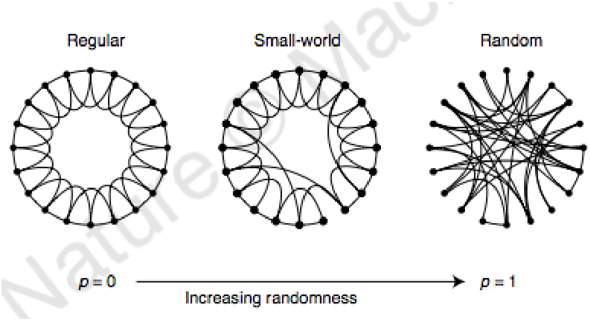
Small world

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

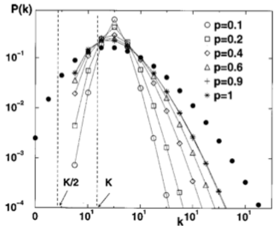
Small world

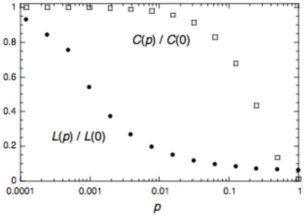


Watts, 1998

Small world model

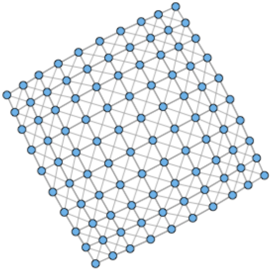
- Node degree distribution:
  Poisson like
- Ave. path length $\langle L(p) \rangle$:
  $p \to 0$, ring lattice, $\langle L(0) \rangle = 2n/k$
  $p \to 1$, random graph, $\langle L(1) \rangle = \log(n)/\log(k)$
- Clustering coefficient $C(p)$:
  $p \to 0$, ring lattice, $C(0) = 3/4 = const$
  $p \to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

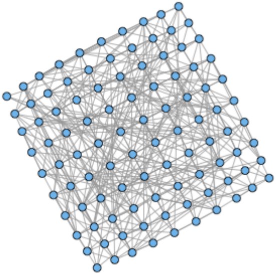




Watts, 1998

Small world model





20% rewiring:

ave. path length = 3.58 -> ave. path length = 2.32

clust. coeff = 0.49 -> 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

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