Barabási-Albert Model

Growth and Preferential Attachment Model

Topics

- Introduction
- Growth and Preferential Attachment
- The Barabási-Albert Model
- Degree Dynamics and Degree Distribution
- Measuring Preferential Attachment
- Competing 'Preferential Attachment' Models
- Diameter and Clustering Coefficient
- Summary and Limitations

Introduction

Two dundamental questiona behind existence of Scale-Free Networks:

- Why do such different systems as the WWW and the cell converge to a similar scale-free architecture?
- Why does the random network model of Erdős and Rényi fail to reproduce the hubs and the power laws observed in real networks?

Image source: http://salon.seedmagazine.com/img/salon/portrait_barabasi_240x270.jpg

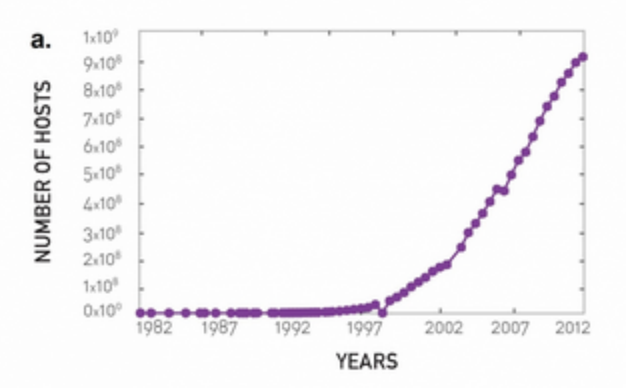
Growth and Preferential Attachment

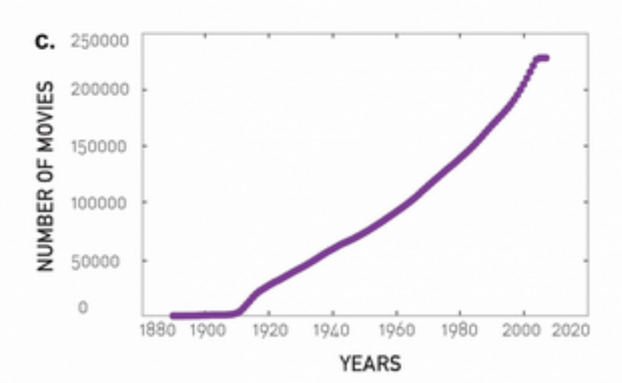
Networks expand through the addition of new nodes (**Growth**).

- Number of web documents (nodes) keeps growing on WWW.
- The actor network continues to expand through release of new movies.
- Number of published research papers continues to grow each time a new paper is published.
- Number of genes continue to grow (from few to over 20,000 in four billion years).

Nodes prefer to link with more connected nodes (**Preferential Attachment**).

- Movie actors: popular actors have higher chance of being considered for a new role.
- More cited paper has higher probability of being read, and, in turn, cited again.
- Web documents are more likely to link to a high-degree node than to a node with only a few links.





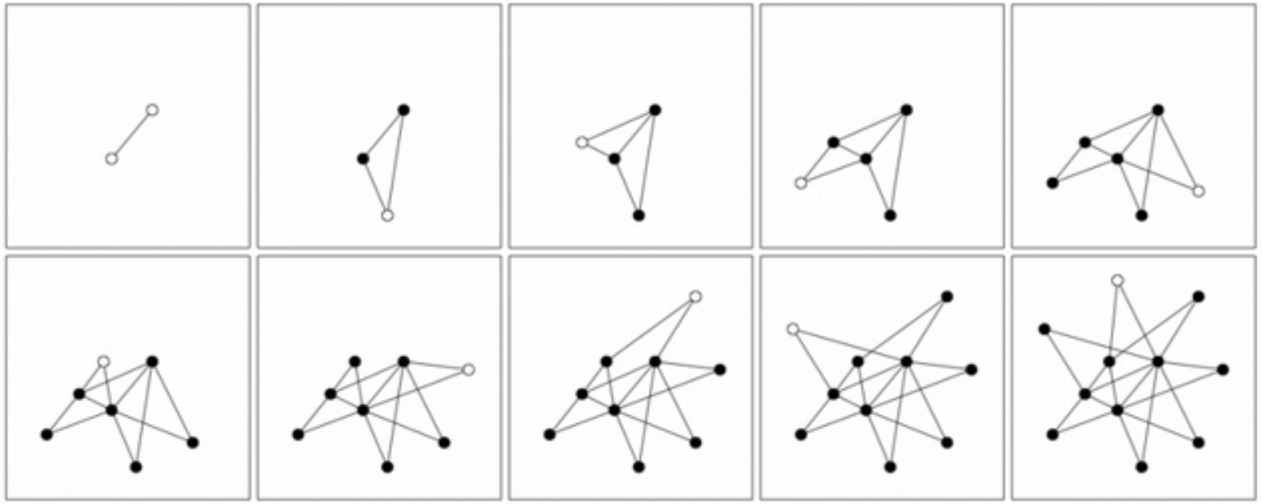
Barabási-Albert Model

Preferential attachment is a probabilistic mechanism:

Given by $\pi(k_i)$

If a new node has a choice between a degree-two and degree-four node, it is twice as likely to connect to the degree-four node.

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



Sequence of ten images: In most real networks new nodes (empty white circles) use preferential attachment when deciding whom to connect to.

Barabási – Albert Model

Model starts with given m_0 nodes.

At each time step, we add a new node with m links (where $m \leq m_0$).

Probability that a link k of the new node connects to node i depends on the degree k_i

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

After t time steps the model generates a network with $N = t + m_0$ nodes and $m_0 + m_t$ links.

Emergence of a Scale-Free Network

Growth and Preferential Attachment

Degree Dynamics (Rich-Gets-Richer)

The degree of each node increases following a power law with exponent = ½

After time t, a node i will have k_i degrees (a new node at event time t_i comes with m links, and β = ½ for Barabási – Albert model - always)

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

Growth in degrees is sublinear (exponent is ½ which is less than 1) – with time the existing nodes compete for links with an increasing pool of other nodes.

First mover advantage: Hubs are large because they arrived earlier.

Degree Distribution

Analytically, it is shown that Barabási - Albert model generates a scale-free network with power law $y = 3$ (for large k).

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

Measuring Preferential Attachment

Attachment probability depends on the node degree.

- Rich-gets-richer

Functional form of the attachment probability is proportion to k^{α}

For internet and citation networks α = 1

- Linear preferential attachment

For actor networks α < 1

- Sublinear preferential attachment
- Contrast this with growth – which is always sublinear with β = 1/2
- Hubs are smaller when α < 1

The Absence of Growth or Preferential Attachment

Both growth and preferential attachment are simultaneously needed for the emergence of the scale-free property.

If they both are present, they do lead to scale-free networks.

Exceptions:

- There any scale-free networks due to some completely different mechanism.
- Link-selection, copying and optimization models that do not have preferential attachment built into them, yet do lead to a scale-free network.

Competing 'Preferential Attachment' Models

Local Mechanism models that lead to scale-free networks without preferential attachment:

- **Link selection model**: a link is randomly selected and the new node is attached to one of the nodes on the either side of the randomly selected link. It is shown that the random selection of a link favors selecting link which has high degree nodes on its ends. Thus it generates a linear preferential model.
- **Copying model**: new node selects a random existing node, and then connects to that selected node's connection rather than connecting to that selected node itself. It is similar to we "copying" friends of our friends. Or web page "copying" links from an existing web page.
- **Optimization**: rational choice theory suggesting that we prefer to transact with a "trustworthy" source which has been trusted by many others.

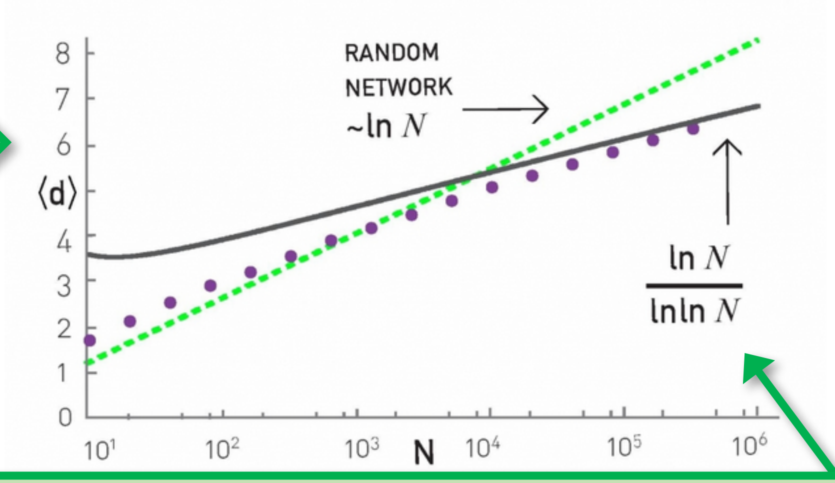
Diameter and Clustering

Diameter of Barabási - Albert models is smaller to diameter of a random graph of a similar size.

- Especially for large N

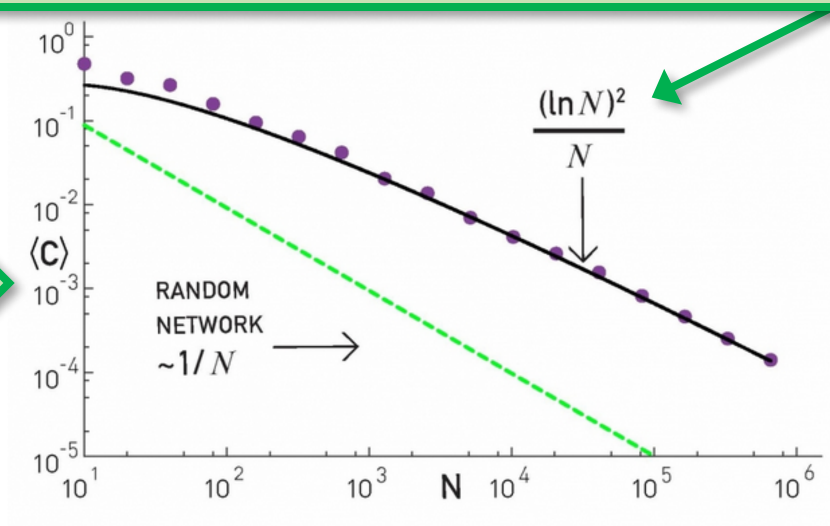
Clustering in Barabási - Albert network is locally more clustered than a random network.

Diameter



Barabási – Albert dimeter and clustering coefficients are proportional to:

Clustering



Barabási – Albert Model: Summary Statistics

Number of Nodes	N = t (where t is number of time events lapsed)
Number of Links	N = mt (where m is the number of links associated with a new node)
Average Degree	<k> = 2m
Degree Dynamics	$k_i(t) = m \left(\frac{t}{t_i}\right)^{\beta}$ (where $\beta = ½$)
Degree Distribution	$P_k \sim k^{-y}$ (where y = 3)
Average Distance	<d> ~ Ln N / Ln Ln N
Clustering Coefficient	<C> ~ $(\text{Ln } N)^2$ / N

Limitations of the Barabási-Albert Model

Unable to describe many characteristics of real systems:

- The model predicts γ=3 while the degree exponent of real networks varies between 2 and 5.
- Many networks, like the WWW or citation networks, are directed, while the model generates undirected networks.
- Could lead to multi-links (all m links connecting to the same existing node).
- Many processes observed in networks such as disappearance of links and nodes are absent from the model.
- The model does not allow us to distinguish between nodes based on some intrinsic characteristics, like the novelty of a research paper or the utility of a webpage.

References

- Albert - László Barabási, *Network Science*, Cambridge University Press, 2016.
- http://www.barabasilab.com/

Questions? Thoughts? Visit the course's online discussion forum.