NATIONAL RESEARCH
UNIVERSITY

Network models

Network Science
Lecture 4

Leonid Zhukov

lzhukov@hse.ru
www.leonidzhukov.net/hse/2022/networks

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

Network models

Empirical network features:

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

Generative models:

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

Motivation

- Citation networks
- Collaboration networks
- Web
- Social networks

Motivation

- Citation networks
- Collaboration networks
- Web
- Social networks

Most of the networks we study are dynamic, they evolve over time, expanding by adding new nodes and edges

Preferential attachment model

Barabasi and Albert, 1999

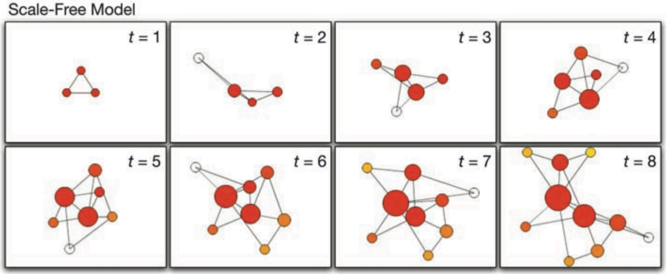
Dynamic growth: start at $t = 0$ with n_0 nodes and $m_0 \geq n_0$ edges

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:

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

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

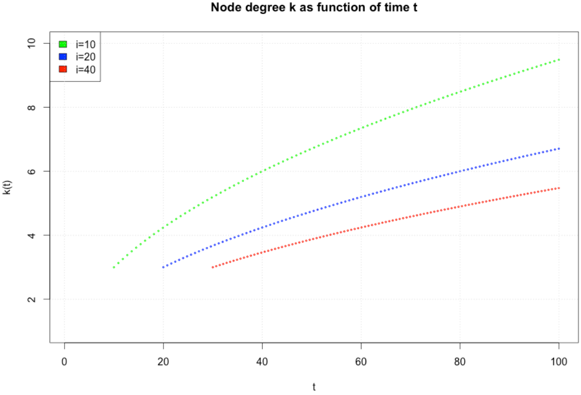
node *i* is added at time t_i: $k_i(t_i) = m$

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

Solution:

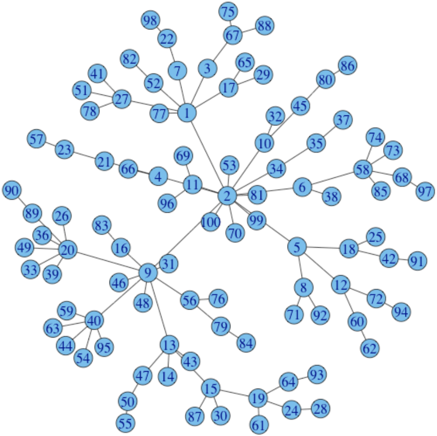
$$k_i(t) = m \left(\frac{t}{t_i} \right)^{1/2}$$

Preferential attachement



$$k_i(t) = m\left(\frac{t}{t_i}\right)^{1/2}; \qquad \frac{dk_i(t)}{dt} = \frac{m}{2}\frac{1}{\sqrt{tt_i}}$$

Preferential attachement



Preferential attachment

Time evolution of a node degree

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

Find probability $P(k' \leq k)$ of a randomly selected node to have $k' \leq k$ at time t (fraction of nodes with $k' \leq k$). Nodes with $k_i(t) \leq k$:

$$m\left(\frac{t}{t_i}\right)^{1/2} \leq k \quad \Rightarrow \quad t_i \geq \frac{m^2}{k^2}t$$

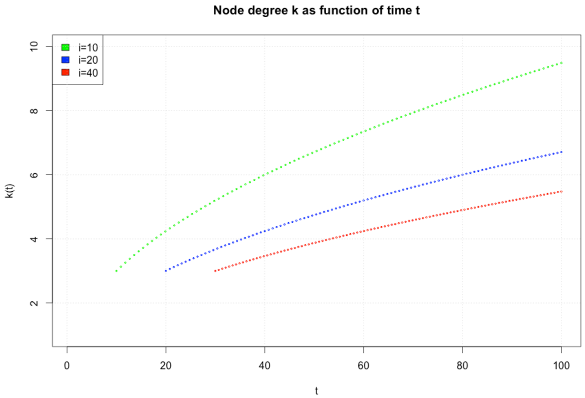
Cumulative function:

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

Distribution function:

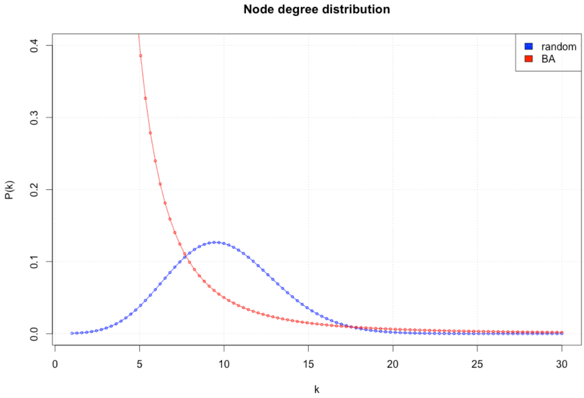
$$P(k) = \frac{d}{dk}F(k) = \frac{2m^2}{k^3}$$

Preferential attachement



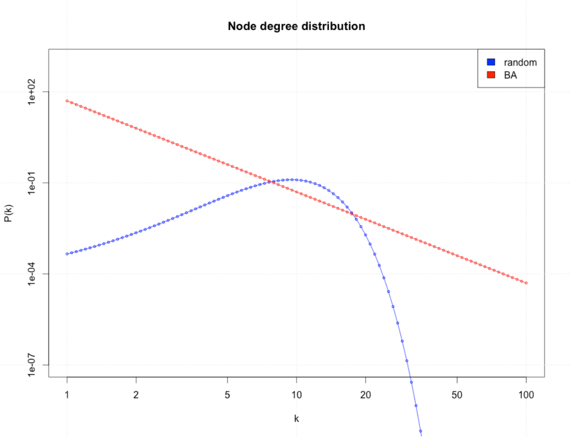
$$m\left(\frac{t}{t_i}\right)^{1/2} \leq k \quad \Rightarrow \quad t_i \geq \frac{m^2}{k^2}t$$

Preferential attachment vs random graph



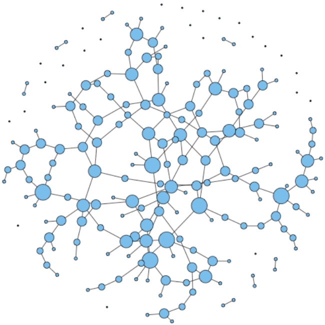
$$BA : P(k) = \frac{2m^2}{k^3}, \quad ER : P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad \langle k \rangle = pn$$

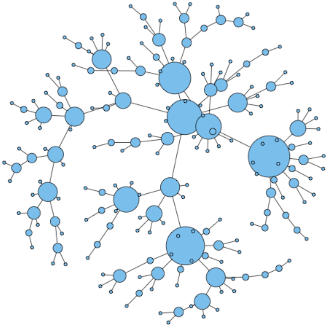
Preferential attachment vs random graph



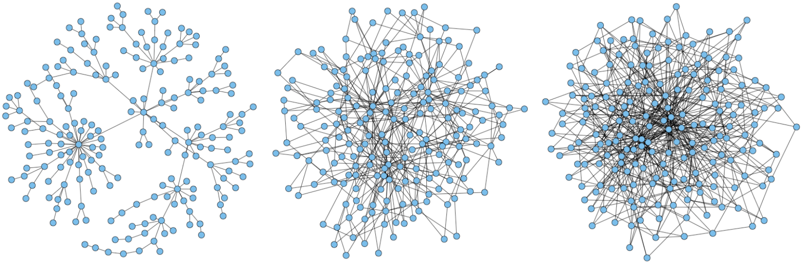
$$BA : P(k) = \frac{2m^2}{k^3}, \quad ER : P(k) = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad \langle k \rangle = pn$$

Preferential attachment vs random graph





Preferential attachment model



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

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. **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):

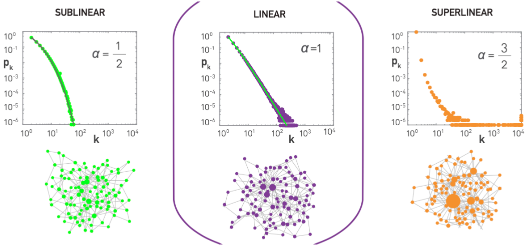
$$C \sim N^{-0.75}$$

Non-linear preferential attachment

- Non-linear preferential attachment models:

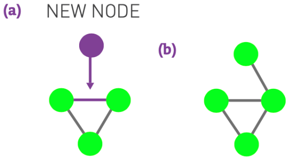
$$\Pi(k) \sim k^{\alpha}$$

- $\alpha = 0$, no hubs, exponential dsitribution
- $0 < \alpha < 1$, sublinear, smaller hubs, stretched exponential
- $\alpha = 1$, scale-free, hubs, power law
- $\alpha > 1$, superlinear, super hubs, hubs-and-spoke



Local growth mechanism:

- Growth: at each time step add a new node
- Link selection: select link at random and connect to one of two nodes at the ends



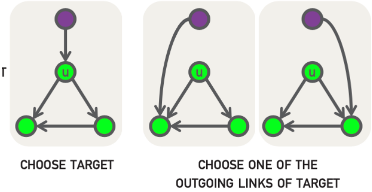
Probability to connect to a node with degree k:

$$\Pi(k) = \frac{k p_k}{\langle k \rangle}$$

Copying model

Local growth mechanism:

- Random connection: with probability p connect to a random node u
- Copying: with probability $1 - p$ randomly choose an outgoing link from u and connect to its target



Probability to connect to a node with degree k :

$$\Pi(k) = \frac{p}{n} + (1 - p)\frac{k}{2m}$$

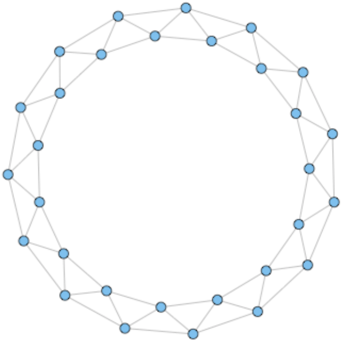
Historical note

- 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

Local random models vs global optimization models

Small world

Motivation: keep high clustering, get small diameter



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

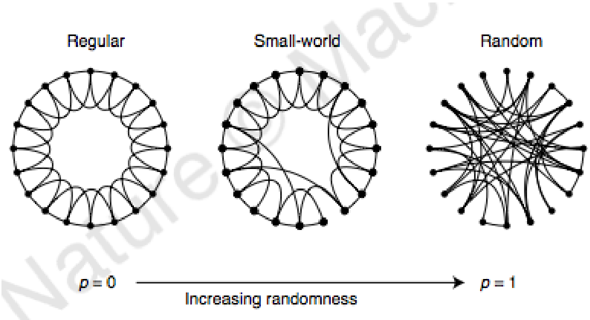
Small world

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

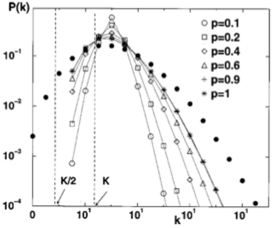
Small world

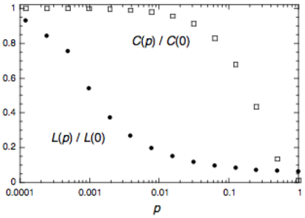


Watts, 1998

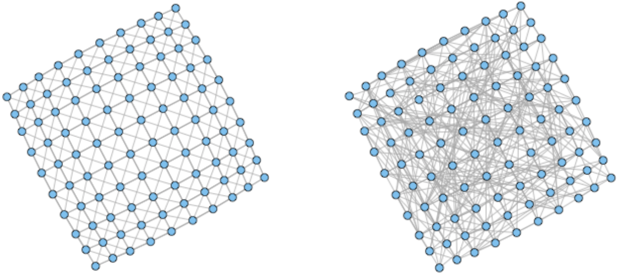
Small world model

- Node degree distribution:
  Poisson like
- Ave. path length $\langle L(p)\rangle$:
  $p \to 0$, ring lattice, $\langle L(0)\rangle = n/2k$
  $p \to 1$, 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, $C(1) = k/n$





Small world model



20% rewiring:

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

clust. coeff = 0.49 $\rightarrow$ 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