Random network models

Introduction to complex networks theory

ADA 2022/2023

Marija Mitrović Dankulov mitrovic@ipb.ac.rs
Aleksandra Alorić aleksandra.aloric@gmail.com

Aim

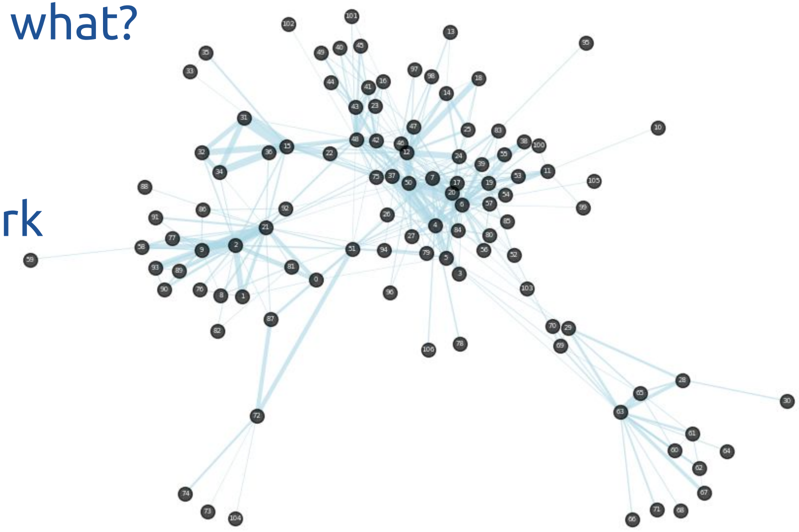
- Introduce random network models and their properties

- Models we'll cover today:
  - Erdős–Rényi
  - Watts–Strogatz
  - Barabási–Albert

Network measurement

- We learned how to spot a network, and also some ways to measure its properties and now what?

- Remember this Game of Thrones network
    - N= 107
    - L = 352
    - $\langle k \rangle = 6.6$
    - $\langle l \rangle = 2.9$
    - $\langle c \rangle = 0.6$



Are those values typical?

- If we look at other network with same number of nodes and edges, do they have the same average path/clustering? Are those values low/high?
- What would be the property of a similar regular network?
  - Is that a good benchmark? Why?
- What about some more heterogeneous network (not all the nodes are the same), let’s create a random network that has some of the properties the same and investigate if other network metrics hold
- How do we randomise?

Erdős–Rényi random graphs

- For a network with N nodes and L links, one can make a random graph G(N, L) where L pairs of nodes are chosen at random and connected
  - Graphs created this way will always have exactly N nodes and L links, but different pairs of nodes will be connected in different random realisation
- Alternatively, for every pair of nodes, we draw a link between them with probability p. This graph is usually denoted with G(N, p). To map to previous model, p needs to be 2L/(N(N-1))
  - Graphs generated this way will have N nodes but **variable** number of links. The expected value of number of links will be L (for p chosen as before), but every graph realisation will have random number of nodes. Consequently, average node degree will be a random variable.

Erdős–Rényi network construction

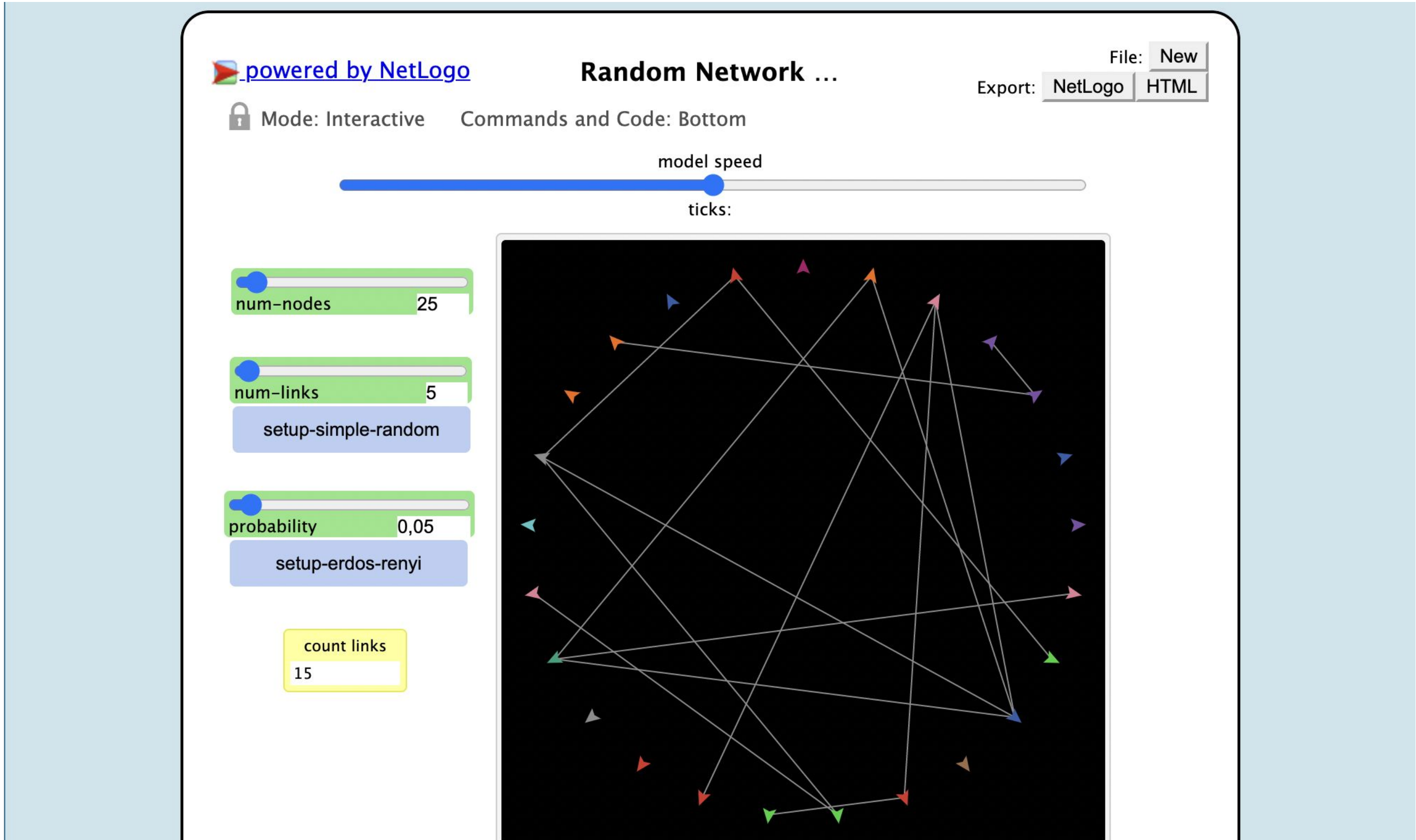
- G(N, p) - N nodes, p is probability that an edge exists
- Loop through N(N-1)/2 node pairs
  - Toss a biased coin (p is probability of head outcome)
  - If head: connect the two nodes


$p = 0.1$


$p = 0.25$


$p = 0.5$

Erdős–Rényi network examples



Properties of Erdős–Rényi network

- Average number of links: $< L >= p\frac{N(N-1)}{2}$

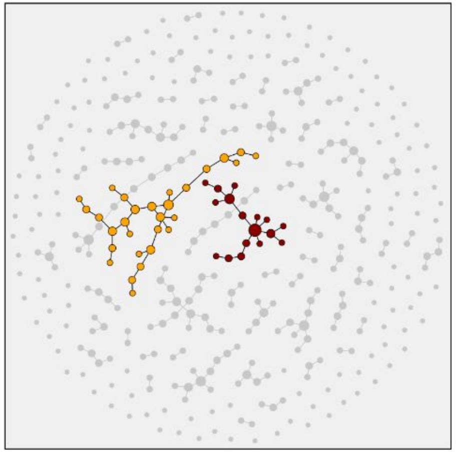
- Average degree is: $< k >= p(N-1)$

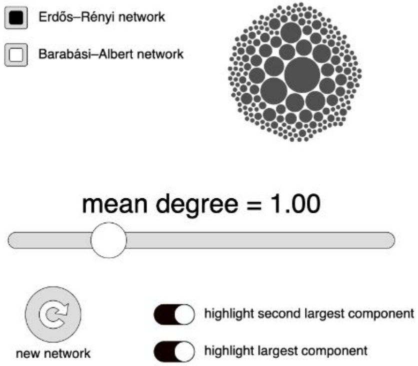
  Think of it either from perspective of a single node that can connect to at most N-1 other nodes, to each with prob p. Alternatively, think about definition of average degree using number of nodes and links.

- Degree distribution is Binomial, that is why you'll often find this model under the name binomial random graph:

$$P(\deg(v) = k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

A network's giant component





based on Complexity Explorable:

*" **The Blob -** The emergence of a giant component in a complex network"*

Paths in Erdős–Rényi network

- This simple network reproduces small world property present in many real world network, and absent in regular graphs that could've been our first benchmark

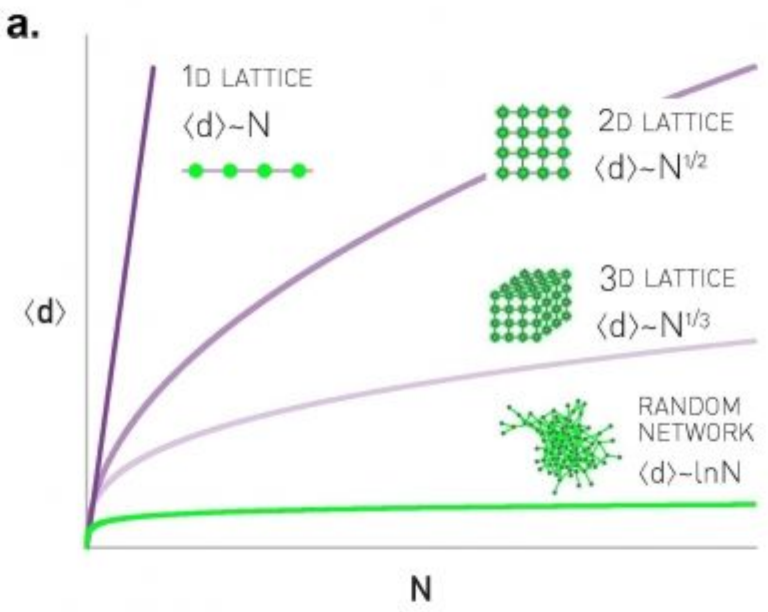


Image source:
http://networksciencebook.com/chapter/3#small-worlds

Clustering coefficient

- If node degree is k_i what is the expected number of links between those neighbours?
  - What is then the expected clustering coefficient of the node i?
  - What about the average clustering coefficient?

- Downsides of these results:
  - For fixed average degree, clustering coefficient diminishes with N
  - The local clustering and degree are unrelated (contrary to our observations)

Random graphs in python

Watts–Strogatz model

- Motivation for Watts–Strogatz model:
  - Interpolate between two extreme examples we've mentioned:
    - Completely regular networks
    - Completely random (ER)
  - Why?
    - Clustering coef is high in regular networks, but very low in ER graphs
    - Average path length is high in regular networks, but short in ER graphs
    - Real networks show short paths (closer to ER random networks) but high clustering (closer to regular networks)

Watts–Strogatz network construction

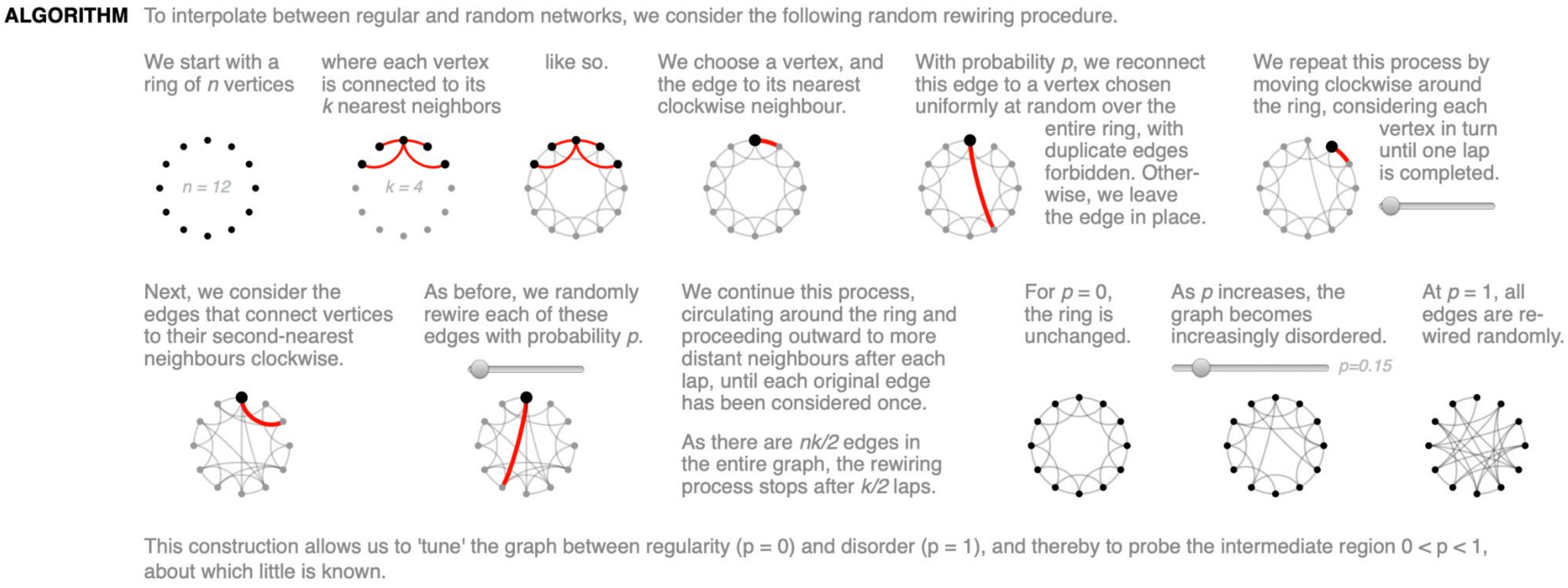


Image source:
http://worrydream.com/ScientificCommunicationAsSequentialArt/ScientificCommunicationAsSequentialArt.pdf

Properties of Watts–Strogatz network

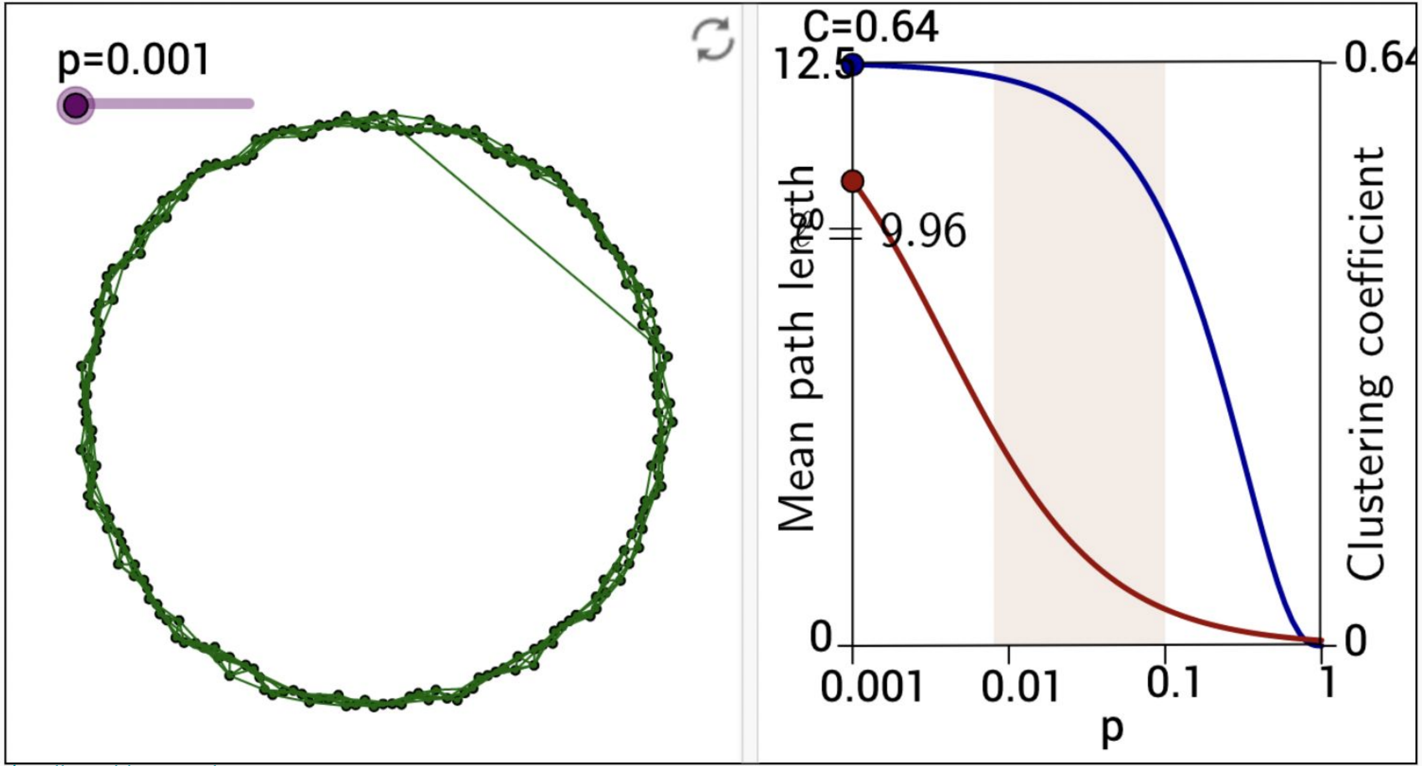


Image source:
https://mathinsight.org/small_world_network

ER/WS random networks don't predict existence of hubs

A few examples of real degree distributions (violet) and degree distribution of a Erdős–Rényi network with the same average degree (green).

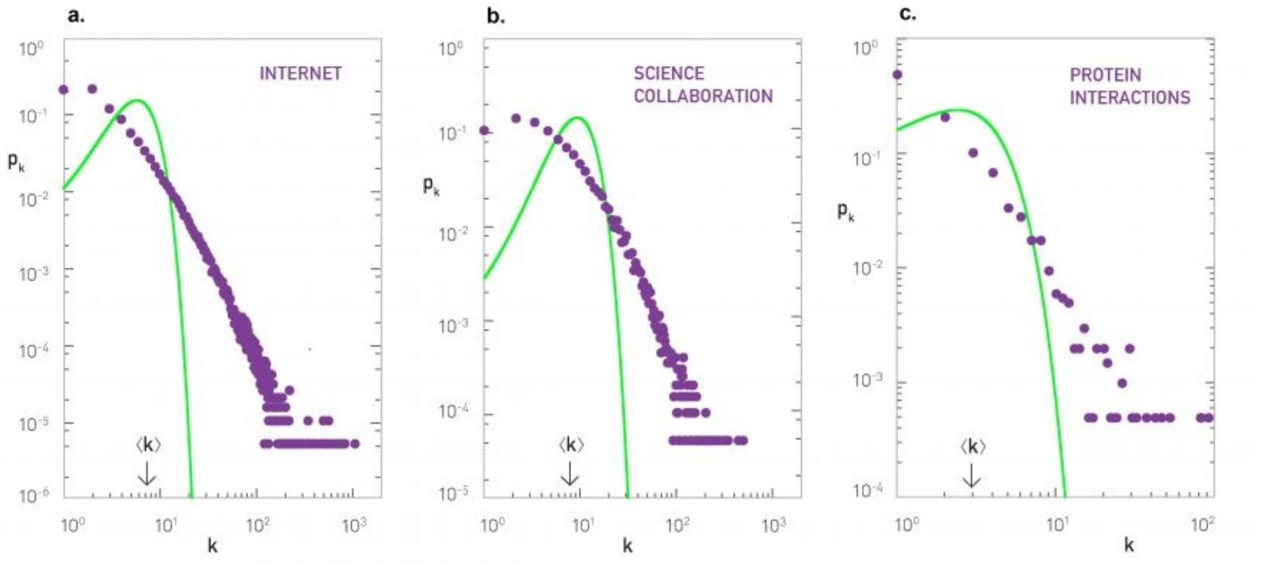


Image source:
http://networksciencebook.com/chapter/3#not-poisson

Preferential attachment

Motivation:

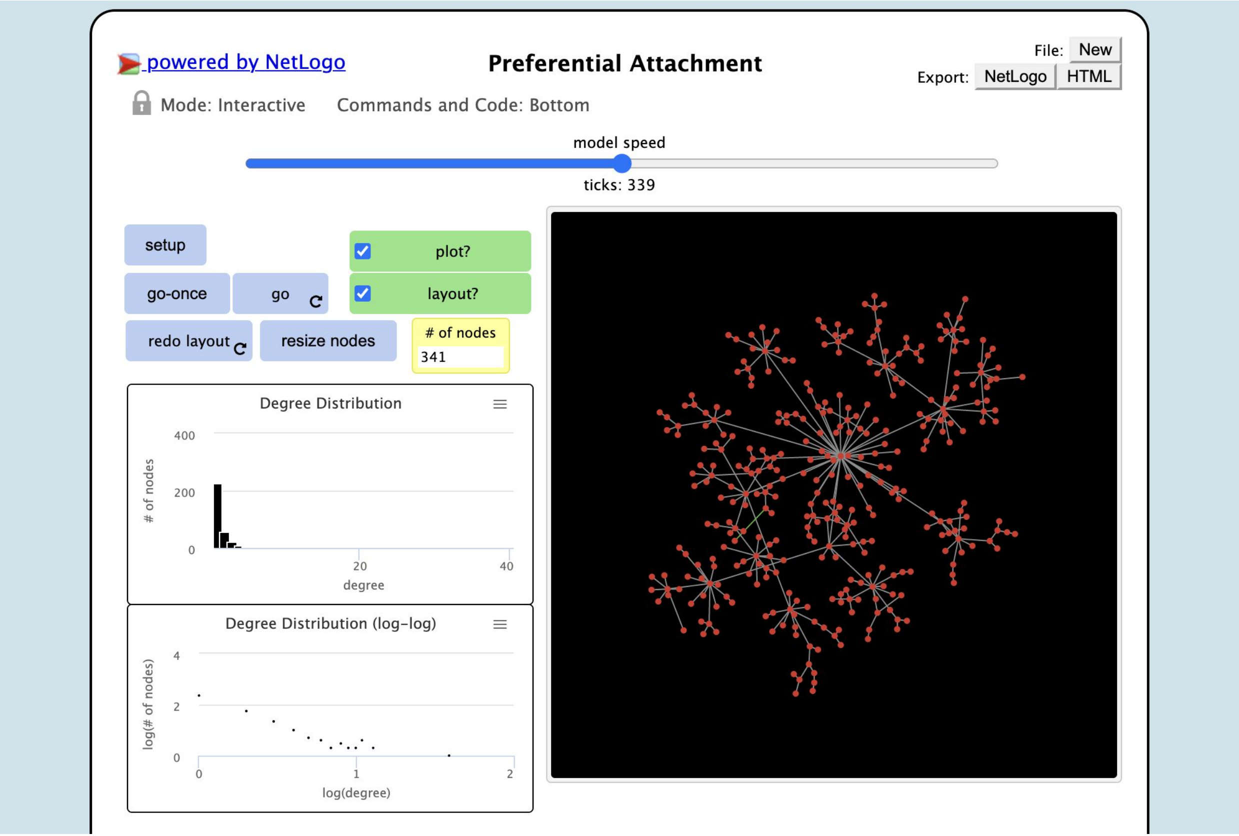
- Regular networks makes assumption all nodes are equal (equal degree)
- With ER networks we've made assumption that existence of any link (between any two nodes) is equally likely
- We managed to recover short paths and clustering coefficients properties we see in some realistic networks, but the heterogeneity in the degrees and particularly the shape of distribution is very different in real networks
- Instead of aiming to describe only the final network and its property, think about how network grew and if some current network properties could be consequences of its growth process

Barabási–Albert network construction

- Start with a regular network of m_0 nodes
- Every time step:
  - New node is added to the graph and connected to m existing nodes($m \leq m_0$)
  - Connections to existing nodes follow preferential attachment rule, e.g. probability to connect to a node i is proportional to node's degree k_i

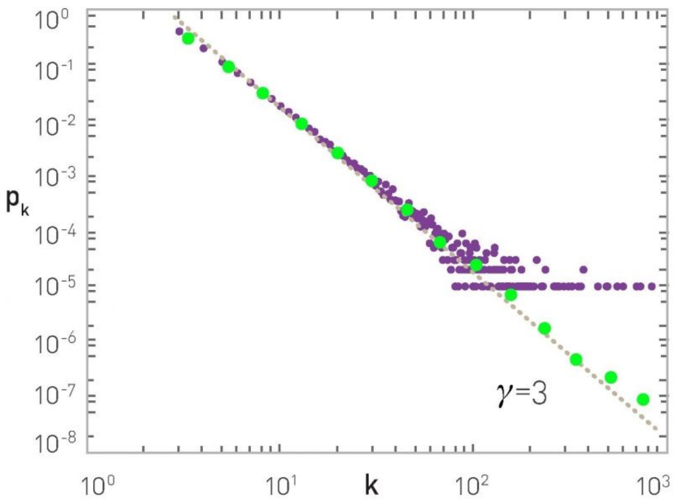
$$p_i = \frac{k_i}{\sum_j k_j}$$

Barabási–Albert model



Properties of Barabási–Albert network

- Number of nodes = t (timesteps needed to grow the network)
- Number of links = mt (every time step, new m links are added)
  - What's average degree?
- Short paths, e.g. small world property persists
- Degree distribution $p(k) \sim k^{-\gamma}$
  - Scale free property
  - Existence of hubs
- But clustering coefficient not explained



Random models take away messages

- Not intended to be predictive models or realistic models of networks, but useful toy examples to guide our network intuition
- They can serve as useful benchmark to understand how much of network properties can be explained by simple characteristics such as number of nodes/links and randomness alone and how much is it left to explain by investigating further
- In case of preferential attachment, the model serves as a useful intuition builder for power law degree distributions but it should not be interpreted as the only mechanisms that lead to scale free property

Further reading

- ER and WS models - chapter 3
  - WS original paper
- BA model - chapter 5
  - BA original paper
- Netlogo
- Complexity explorables

Homework

Submission link