DATA ANALYSIS II

Evolving Networks

2021/22

References

- Barabási, L-A. (2016) *Network Science*. Cambridge University Press.
  [The Barabási-Albert Model: http://networksciencebook.com/chapter/5]
  [Evolving Networks: http://networksciencebook.com/chapter/6]

- Bianconi, G., Barabási, A. L. (2001). *Competition and multiscaling in evolving networks*. EPL (Europhysics Letters), 54(4), 436.

Motivation

- Why do so different systems as the WWW or the cell converge to a similar scale-free architecture?

- Why does the random network model of Erdős-Rényi fail to reproduce the hubs and the power laws observed in real networks?

- Growth and Preferential Attachment

  - Networks Expand Through the Addition of New Nodes

  - Nodes Prefer to Link to the More Connected Nodes

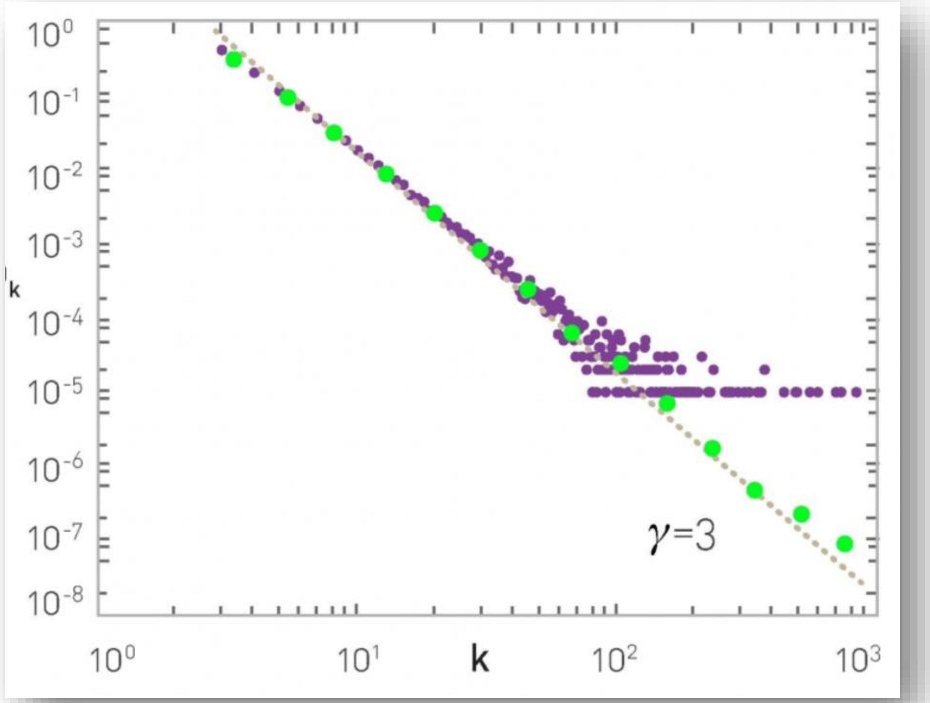
Barabási-Albert Model

- **Growth.** At each timestep we add a new node with m ($\leq m_0$) links that connect the new node to m nodes already in the network.
- **Preferential attachment.** The probability $\Pi(k)$ that a link of the new node connects to node i depends on the degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}$$

Degree Distribution

- The degree distribution of a network generated by the Barabási-Albert model. The figure shows p_k for a single network of size N=100,000 and m=3.



Degree Dynamics

- In the model an existing node can increase its degree each time a new node enters the network. This new node will link to *m* of the *N(t)* nodes already present in the system.

- The rate at which an existing node *i* acquires links as a result of new nodes connecting to it is

$$\frac{dk_i}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}$$

$$\beta = \frac{1}{2}$$

$$\sum_{j=1}^{N-1} k_j = 2mt - m$$

$$\frac{dk_i}{dt} = \frac{k_i}{2t-1}$$

$$\frac{dk_i}{k_i} = \frac{1}{2}\frac{dt}{t}$$

We call β the dynamical exponent.

Predictions

- The degree of each node increases following a power-law with the same dynamical exponent $\beta = 1/2$. *Hence all nodes follow the same dynamical law.*

- The growth in the degrees is sublinear (i.e., $\beta < 1$). This is a consequence of the growing nature of the Barabási-Albert model: Each new node has more nodes to link to than the previous node. *Hence, with time the existing nodes compete for links with an increasing pool of other nodes*.

- The earlier node *i* was added, the higher is its degree $k_i(t)$. *Hence, hubs are large because they arrived earlier, a phenomenon called first-mover advantage in marketing and business.*

Degree Distribution

- The distinguishing feature of the networks generated by the Barabási- Albert model is their power-law degree distribution (continuum theory).

$$p(k) \approx 2m^{1/\beta} k^{-\gamma}$$

with

$$\gamma = \frac{1}{\beta} + 1 = 3$$

- Therefore, the degree distribution follows a power law with degree exponent γ=*3*, in agreement with the numerical results

Degree Distribution

- For large *k* it reduces to $p_k \sim k^{-3}$, or $\gamma = 3$.

- The degree exponent γ is independent of *m*, a prediction that agrees with the numerical results.

- In summary, the analytical calculations predict that the Barabási-Albert model generates a scale-free network with degree exponent $\gamma=3$.

  - The degree exponent is independent of the *m* and m_0 parameters.

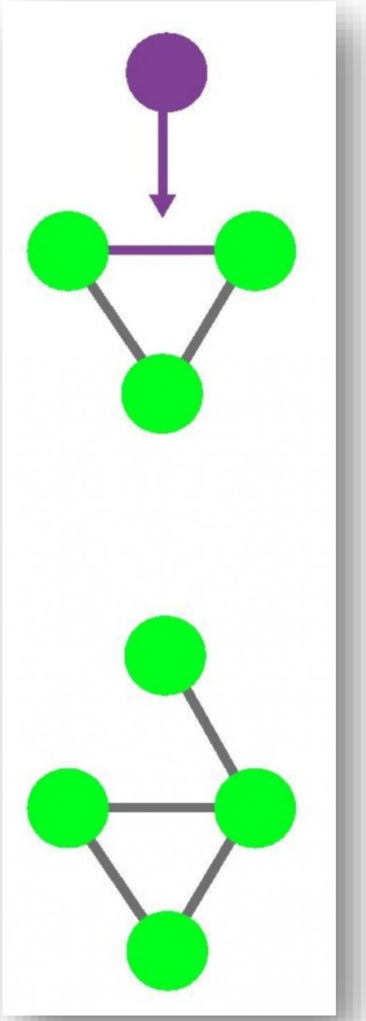
  - Furthermore, the degree distribution is stationary (i.e. time invariant), explaining why networks with different history, size and age develop a similar degree distribution.

Origins of Preferential Attachment

- We can build models that generate scale-free networks apparently without preferential attachment.
- They work by generating preferential attachment.
  - Link Selection Model
  - Copying Model

Link Selection Model

- **Growth.** At each time step we add a new node to the network.
- **Link Selection.** We select a link at random and connect the new node to one of the two nodes at the two ends of the selected link.
- The model requires no knowledge about the overall network topology hence it is inherently local and random. It generates preferential attachment.



Properties

- The higher is the degree of a node, the higher is the chance that it is located at the end of the chosen link.

- The more degree-*k* nodes are in the network (i.e., the higher is p_k), the more likely that a degree-*k* node is at the end of the link.

Copying Model

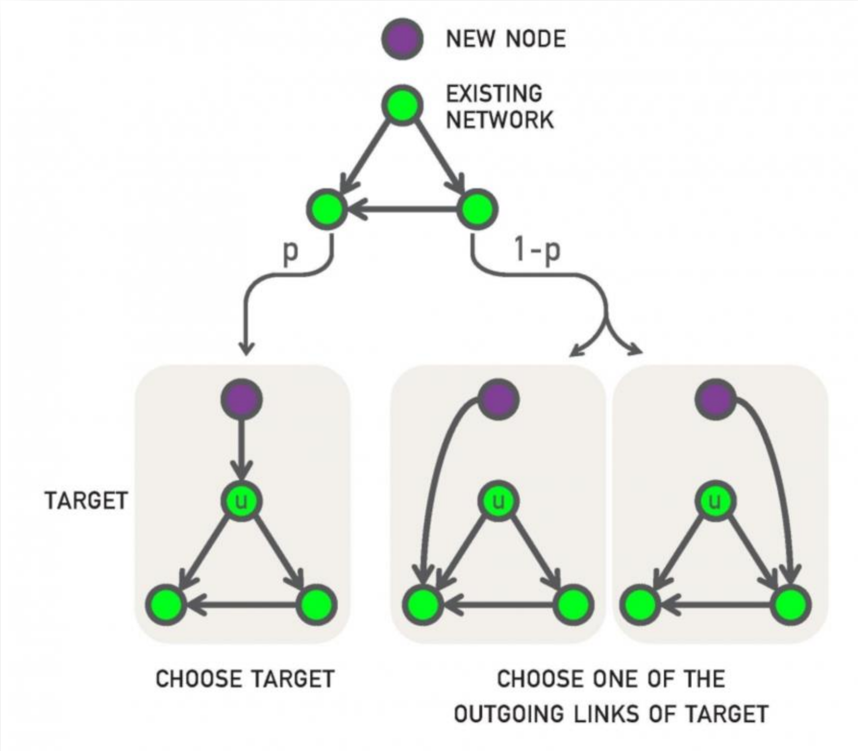
- The main steps of the copying model.
  - A new node connects with probability p to a randomly chosen target node u, or…
  - …with probability $1\text{-}p$ to one of the nodes the target u points to. In other words, with probability $1\text{-}p$ the new node copies a link of its target u.

Two Steps

- **Random Connection.** With probability *p* the new node links to *u*, which means that we link to the randomly selected web document.

- **Copying.** With probability *1-p* we randomly choose an outgoing link of node *u* and link the new node to the link's target.

- In other words, the new webpage copies a link of node u and connects to its target, rather than connecting to node *u* directly.



Bianconi-Barabási Model

- **Growth.** In each timestep a new node j with m links and fitness η_j is added to the network, where η_j is a random number chosen from a fitness distribution $\rho(\eta)$. Once assigned, a node's fitness does not change.

- **Preferential Attachment.** The probability that a link of a new node connects to node i is proportional to the product of node i's degree k_i and its fitness η_i,

$$\Pi_i = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

Degree Dynamics

- Let us assume that the time evolution of k_i follows a power law with a fitness-dependent exponent $\beta(\eta_i)$.

$$\frac{\partial k_i}{\partial t} = m \frac{\eta_i k_i}{\sum_j \eta_j k_j}$$

$$k(t, t_i, \eta_i) = m\left(\frac{t}{t_i}\right)^{\beta(\eta_i)}$$

$$\beta(\eta) = \frac{\eta}{C}$$
$$C = \int \rho(\eta) \frac{\eta}{1-\beta(\eta)} d\eta$$

- In the Barabási-Albert model we have $\beta = 1/2$, hence the degree of each node increases as a square root of time.

- In the Bianconi-Barabási model the dynamic exponent is proportional to the node's fitness, η, hence each node has its own dynamic exponent.

- Consequently, a node with a higher fitness will increase its degree faster.

Other Principles in Evolving Networks

- Internal Links
  - New links do not only arrive with new nodes but are added between pre-existing nodes.
- Node Deletion
  - Nodes and links can disappear.
- Accelerated Growth
  - The number of links increases linearly with the number of nodes. In real-world networks, the number of links can grow faster than N.
- Aging
  - Nodes can have a limited lifetime.

Assignment

- Implement network generators based on Link Selection Model and Copying Model. Generate networks with more then thousands nodes. Visualize these networks and determine their properties (e.g., average degree and distribution, average clustering coefficient).

- Implement at least one of the other principles as a part of the generators above. Compare properties of resulting networks with the previous ones.