CS4973

network analysis

Recall Power Law Networks

- $P(k) \propto k^{-\alpha}$
- Power law exponent (negative slope) typically in range $[2,3]$.
- The Barabási-Albert (BA) model is a simple generation mechanism that produces networks with power law degree distributions.

Recall Power Law Networks

- **Growth:** *in real networks the number of nodes continually grows thanks to the addition of new nodes.*

  - The WWW started with 1 webpage, built by Tim Berners-Lee.
  - Today, it has on the order of ~100B – 1T pages.

  - Collaboration networks grow as new papers are published.

  - The actor network grows with new movies.

  - The number of genes in a human cell grew to more than 20k over ~4B years.

Recall Power Law Networks

- **Preferential Attachment:** *most real networks new nodes prefer to link to the more connected nodes.*
  - We are more likely to link to popular websites (Facebook, X, Instagram)
  - We are more likely to read, and therefore, cite, popular papers.
  - Famous actors are more likely to be considered for a new role.
  - *Often called the rich-get-richer phenomenon.*

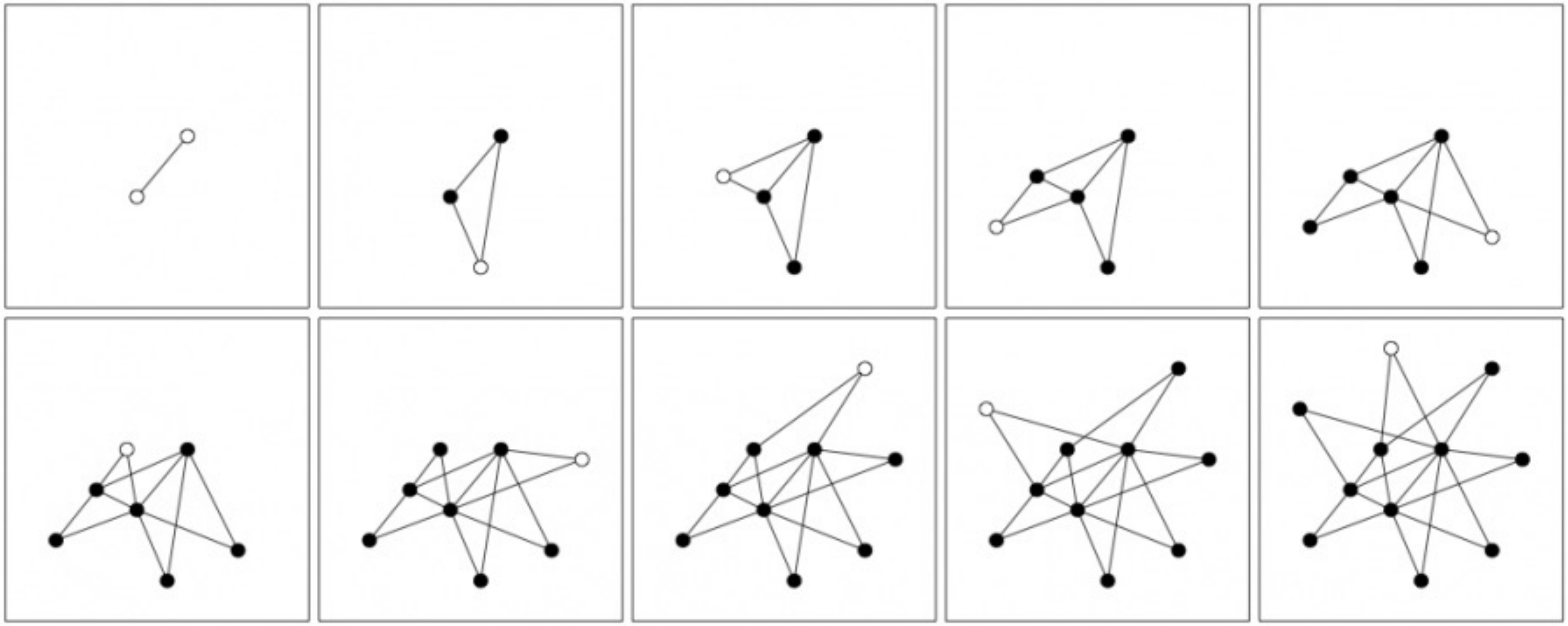
Barabási-Albert (BA) model:

- Start with m_0 nodes, randomly wired, such that each node has degree ≥ 1.
- At each timestep, add a new node v with $m \leq m_0$ links that connect to m nodes already in the network.
- The probability of the node v connecting to a node u in the network is:

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

- This is a probabilistic model, that doesn't take say anything about connecting to hubs explicitly.

Barabási-Albert (BA) model:



BA Model Web Demo

Evolution of the Barabási-Albert Network

$m_0 = 2, m = 2$

Barabási-Albert (BA) model:

- Are both growth and preferential attachment necessary?

- In-class activity:

    - Simulate a network using only the *growth* element.

      Plot the degree distribution.

    - Simulate a network using only the *preferential attachment* element.

      Plot the degree distribution.

Degree Distribution for BA model

- Let k_i be the degree of node i. The probability of a new edge connecting to i is

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

- On average, the expected increase in degree is proportional to the degree:

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

- Solving the differential equation, we get:

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{0.5}$$

where t_i represents the timestep when node i was added to the network.

Degree Distribution for BA model

- Cumulative density function (CDF)

$$P(k_i(t) < k) = P\left(t_i^2 > \frac{m^2 t}{k^2}\right) = 1 - P\left(t_i^2 \leq \frac{m^2 t}{k^2}\right) = 1 - \frac{m^2 t}{k^2} \frac{1}{(t + m_0)}$$

(assuming uniform intervals of adding nodes)

- Probability density function (PDF) is the derivative of the CDF:

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^2 t}{d^3 (t + m_0)}$$

$$\text{As } t \to \infty, P(k) = \frac{2m^2}{d^3}$$

Network Measures for BA model

- $C = \frac{m_0 - 1}{8} \frac{(\ln t)^2}{t} \sim \frac{(\ln t)^2}{t}$
- Average path length: $l \sim \frac{\ln|V|}{\ln(\ln|V|)}$
- Nonlinear Preferential Attachment:

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

- Estimating exponent from data? (HINT: linear regression/MLE)