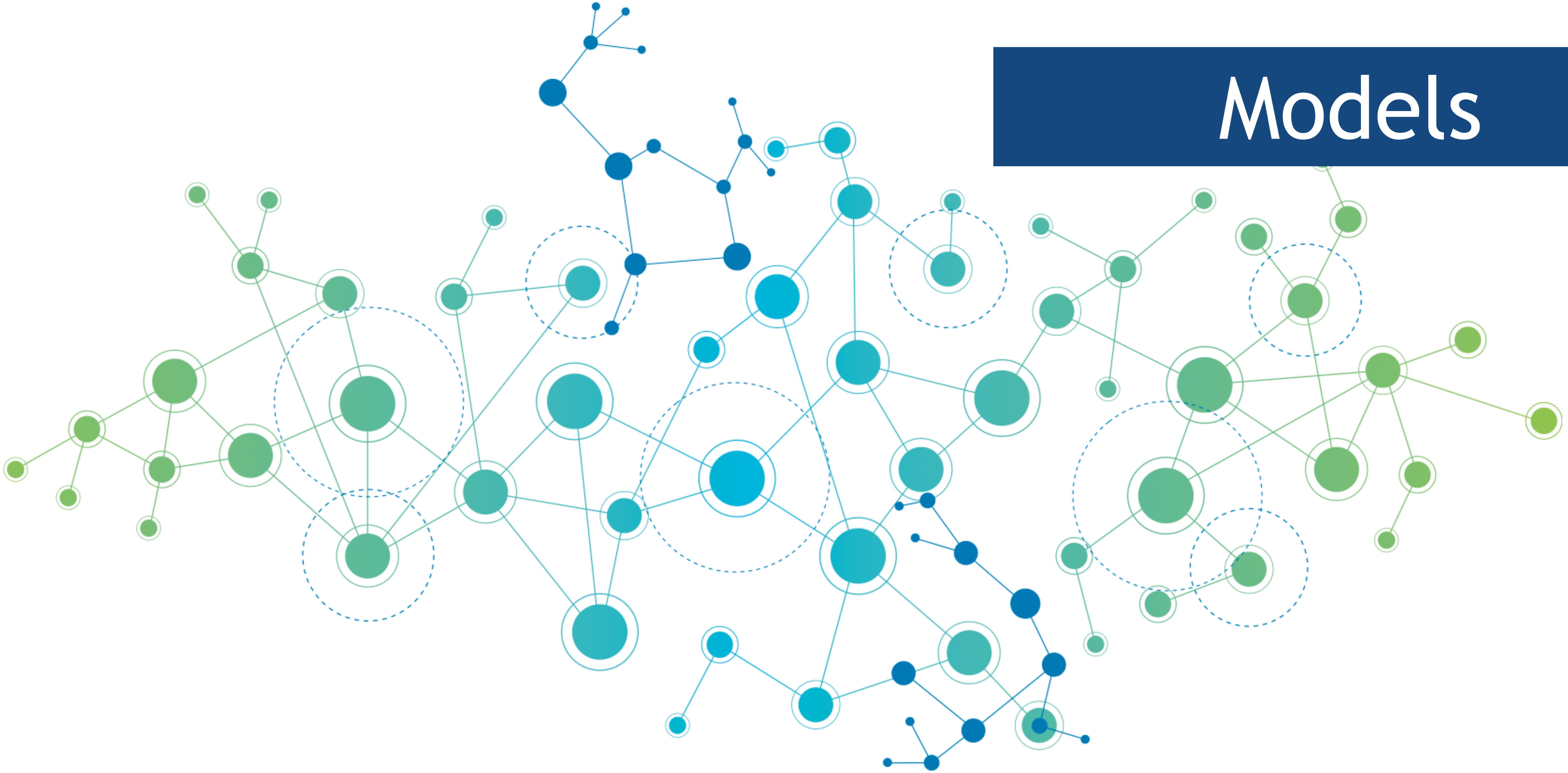
Models



Center for Public Health
Systems Science
Brown School
Washington University in St. Louis

Goals

- Describe challenge of statistically modeling networks
- Explore reasons for modeling networks
- Study three basic network models
  - Erdös-Rényi (random)
  - Small world
  - Scale free

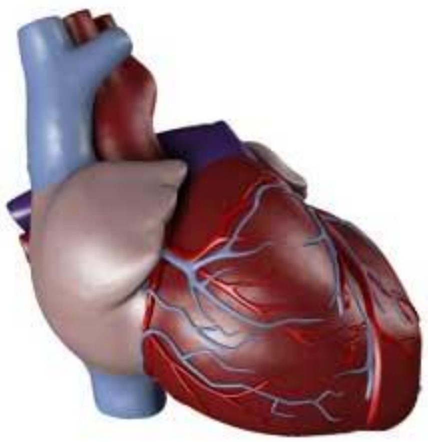
Models

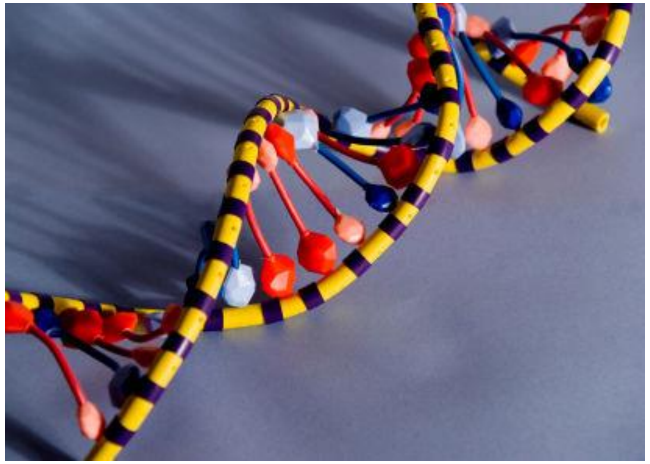
The promise and challenge of models for understanding network structures and processes

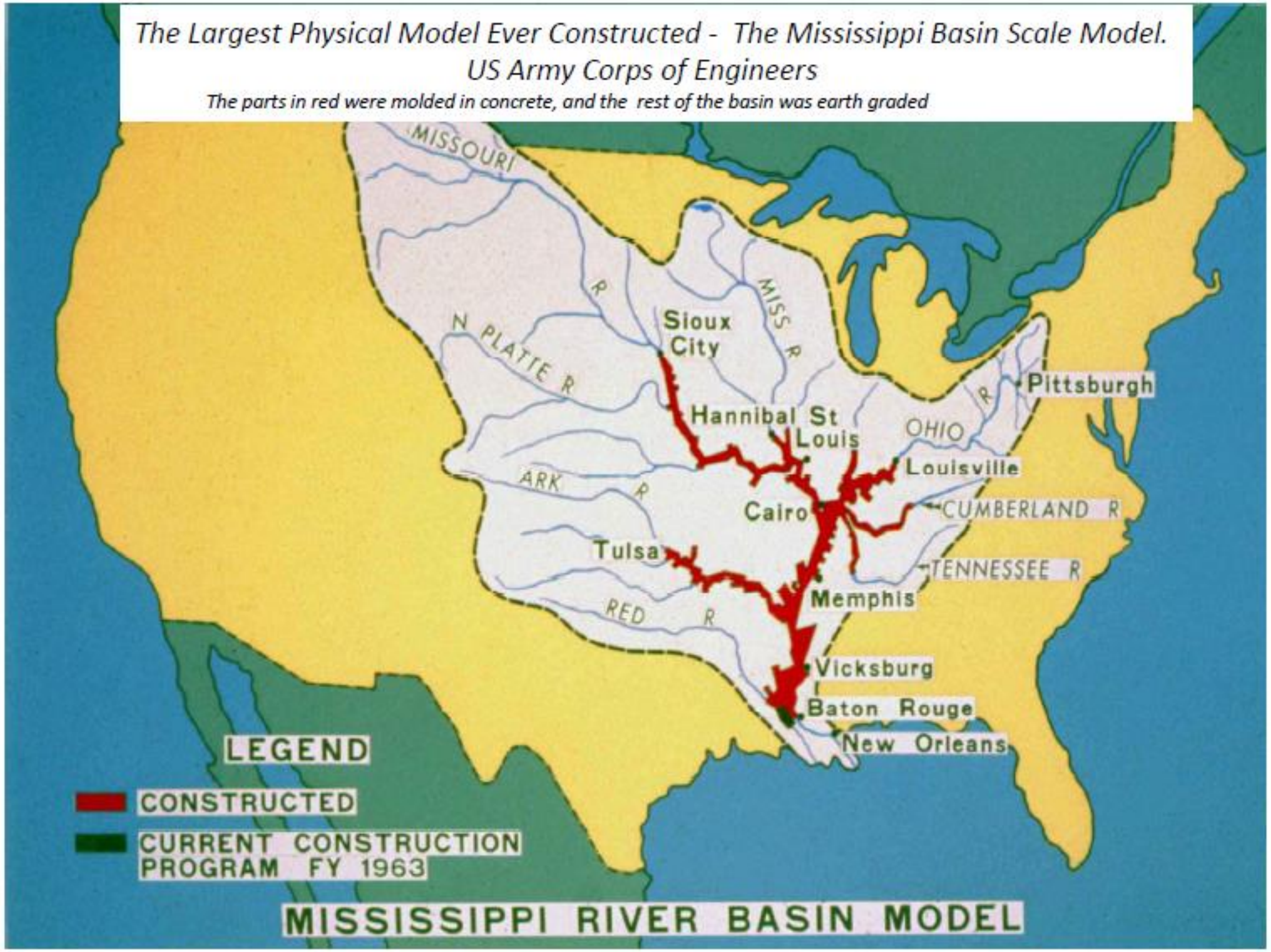
What is a scientific model?

- A simplified representation of a system that is used to explain or predict the behavior of that system
- Models can be:
  - Conceptual, physical, mathematical, statistical, computational

George Box: *…essentially, all models are wrong, but some are useful…*







From Burke, 2013, AAHB Conference, Santa Fe

From Burke, 2013, AAHB Conference, Santa Fe

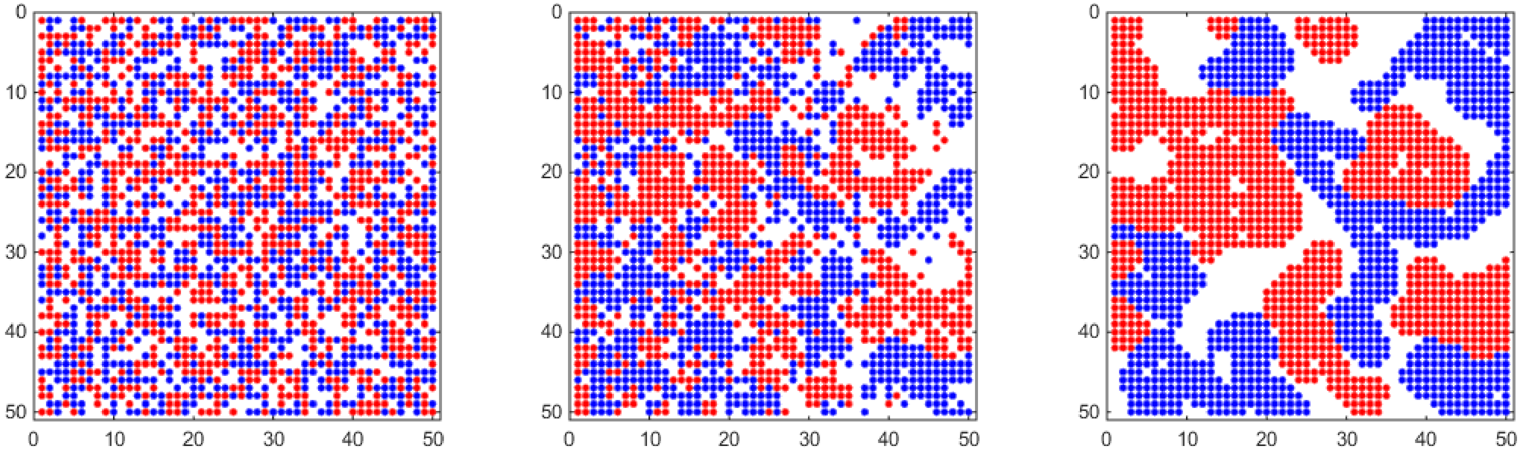
Mathematical and simulation models

2.1. *Preferences and demand*

Preferences are defined over a continuum of differentiated varieties indexed by $i \in \Omega$, and a homogenous good chosen as numeraire. All consumers share the same utility function given by

$$U = q_0^c + \alpha \int_{i\in\Omega} q_i^c di - \frac{1}{2}\gamma \int_{i\in\Omega} (q_i^c)^2 di - \frac{1}{2}\eta \left(\int_{i\in\Omega} q_i^c di \right)^2, \qquad (1)$$

Economics utility model: Melitz and Ottaviano (2008)



Schelling segregation model:
https://camo.githubusercontent.com

Reasons for modeling networks

- Complexity of social behavior.
  - Models allow us to *discover* the regularities in the processes leading to network ties. (Robins, et al, 2007).
- *Describing* real-world networks
  - Models allow us to compare hypothesized networks to actual observed networks
- *Prediction* of future behavior
  - Models allow us to predict future patterns of behavior and social structure

More specific reasons

- Models also allow us to understand the uncertainty (i.e., error) associated with observed outcomes.
- Ability to model complex network processes (e.g., network dynamics)
- Generative models: how do local process, structures, and characteristics form global network patterns?

Formation of network ties

- Networks are not random
- What underlying processes, rules, mechanisms account for the particular patterns of network ties?
- Orienting example:
    - Many networks are scale-free (power-law degree distribution). The Internet is scale-free, for example. What causes this? Some have suggested that this is a 'rich-get-richer' phenomenon, but this still begs the question. Fundamental goal of network modeling is to develop and test theoretical models of network structure and formation.

Network formation concepts

- Proximity/Contact
- Similarity (homophily, assortative/disassortative mixing)
- Reciprocity
- Contagion
- Social Exchange/Capital
- Transitivity
- Structural Equivalence

Three important models

Random networks, Small-world networks, Scale-free networks

Random graphs

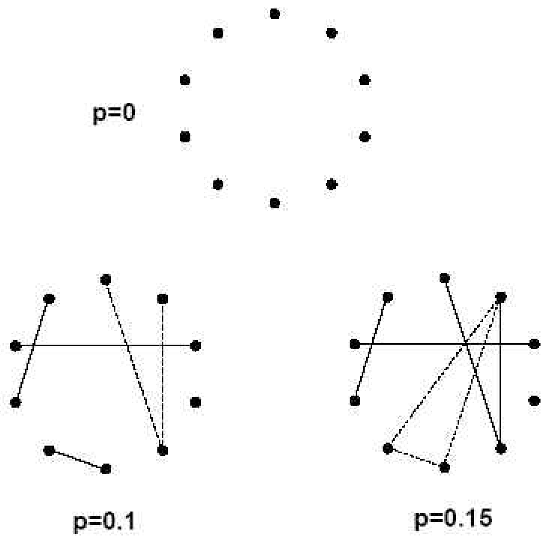
- Erdös-Rényi random graphs
- Random graph – take a set of nodes (vertices), and randomly connect pairs of vertices with ties (edges)
  - Important finding – only a small percentage of pairs need to be randomly connected before the entire graph is connected, and this percentage goes down as the network gets larger
  - E.g., for the Earth's population, people only need to have around 24 connections to have the entire planet included in one social network

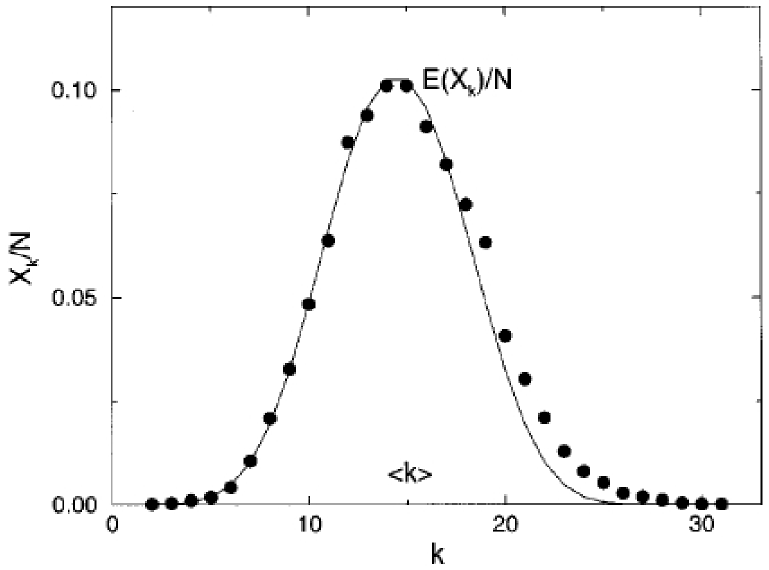
Sidebar – Paul Erdös

- Most prolific mathematician in history
- Itinerant vagabond
- Erdös numbers
  - http://www.oakland.edu/enp/

http://www.public.asu.edu/~mmcbeath/mcbeath.lists/erdos.html

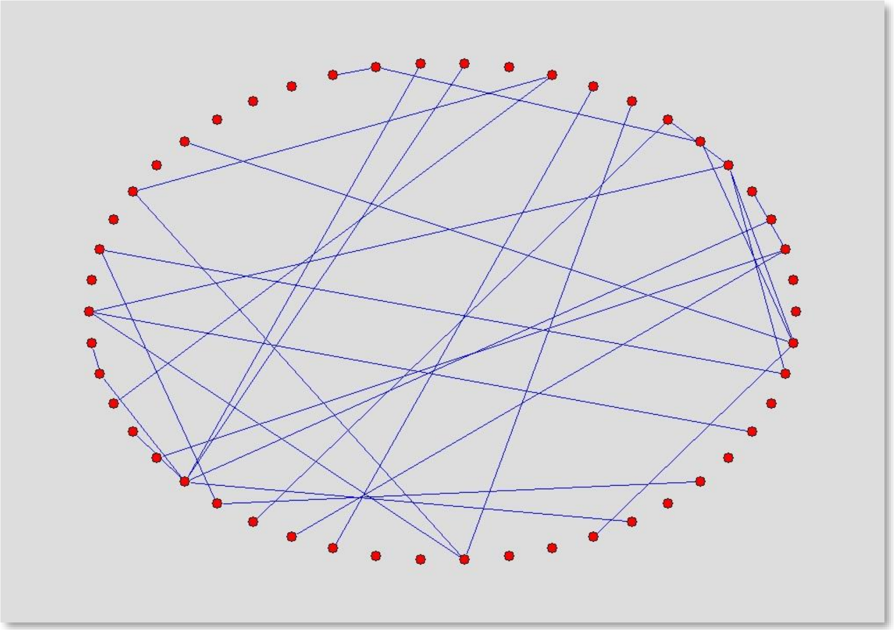
Erdös-Rényi random graphs



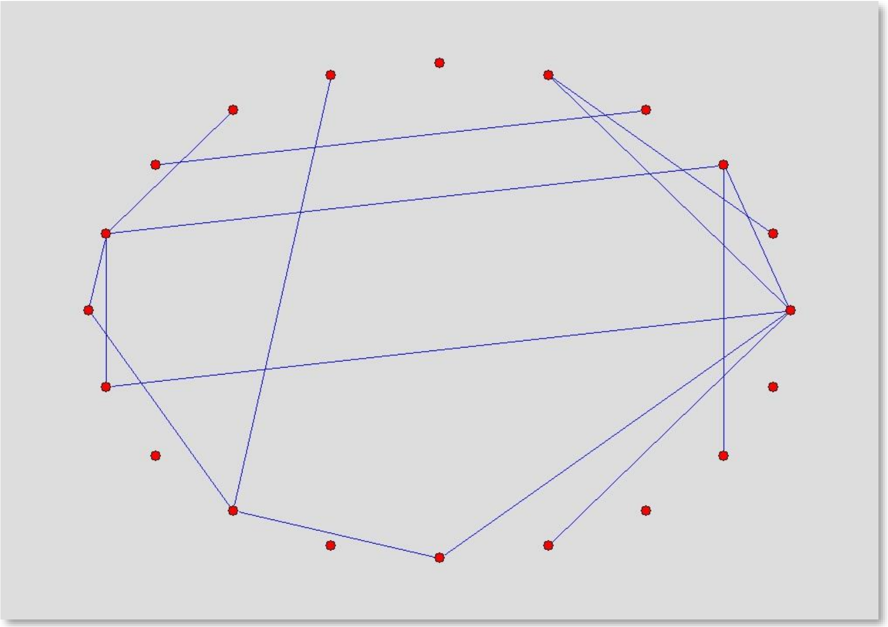


Erdös-Rényi – 3 sizes
Avg. degree = 1.5

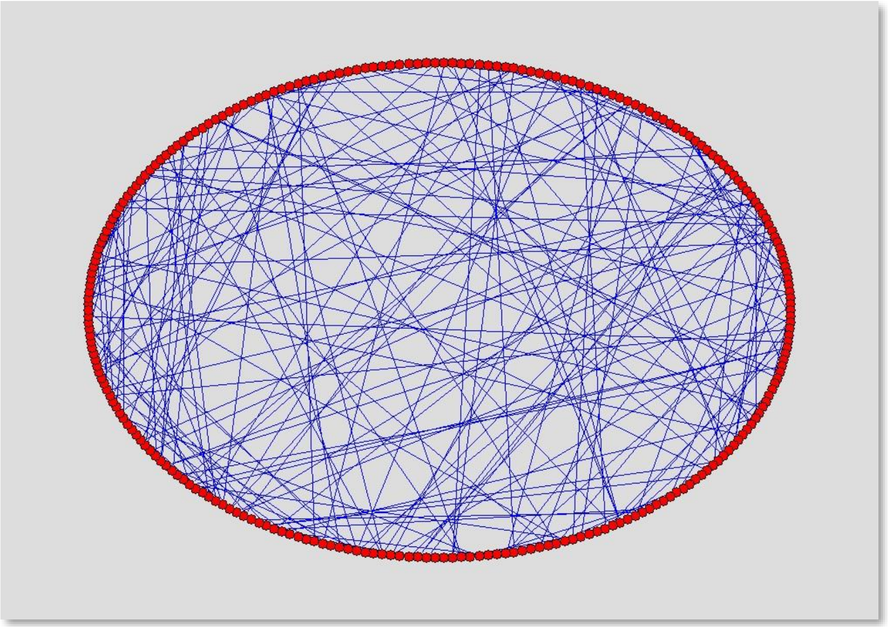
(k = 50, density = .024)



(k = 20, density = .075)



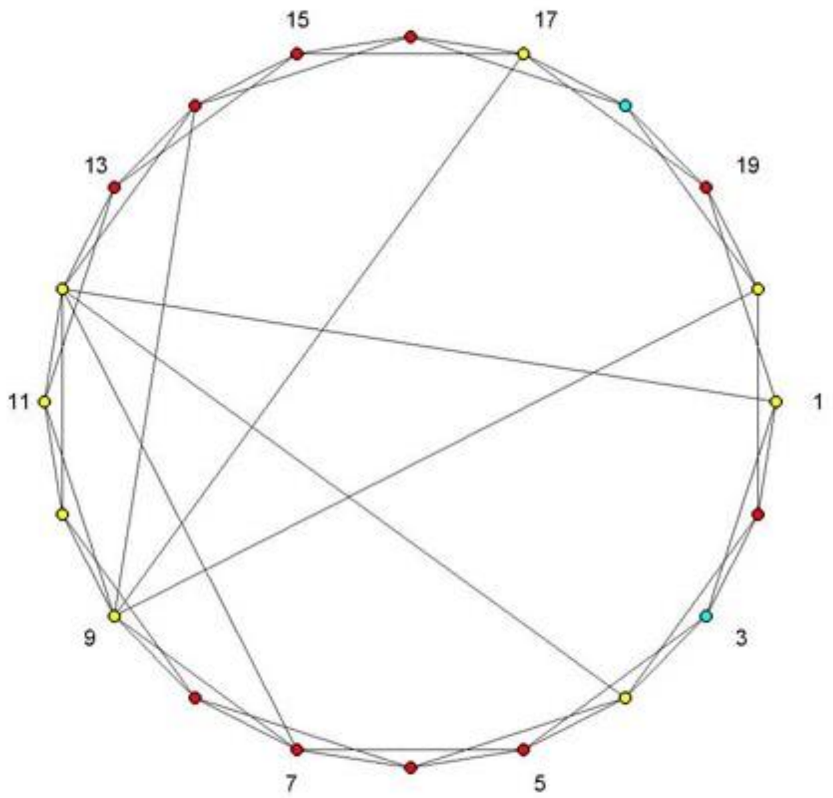
(k = 250, density = .006)



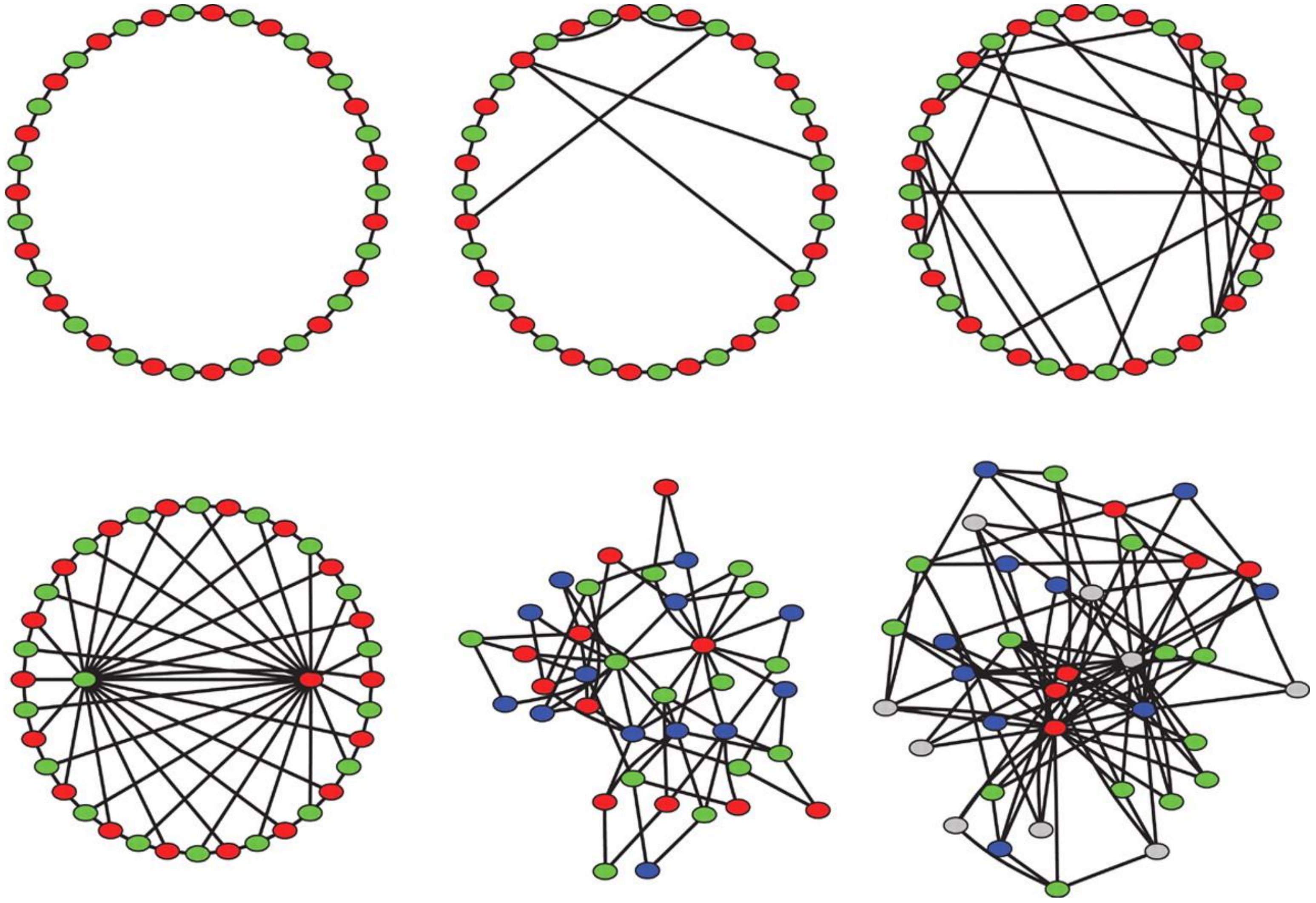
Small-world model

- Watts, Strogatz – mathematicians trying to understand how large, widely distributed networks could show coordinated activity
  - e.g., firing neurons, chirping crickets, clapping crowds
- Discovered that large structured networks suddenly exhibit ‘small-world’ properties when a small number of random ties are added to the structured network

Example – small world network



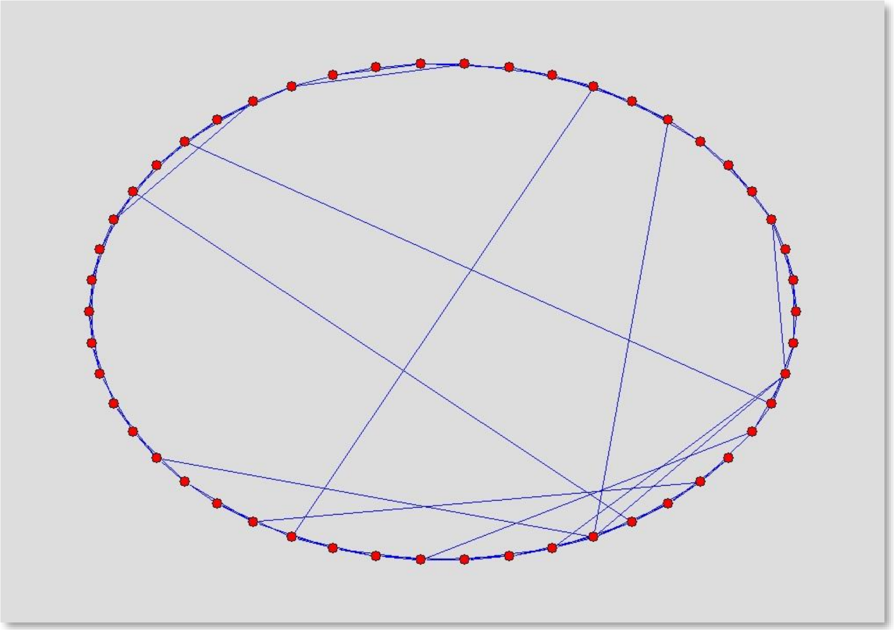
Moving from lattices to more realistic networks



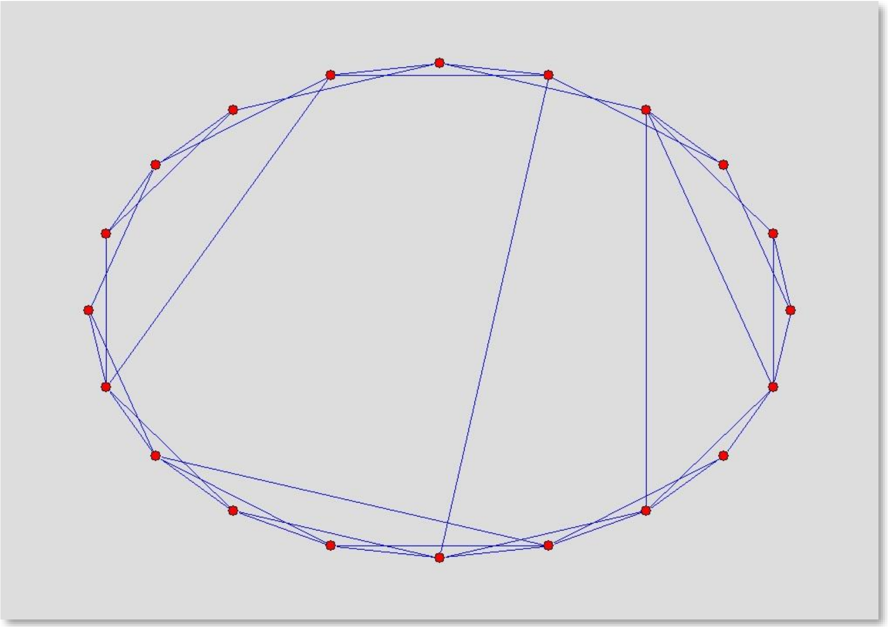
Small world – 3 sizes

Nodes connect to 4 neighbors,
Rewiring = 0.1

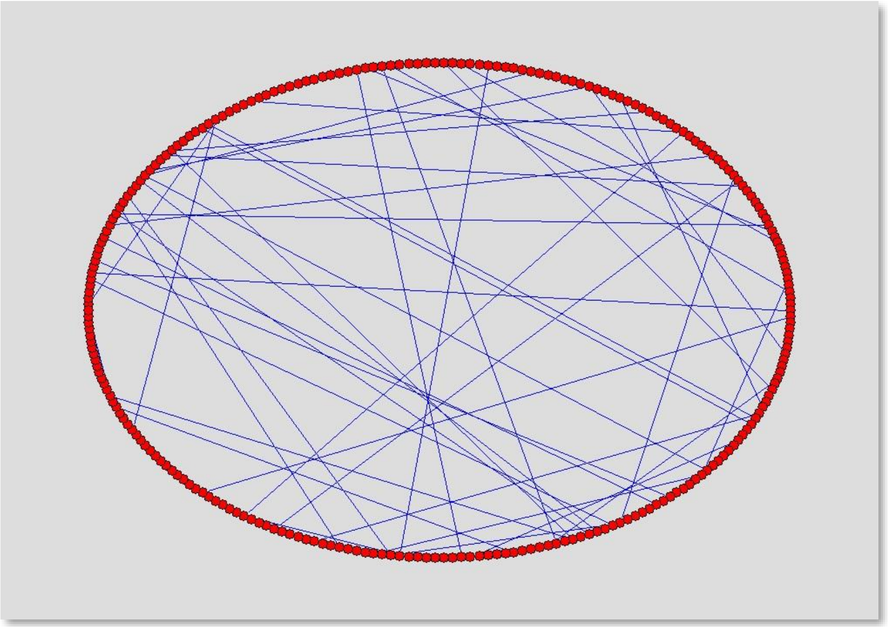
(k = 50, d = .08, dia = 8)



(k = 20, d = .21, dia = 5)

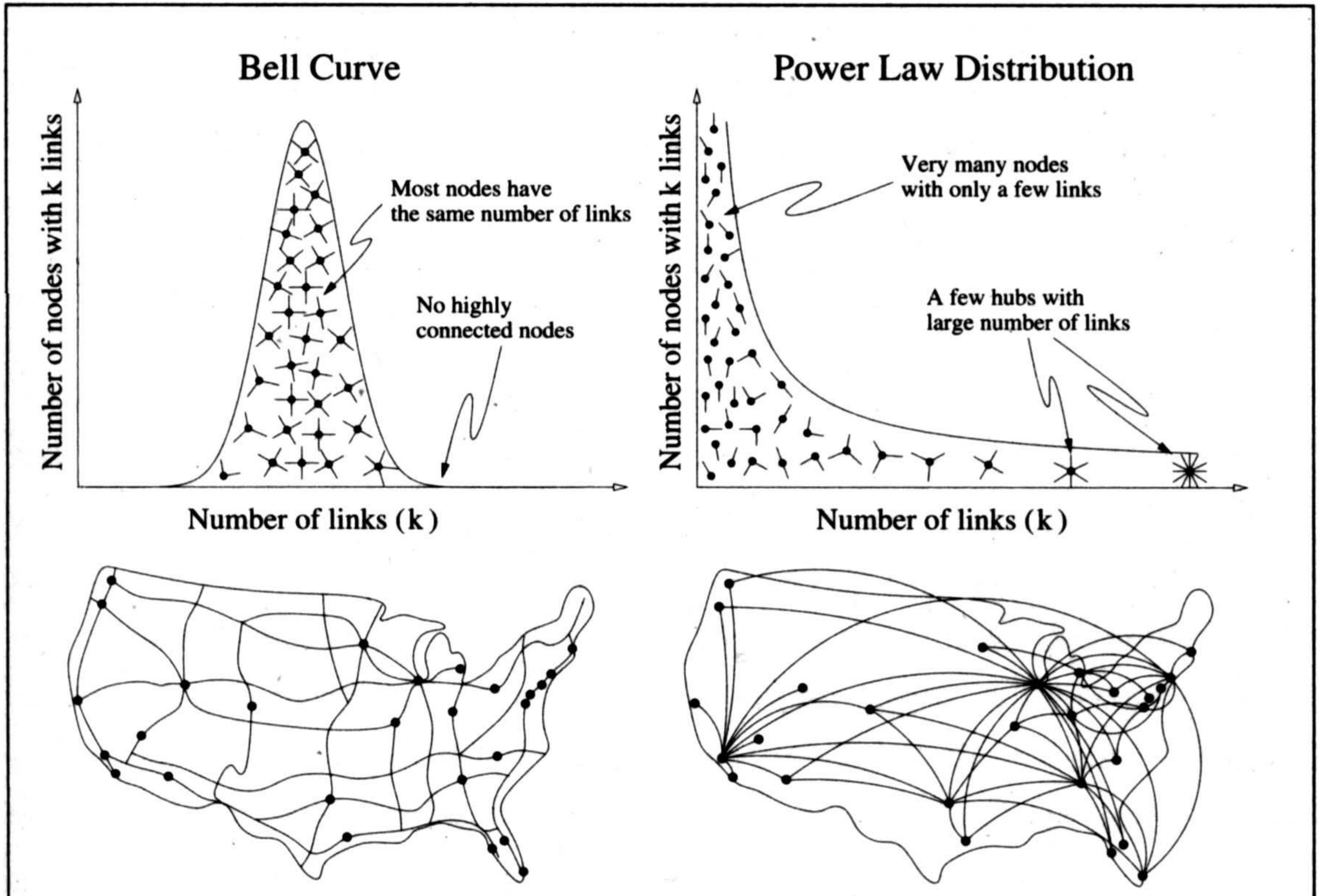


(k = 250, d = .02, dia = 13)



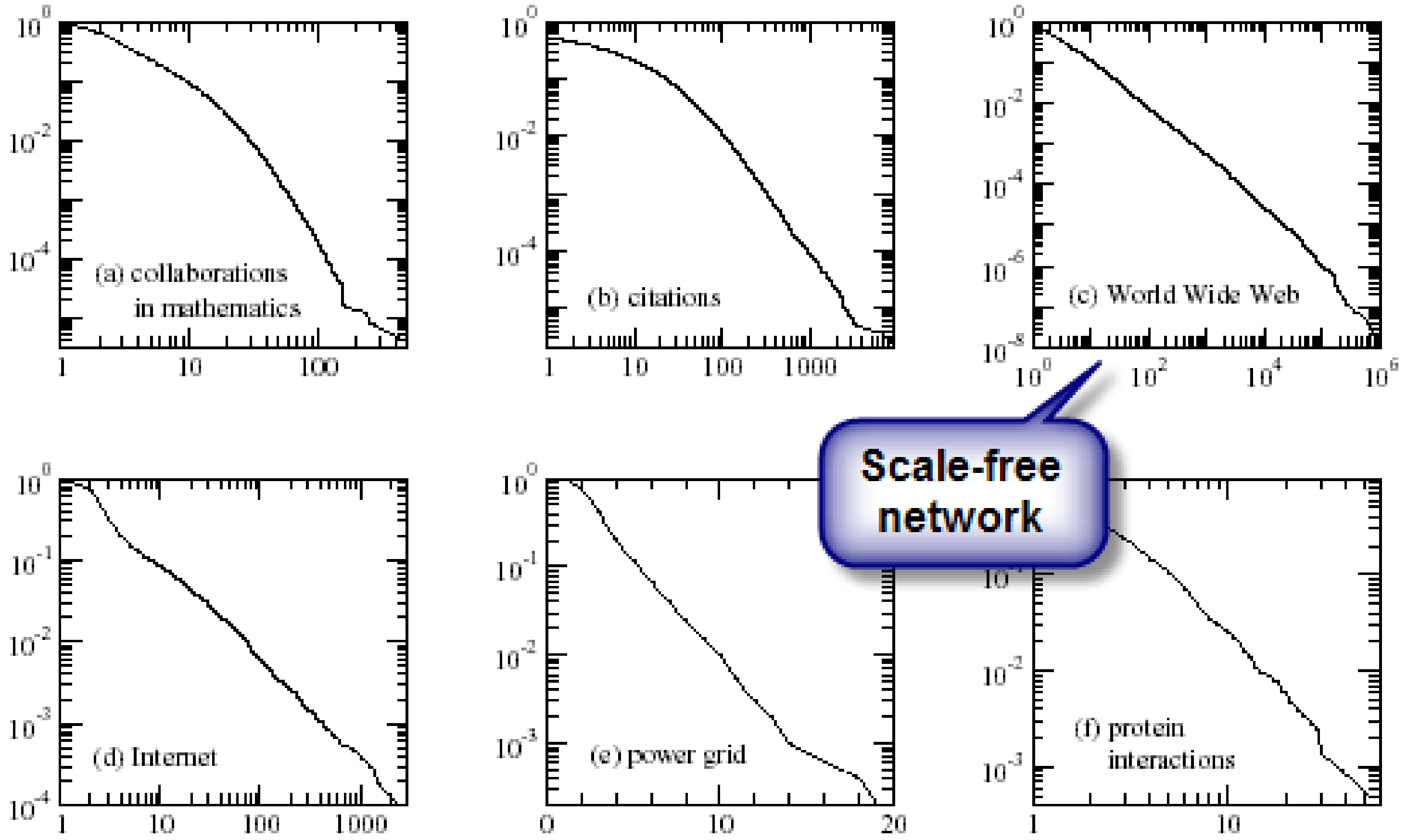
Scale-free networks

- Distribution of node degree follows a power law (scale free)
- Current understanding is that many networks grow according to a process of cumulative advantage (‘rich get richer’), this leads to the distinctive power law relationship
- See NetLogo model example



Reprinted from Linked: The New Science of Networks by Albert-Laszlo Barabasi

Degree distributions for some real networks



From Newman (2003)

Sex is scale-free

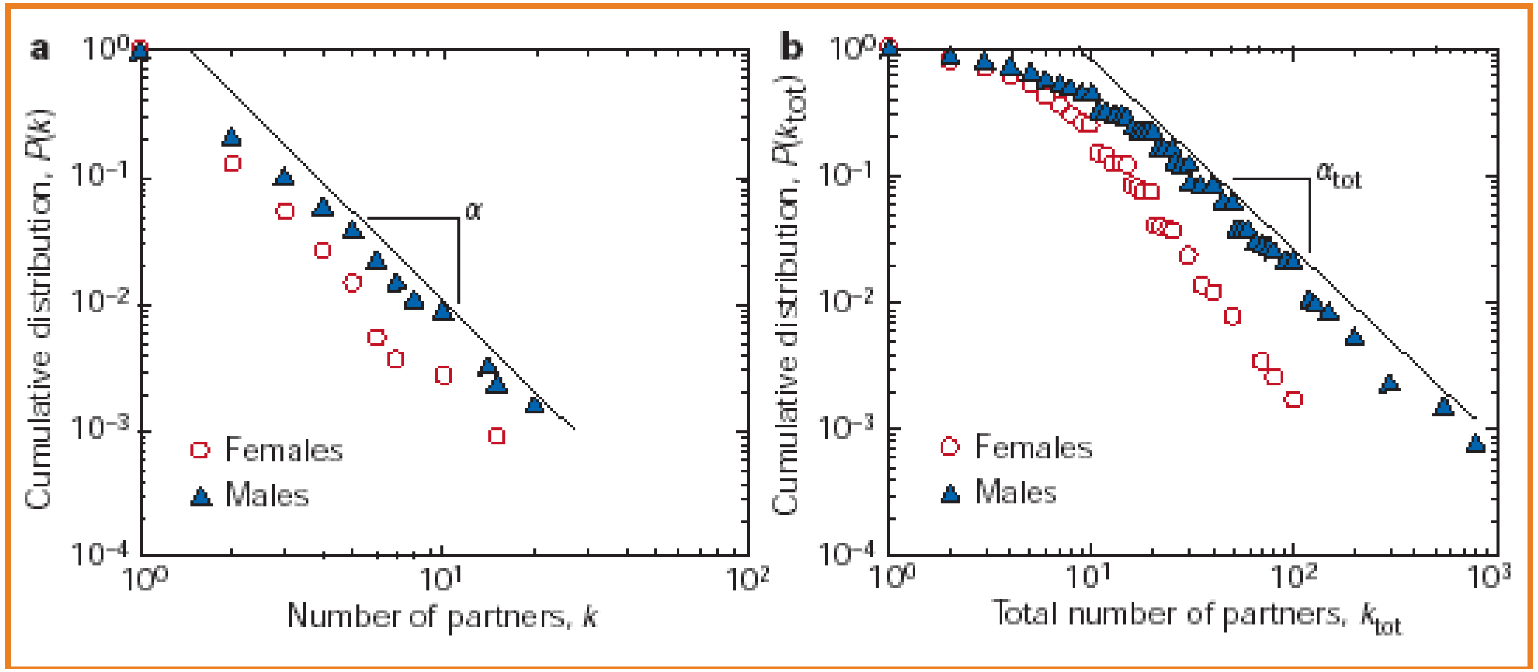


Figure 2 Scale-free distribution of the number of sexual partners for females and males. **a**, Distribution of number of partners, k, in the previous 12 months. Note the larger average number of partners for male respondents: this difference may be due to 'measurement bias' — social expectations may lead males to inflate their reported number of sexual partners. Note that the distributions are both linear, indicating scale-free power-law behaviour. Moreover, the two curves are roughly parallel, indicating similar scaling exponents. For females, $\alpha = 2.54 \pm 0.2$ in the range $k > 4$, and for males, $\alpha = 2.31 \pm 0.2$ in the range $k > 5$. **b**, Distribution of the total number of partners k_{tot} over respondents' entire lifetimes. For females, $\alpha_{tot} = 2.1 \pm 0.3$ in the range $k_{tot} > 20$, and for males, $\alpha_{tot} = 1.6 \pm 0.3$ in the range $20 < k_{tot} < 400$. Estimates for females and males agree within statistical uncertainty.

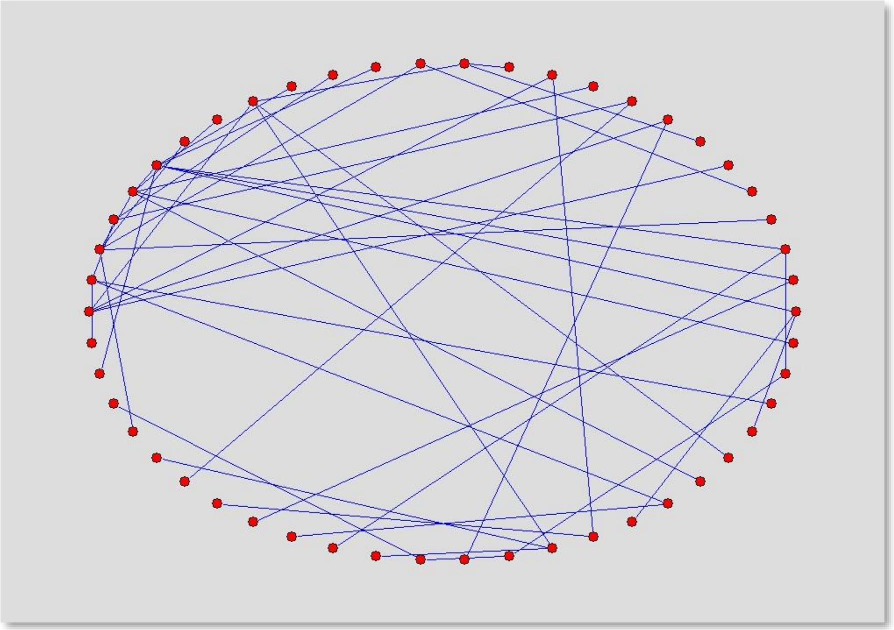
From Liljeros, et al. (2001). The web of human sexual contacts. *Nature, 411,* 907-908.

Scale free– 3 sizes

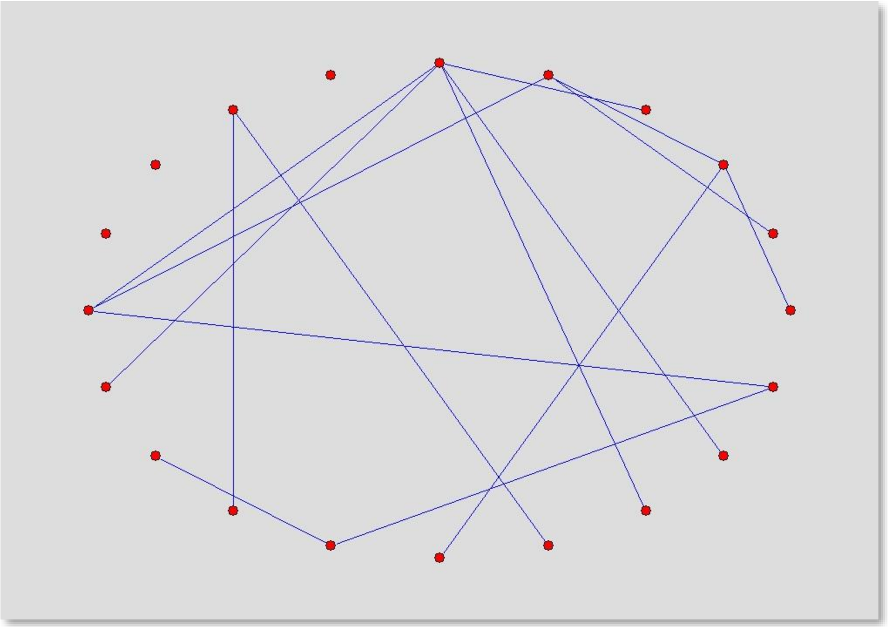
Avg. degree = 1.5,

Alpha = 0.25

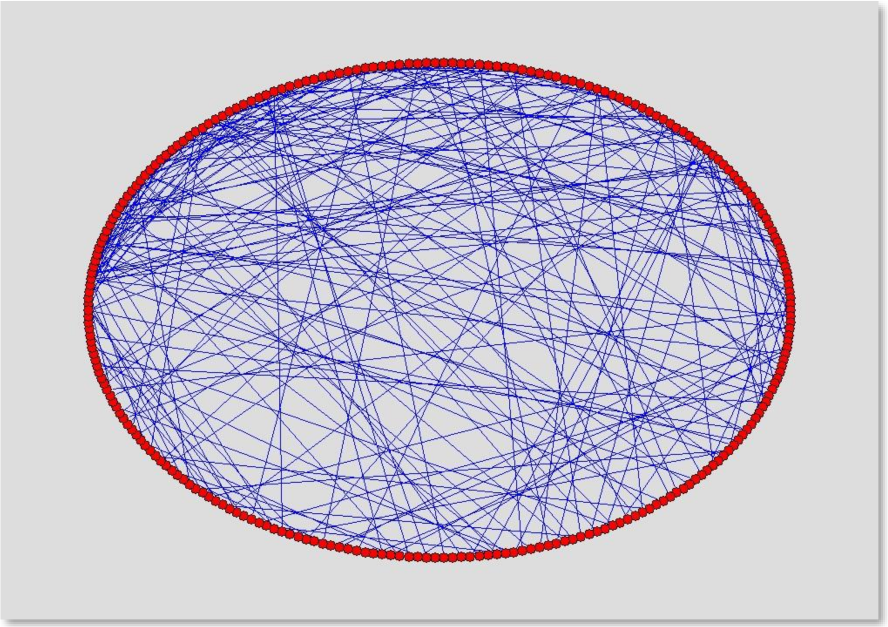
(k = 50, d = .04, high d = 7)



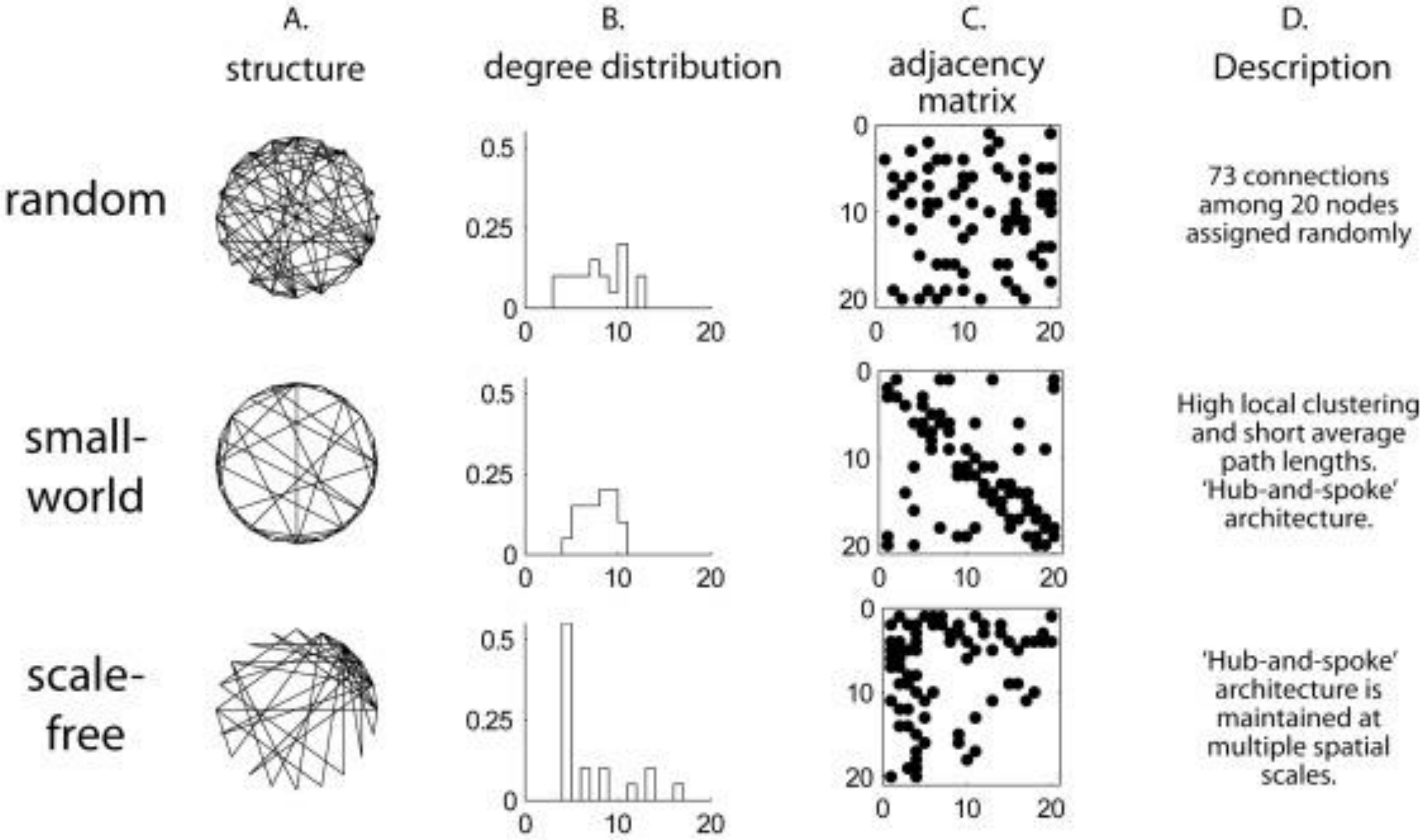
(k = 20, d = .08, high d = 5)



(k = 250, d = .01, high d = 13)



Different models imply different network characteristics

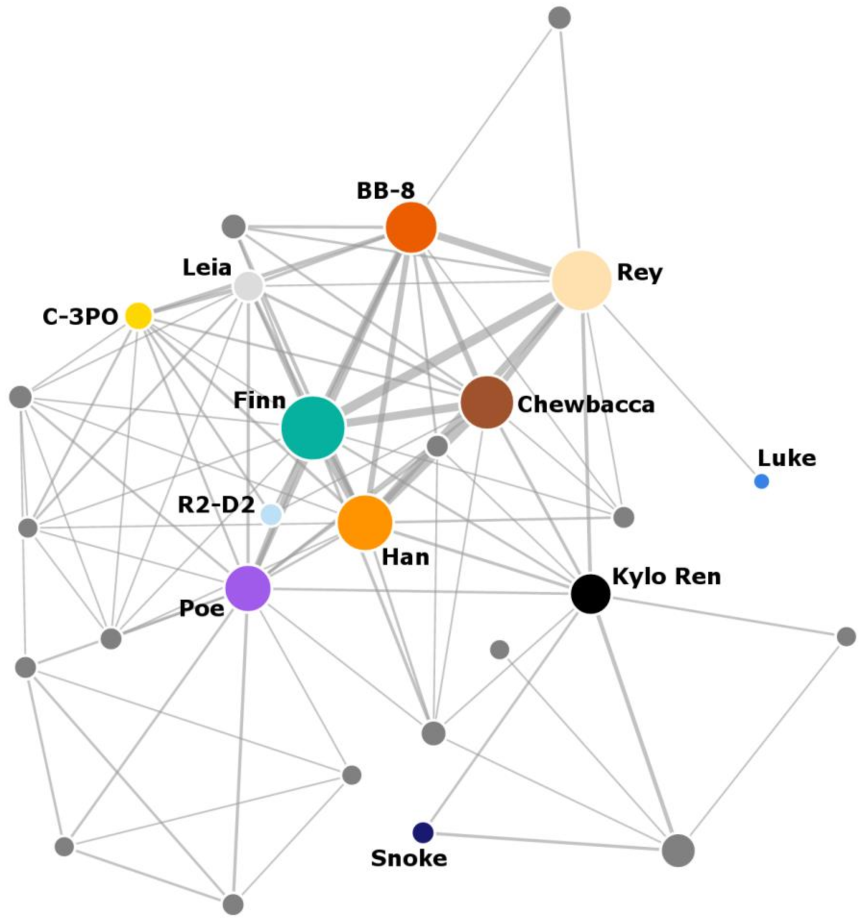


From Stobb, et al., (2012). PLoS ONE 7(5):e37292

Challenge of modeling 'realistic' social networks

Table 1. Properties of social networks, compared to three network generation models.

Property	Social Network	Random Net	Preferential Attachment	Small World
Low average distance	✓	✓	✓	✓
Moderate clustering coefficient	✓			✓
Power laws for degrees	✓		✓	



Dekker, A. H. (2007, May). Realistic social networks for simulation using network rewiring. In *International Congress on Modelling and Simulation* (pp. 677-683).

http://evelinag.com/blog/2016/01-25-social-network-force-awakens/

Network model functions in R

- *Statnet*
  - rgraph
  - ergm (statistical modeling, more on this later)
- *igraph*
  - sample_gnm, sample_gnp
  - sample_smallworld
  - sample_pa, sample_pa_age
  - (plus many more)