Preferential Attachment (BA Model)

Social Networks Analysis and Graph Algorithms

Prof. Carlos "ChaTo" Castillo — https://chato.cl/teach

**Universitat Pompeu Fabra**
Barcelona

Contents

- The BA or preferential attachment model

- Degree distribution under the BA model

- Distance distribution under the BA model

- Clustering coefficient under the BA model

Sources

- A. L. Barabási (2016). Network Science – Chapter 05

- R. Srinivasan (2013). Complex Networks – Chapter 12

- D. Easley and J. Kleinberg (2010): Networks, Crowds, and Markets – Chapter 18

- Data-Driven Social Analytics course by Vicenç Gómez and Andreas Kaltenbrunner

The number of nodes N increases: we need models of network growth

Number of hosts in WWW

Number of movies in IMDB

YEARS

YEARS
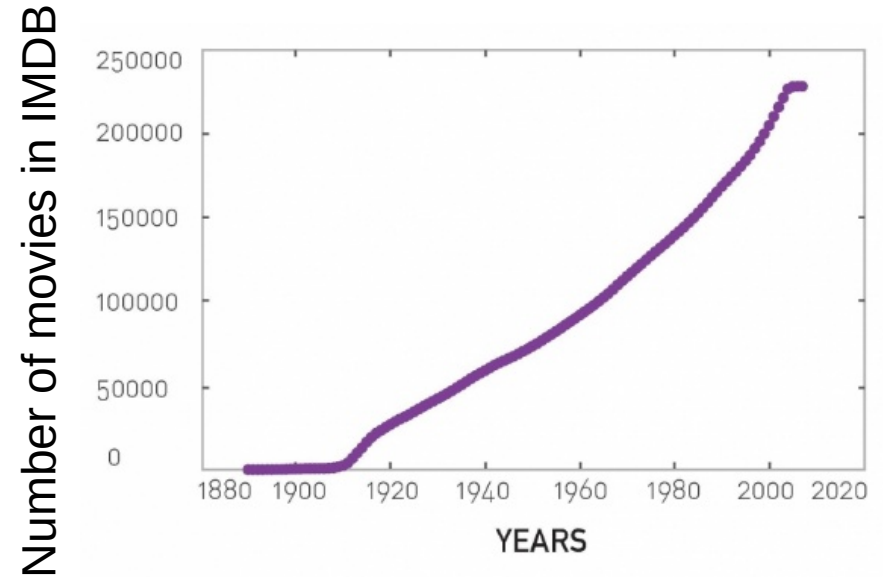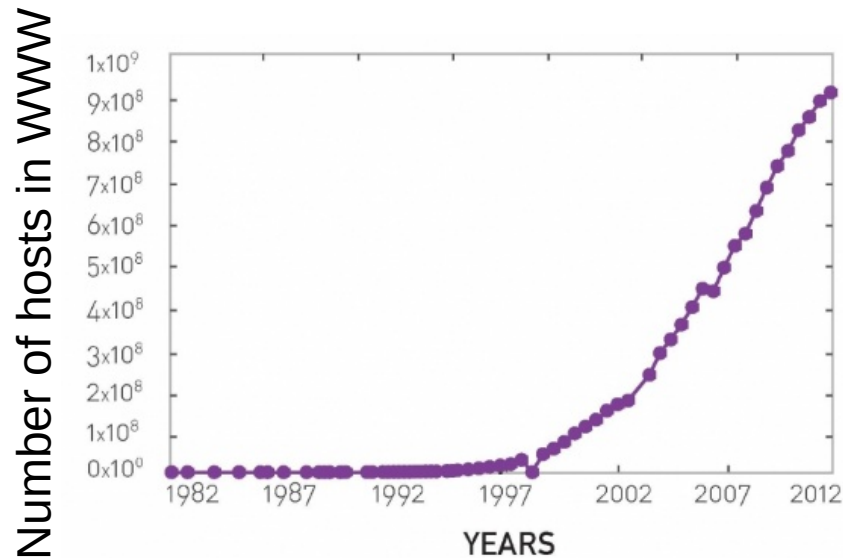
# Growth of an Open Source Project: Python

Gource
visualization

We have seen *what* but not *how*, or *why*

- Power-law degree distributions are prevalent

- We will give a possible answer to *how*

- For now, we will not answer *why*
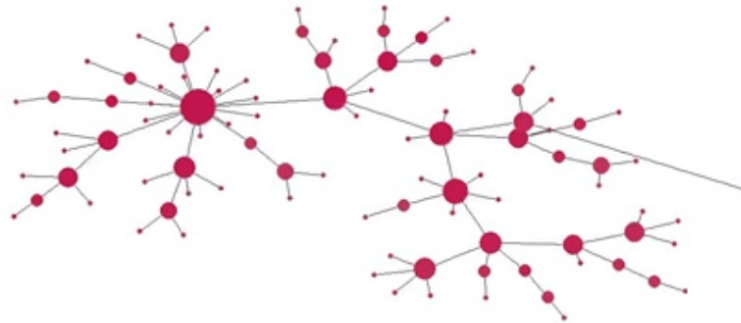
Preferential Attachment

Growth

- Suppose there are two web pages on a topic, one with many inlinks the other with few, which one am I most likely to link to?

- Which scientific papers are read?

- Which book authors sell more?

- Which actors are more sought after?
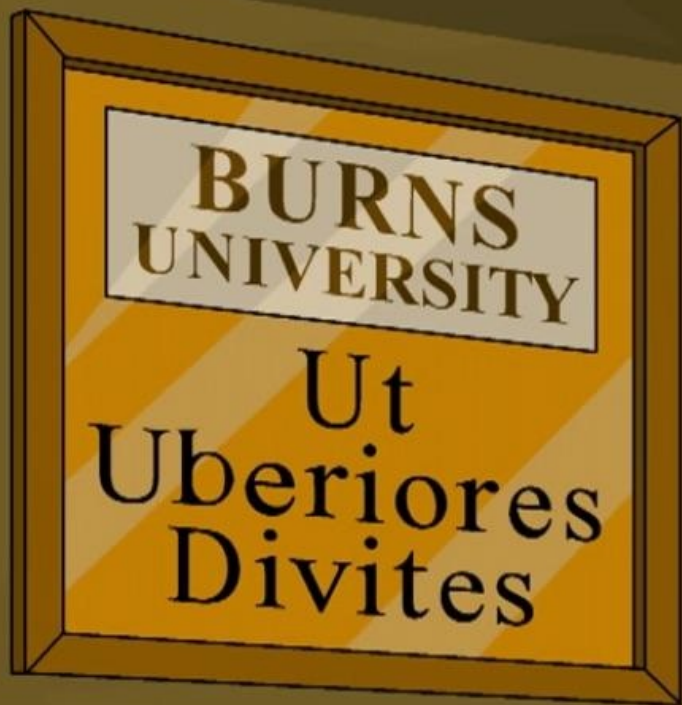
Our motto: *Ut uberiores divites.*

Preferential attachment simulation

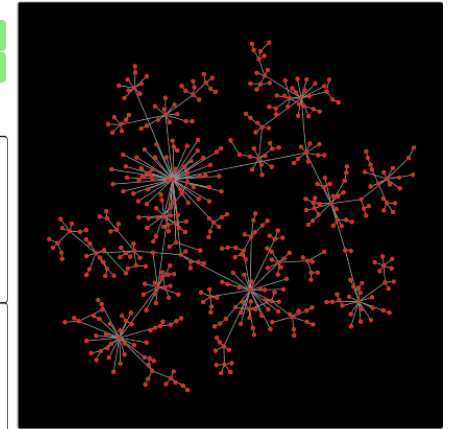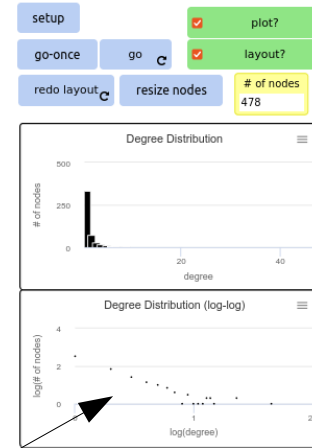https://www.youtube.com/watch?v=4GDqJVtPEGg

Exercise
Slope of degree distribution

Go to netlogoweb.org/launch and select:
"Sample Models / Networks / Preferential Attachment"

- Execute in Netlogo Web the

  "Preferential Attachment" program:

  – Click "setup"

  – Click "go"

  – Let it run to ~500 nodes

- **Guess the slope of the degree distribution in log-log scale**

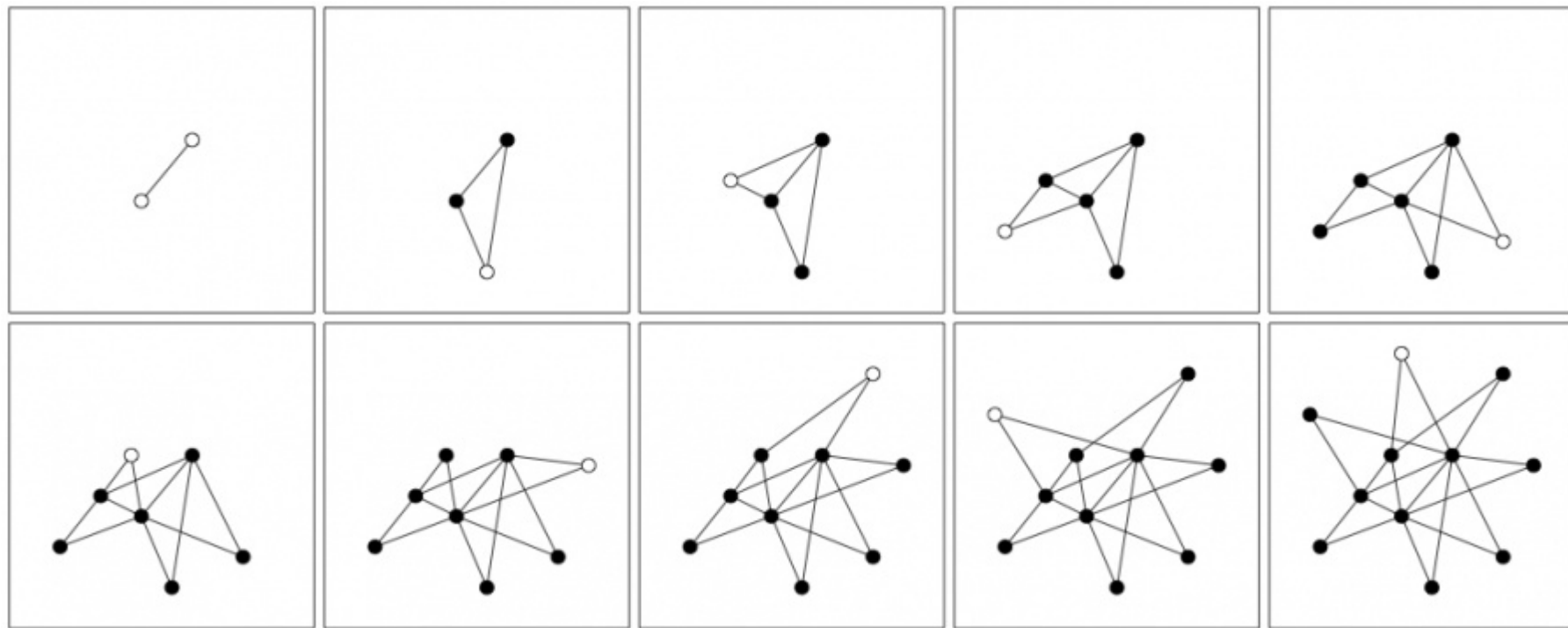Pin board: https://upfbarcelona.padlet.org/chato/y8kw9jcjlluo2p8c

The Barabási-Albert (BA) model

- Network starts with m_0 nodes connected arbitrarily as long as their degree is ≥ 1

- At every time step we add 1 node

- This node will have $m \leq m_0$ outlinks

- The probability of an existing node of degree k_i to gain one such link is
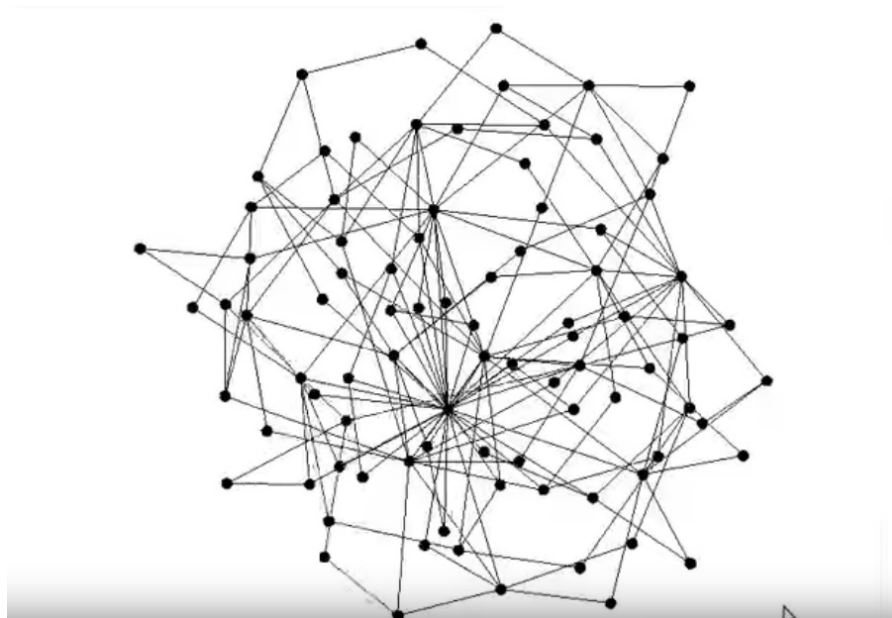
$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

In an ER network, $\Pi(k_i) = \frac{1}{N-1}$

Example ($m_0 = 2$; $m=2$)

Network growth with $m=2$

https://www.youtube.com/watch?v=wocaGeNKn7Y

The Barabási-Albert (BA) model

- Network starts with m_0 nodes connected arbitrarily as long as their degree is ≥ 1

- At every time step we add 1 node

- This node will have m outlinks $(m \leq m_0)$

- The probability of an existing node of degree k_i to gain one such link is
$$\Pi(k_i) = \frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

Write the formula for *N(t)* and *L(t)*: at *t=0* the network has *m_0* nodes and *L(0)* links

Summary

Things to remember

- Preferential attachment

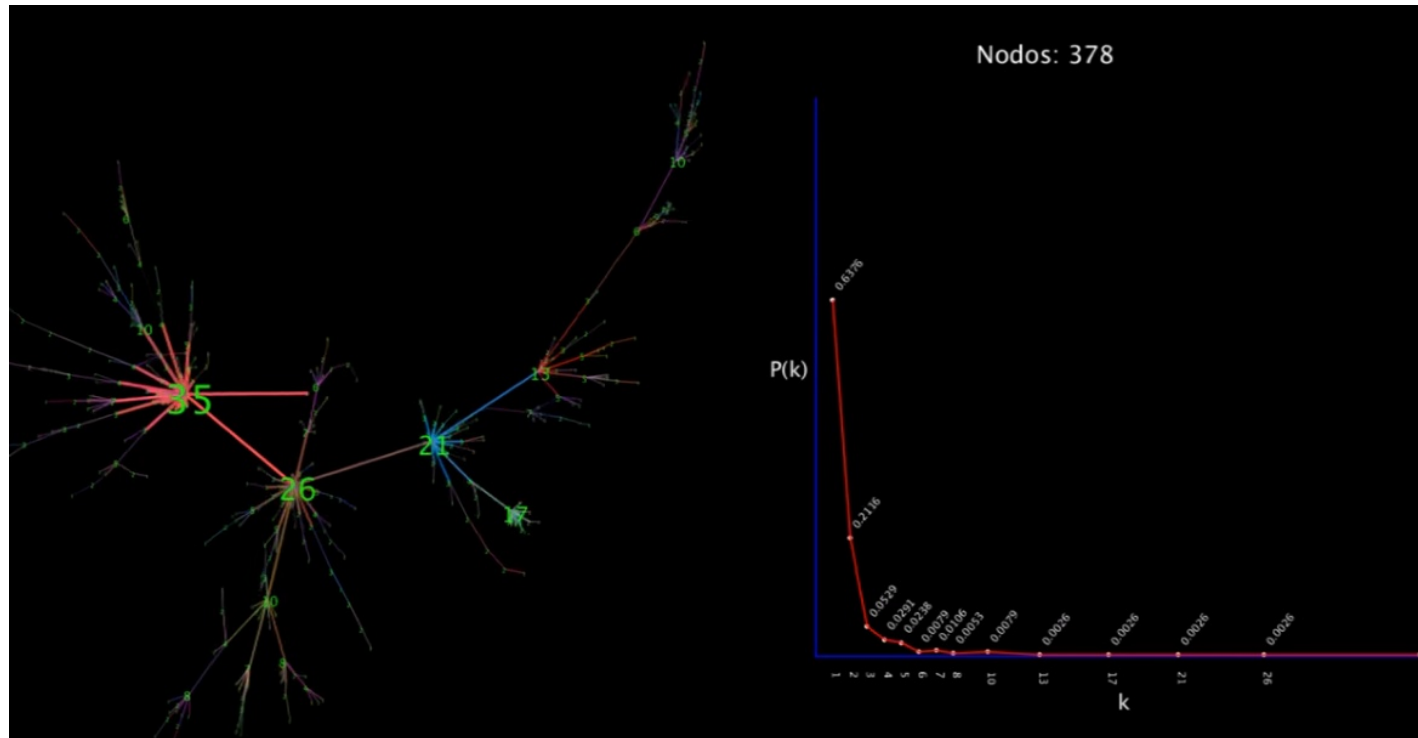- How to create a BA network step by step

Practice on your own

- Describe step by step in pseudocode how to create a Barabási-Albert graph with N nodes having m_0 starting nodes and m outlinks per node.

- For your pseudocode to be valid, if at any point there is a randomized step, you must indicate **what is the probability of each possible outcome**

Additional contents

Video of degree distribution

https://www.youtube.com/watch?v=5RIQweqPT6A