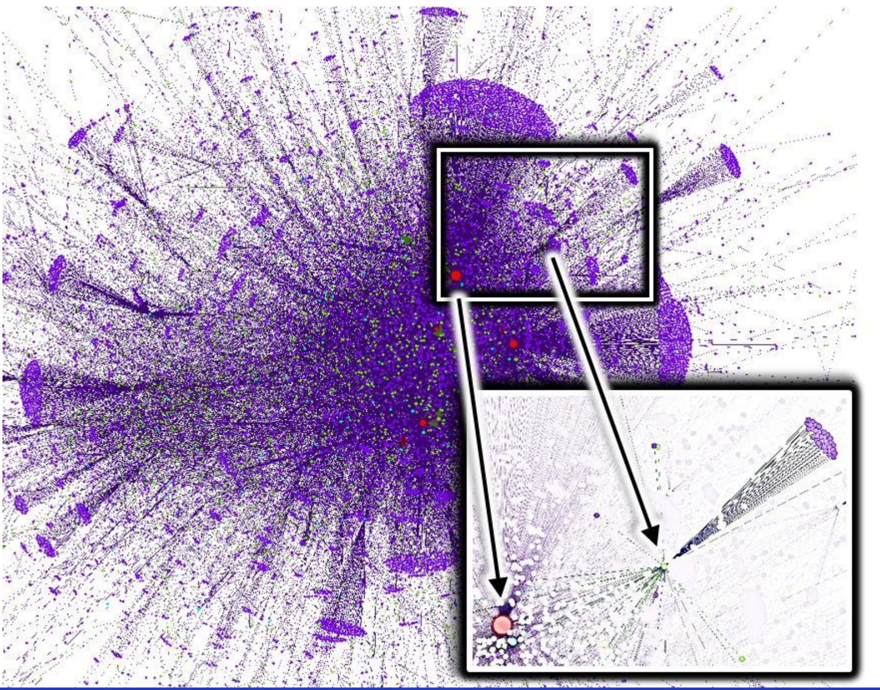
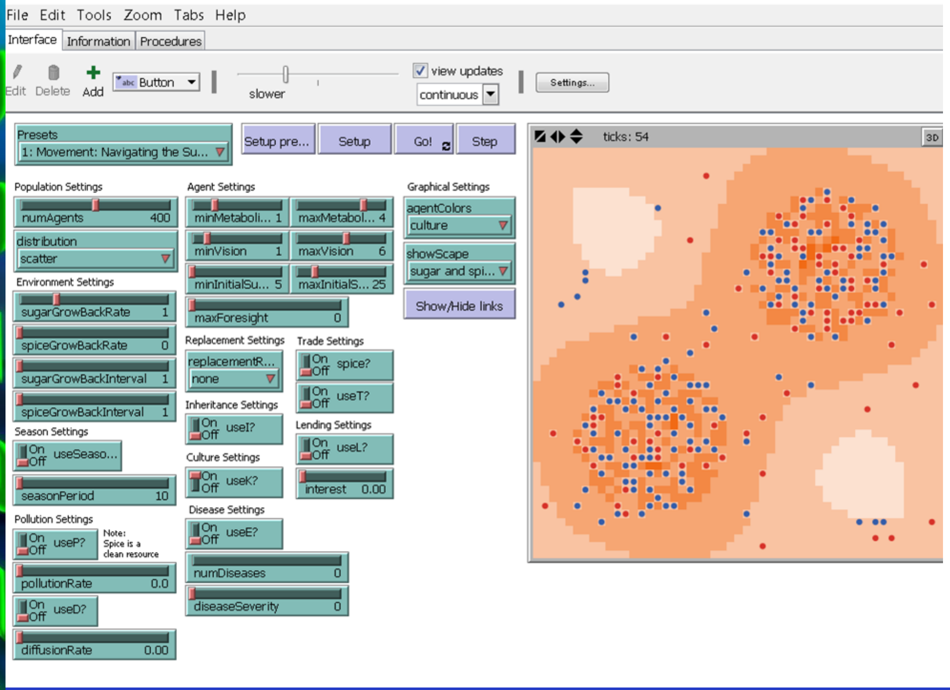




Complex Social Systems

a guided exploration of concepts and methods

SOCIAL NETWORK ANALYSIS (SNA part2)

Network Dynamics

Martin Hilbert (Dr., PhD)

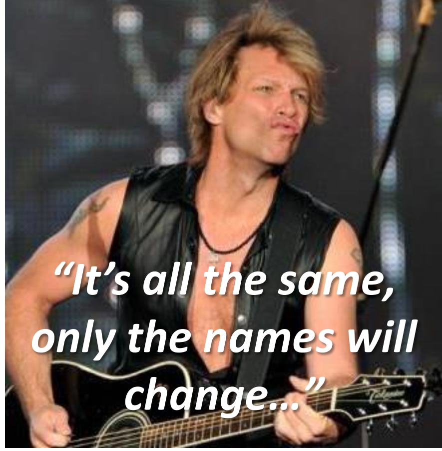
Today's questions

I. How can we find out if our model of the network is correct?

II. How can we model network dynamics over time?

III. How can we use networks to make predictions?

Network Science



Description

Find a real-world phenomenon and describe its properties

Online friend network with different global, local and position measures

Look for other properties

Analysis

Analyze these properties and look for patterns in the noise

Some have many more friends than others

Adjust theory

Science

Theory

The ones that play sports have more friends

Hypothesize a generative mechanism of the functionality / *modus operandi* that produces this phenomenon

?

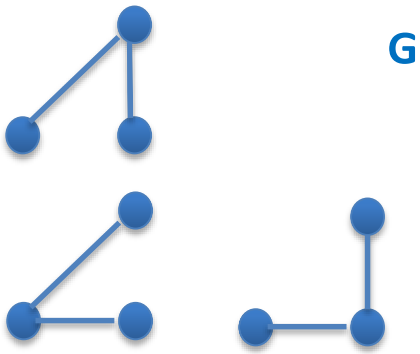
Might it be just random chance?

Modeling / Testing

How to test if our ideas are correct or if we are just daydreaming?

Erdos-Renyi (1959-1960) Random Graphs

benchmark **G(n,p)** or **G(n,M)**



G(n,M) = G(3,2)

4 degrees
3 nodes
= 1.33…

- Start with **n** nodes
- Form independent links with probability for each node **p**
- Form **M** independent links
  - ⇒ If we want to compare these random networks to our real-world network, we need to know the properties of the random network

- Property examples:
  - What's the average degree of the network?
    - *numerical solution => simulate and count!*
    - analytical solution => ???

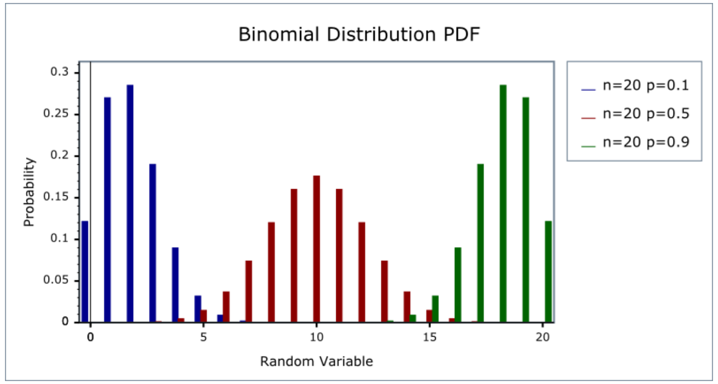
G(n,p) = G(3 , $^2/_3$)

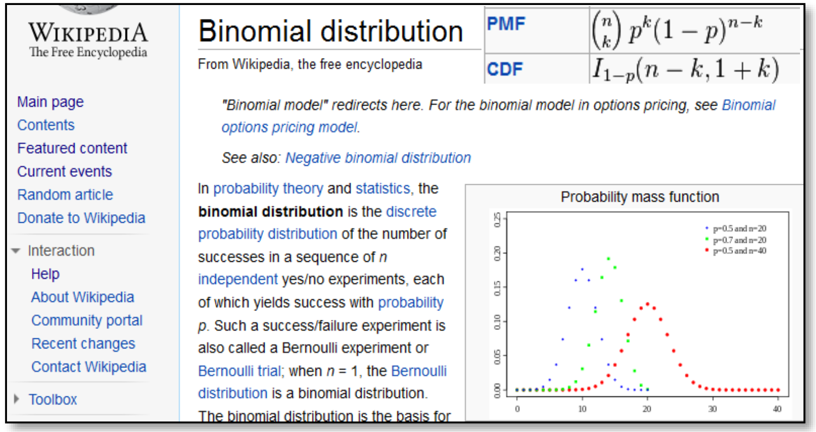
Avg. degree = (n-1)*p

(3-1) * $^2/_3$ = 1.33…

- The nice thing is that we get "almost sure" properties with specific G(n,p) setups
  - What is the average degree of the random network?
  - Does the average degree stay constant, increase, or decrease with number of nodes?
  - What's the degree distribution of the network?
  - What's the likelihood of getting a hub with almost all links concentrated on it alone?

What's the avg. degree of G(n=9; p=0.5)?





WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Binomial distribution

From Wikipedia, the free encyclopedia

PMF	$\binom{n}{k} p^k(1-p)^{n-k}$
CDF	$I_{1-p}(n-k, 1+k)$

"Binomial model" redirects here. For the binomial model in options pricing, see Binomial options pricing model.

See also: Negative binomial distribution

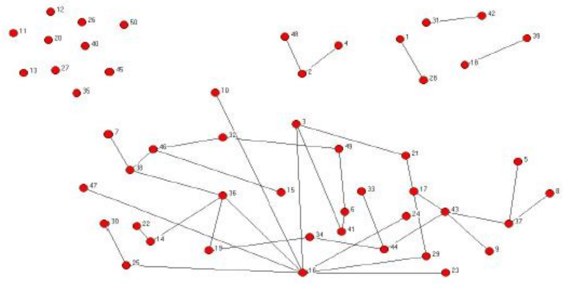
In probability theory and statistics, the **binomial distribution** is the discrete probability distribution of the number of successes in a sequence of *n* independent yes/no experiments, each of which yields success with probability *p*. Such a success/failure experiment is also called a Bernoulli experiment or Bernoulli trial; when *n* = 1, the Bernoulli distribution is a binomial distribution. The binomial distribution is the basis for

Probability mass function

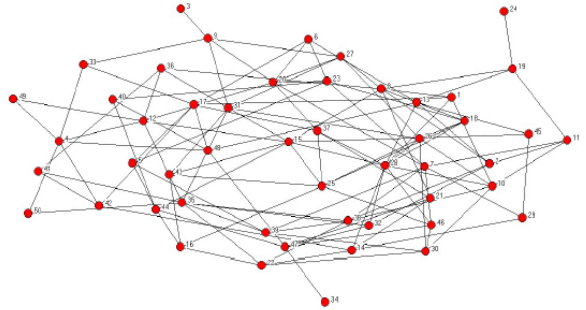
➢ Erdos-Renyi Threshold functions

- The likelihood of getting a property (or not) depends (almost surely) on p (such as avg. degree distribution), but ***non-linear phase transitions emerges*** at certain ***tipping points***: *"more is different"* or *"…merely quantitative differences beyond a certain point pass into qualitative changes…"*
  - e.g.: with $p > {}^{1}/_{n}$ => network WILL have cycles & one unique giant component (with $n^{2/3}$ nodes)!
  - e.g.: with p = ln(n)/n => network WILL be fully connected!

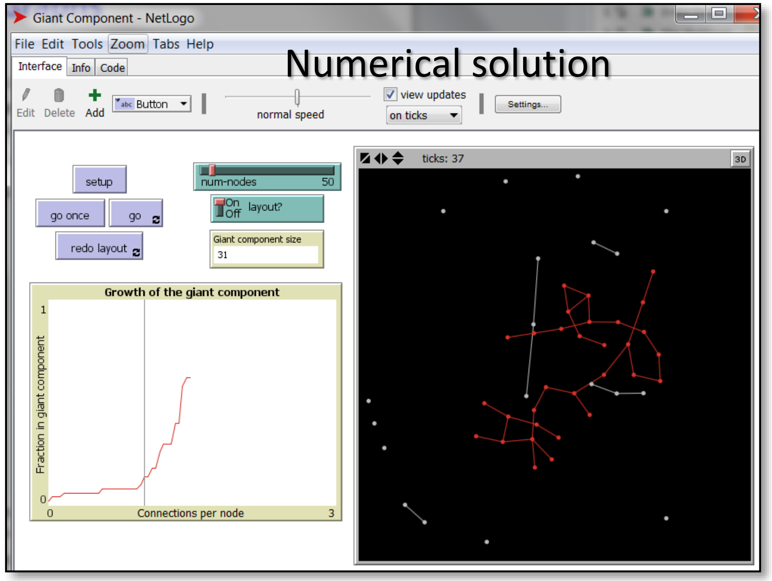
Poisson n = 50; p = 0.03 > $^{1}/_{50}$ = 0.02



Poisson n = 50; p = 0.1 > ln(50)/50 = 0.08

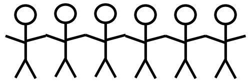


Numerical solution



At what avg. degree does a giant component emerge?

How many friends do your friends need to have for information / disease to spread quickly through most of the network?



Create a random network G(n,p) with the property your are interested in
(proved by some ingenious analytical theorem or by numerical averages)
and compare it to your empirical network!!!

Source: Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera. Wilensky, U. (2005). NetLogo Giant Component model. Center for Connected Learning and Computer-Based Modeling.

Network Formation: growing networks following other rules

➢ Add nodes with uniform likelihood

With Erdos-Renyi we have all nodes given: here we add nodes...

- Start with m nodes fully connected
- Add new nodes with k links to existing nodes with equal likelihood
- At each step, an existing node has a probability [k / #of nodes of getting new link]

$\Rightarrow$ Older nodes will have more links (*"…el viejo diablo…"*)

➢ Add nodes with non-uniform likelihood

- In reality some nodes have many more links than others ("fat-tail" distributions)
- **Preferential Attachment** / Rich get richer
- Likelihood is proportional to number of existing degrees
  - New nodes adds m links to existing nodes

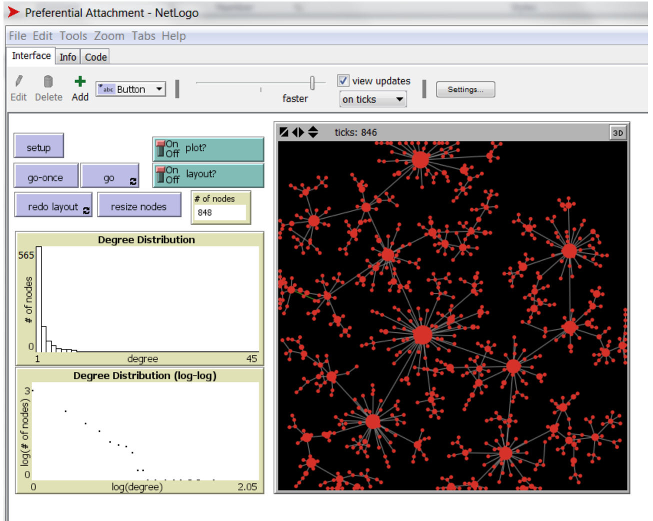
  $\Rightarrow$ At time t, there will be t*m links

  $\Rightarrow$ Total number of degrees = t*m*2

  (each link provide two degrees, one for each connected node)

  $\Rightarrow$ Probability of attaching to node j = $^{[\text{degree of j}]} / _{[2*t*m]}$

How many hubs form?
Can smaller hubs every catch up?



Source: Yule, G. U. (1925). "A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S". Philosophical Transactions of the Royal Society B: Biological Sciences 213 (402–410): 21–23. Price de Solla, D. J. (1965). Networks of Scientific Papers. Science, 149(3683), 510–515. Barabasi, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509–512.

Network Formation: growing networks following other rules

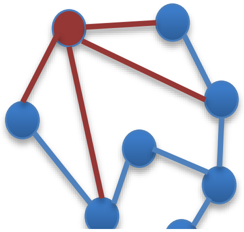
- **"Hybrid" models:** **(between uniform random and preferential attachment)**
  - New nodes creates a fraction of its links with other nodes uniformly at random
    - ***the other fraction*** by looking at the "friends of the friends" of the connected nodes

    $\Rightarrow$ Friends of friends are more likely highly connected

    the more connected nodes are **more often** friends than low degree nodes, simply because they have more links!

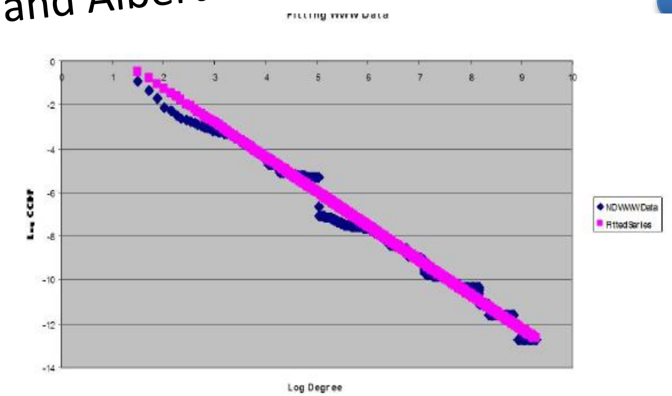
    $\Rightarrow$ The 2nd part is like preferential attachment

    nodes with more connections are easier to find and will get more friends



The best fit of www data has more than 1/3 at random and only 2/3 preferential attachment of Barabasi and Albert

log(freq)



a=.36

R^2=.97

log(degree)

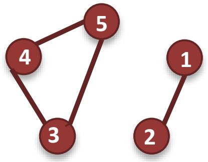
OR:

- …***the other fraction*** by considering "geographic proximity"… etc. etc…

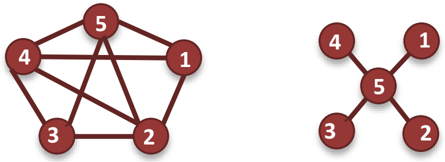
Source: Jackson, M.O., and B.W. Rogers (2007) "Meeting Strangers and Friends of Friends: How Random Are Social Networks?" *American Economic Review 97(3):890–915.*
Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera.

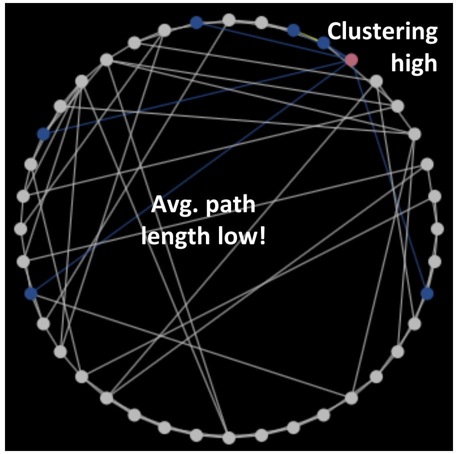
Growing small world networks

- **What small world! From Milgram, to the Hyatt lobby, to power grids, to worm neural networks!**
  - Networks with:
    - high level of clustering: unlikely to quickly reach others, since my friends mainly know my friends

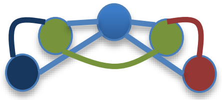


    - small average path length: likely to quickly reach all others on average!





  - Erdos-Renyi misses clustering? => uniform clustering with p
  - Netlogo simulation:

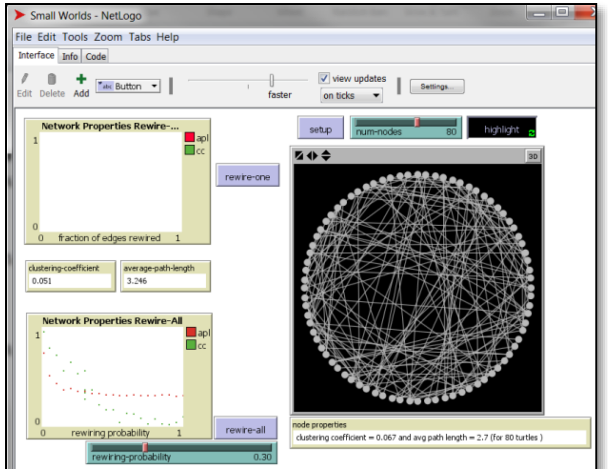


    - each node links to its ***two*** of both neighbors (high clustering & long avg. path)
    - REWIRE changes one end of a connected pair of nodes at random

⇒ few cross-links shrink the avg. path length quickly

⇒ If you don't rewire too many, you keep high clustering

⇒ What's the threshold of the re-wiring probability?



Source: Watts, D.J.; Strogatz, S.H. (1998). "Collective dynamics of 'small-world' networks.". Nature 393 (6684): 409–10.
Wilensky, U. (2005). NetLogo Small Worlds model. Center for CLCBM, Northwestern University.

Exponential Random Graph Models ERGMs (Markov, p*)

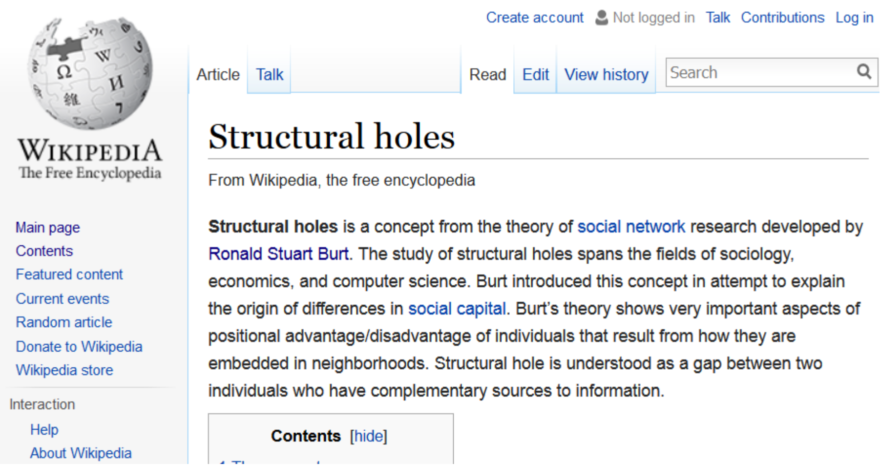
- **Allows to introduce various local features** **(# of links, # of traids, # of isolates, etc)**
- **Also allows for statistical comparisons** **(since we know exponential distributions quite well)**
  - e.g. likelihood of link depends on:
    - node attribute (degree number, etc.)
    - having friends in common (triads)
  - Theorem by Hammersly and Clifford (1971): any network model can be expressed in the exponential family with counts of graph statistics = very general form

$\Rightarrow$ Probability = $e^{(\text{node attribute + triads})} / \sum e^{(\text{all node attribute + all triads})}$

17 nodes can have 272 links = 17*16. If non-weighted: 2^{272} possible networks = 7.6×10^{81},
That's about the estimated number of atoms in the observable universe!
=> Markov-Chain-Monte-Carlo methods are used (sample possible networks) => "Markov models"

- **What's the probability that the actual network (that was observed) would be observed?**
  - Test if local characteristics (like triads) are more or less likely than random
  - PROBLEM: still questioned, since even with MCMC the resulting sample is very small compared to the number of possible networks…

Source: Hammersley, J. and P. Clifford (1971): "Markov fields on finite graphs and lattices,"


Create account Not logged in Talk Contributions Log in

Article Talk Read Edit View history Search

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia

Structural holes

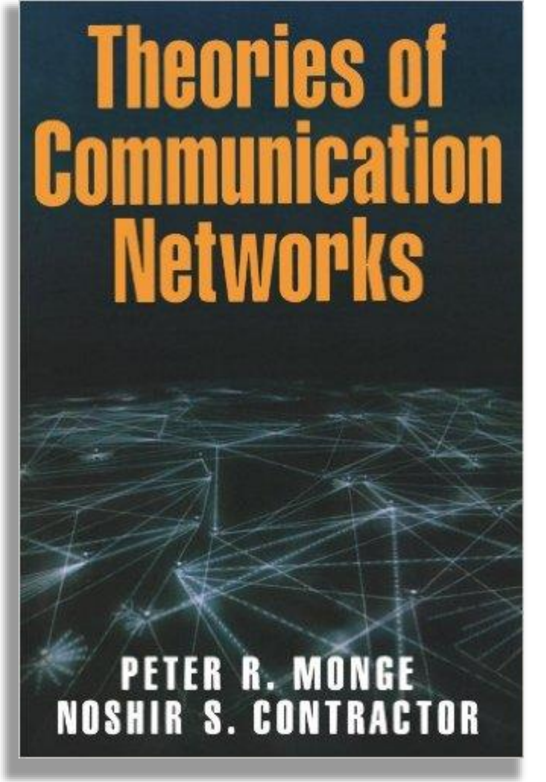
From Wikipedia, the free encyclopedia

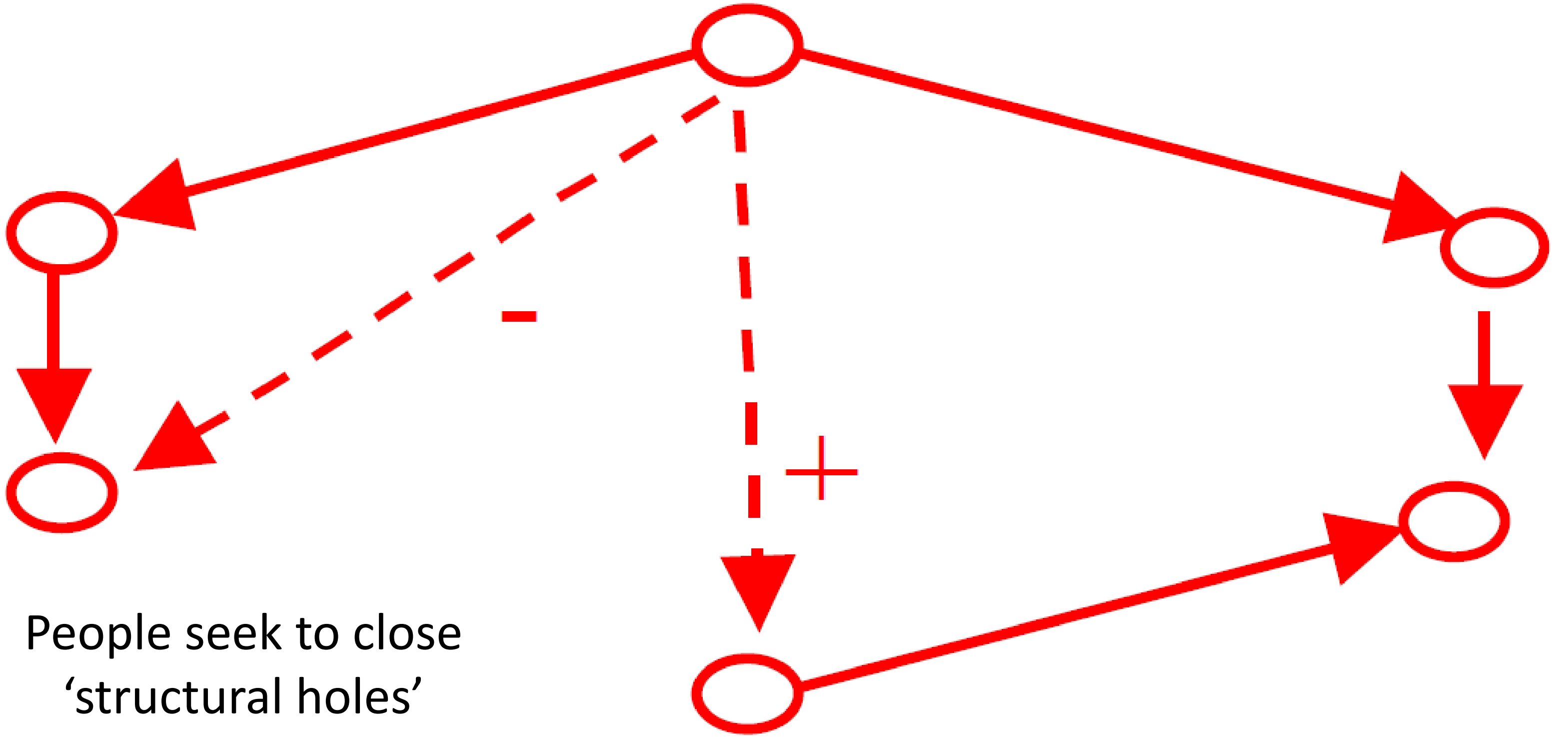
Structural holes is a concept from the theory of social network research developed by Ronald Stuart Burt. The study of structural holes spans the fields of sociology, economics, and computer science. Burt introduced this concept in attempt to explain the origin of differences in social capital. Burt's theory shows very important aspects of positional advantage/disadvantage of individuals that result from how they are embedded in neighborhoods. Structural hole is understood as a gap between two individuals who have complementary sources to information.

Contents [hide]

ERGMs individual level

Network Position metrics





People seek to close
'structural holes'
(= build non-redundant contacts)


Create

Article Talk Read Edit View

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia

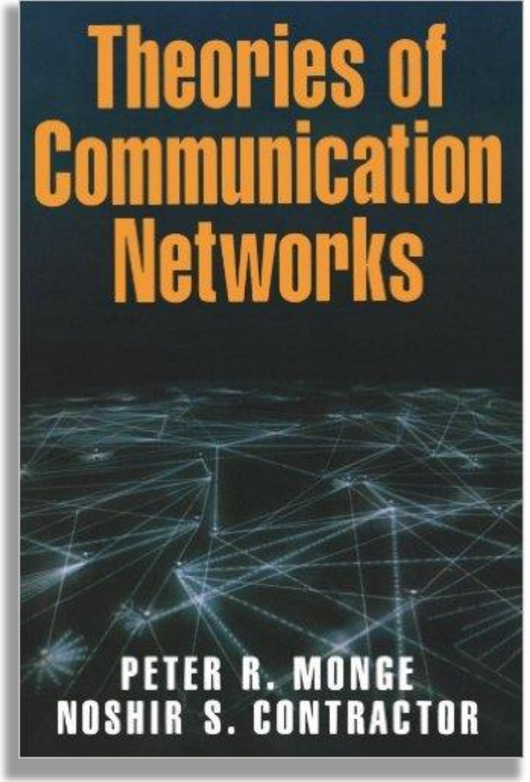
Social exchange theory

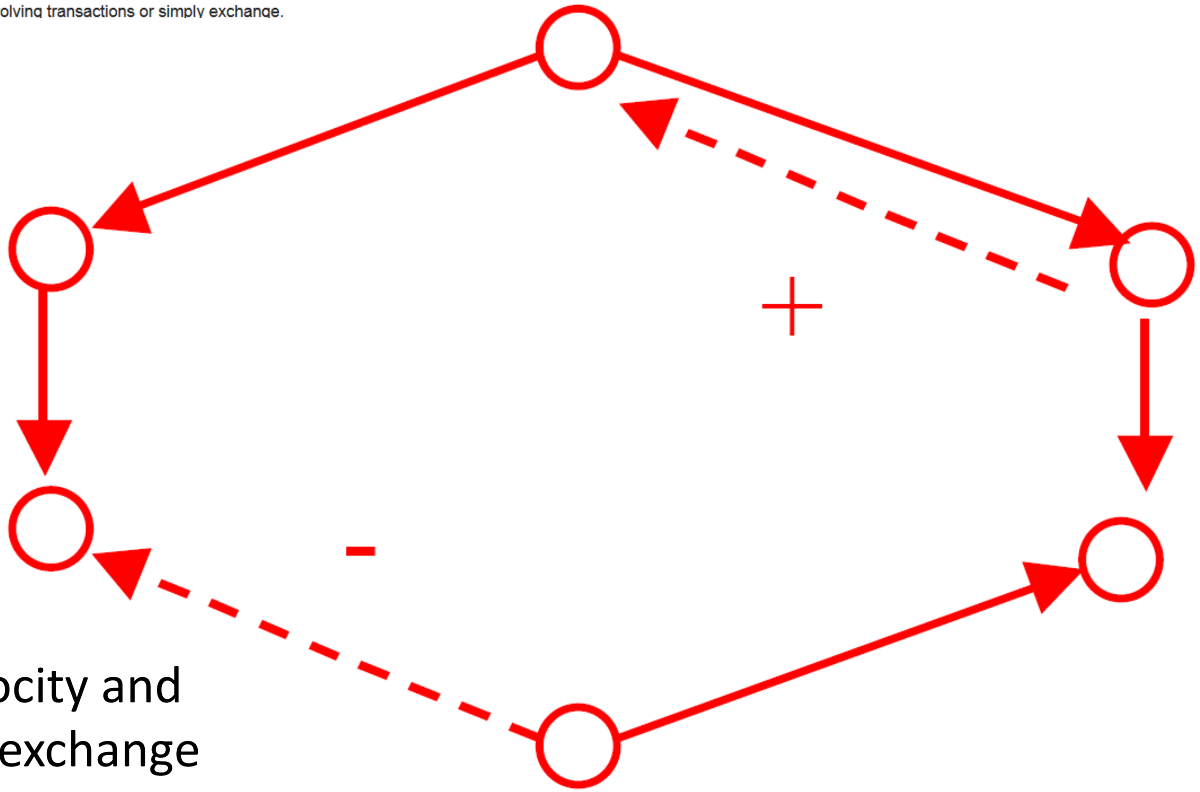
From Wikipedia, the free encyclopedia

Social exchange theory is a social psychological and sociological perspective that explains social change and stability as a process of negotiated exchanges between parties. Social exchange theory posits that human relationships are formed by the use of a subjective cost-benefit analysis and the comparison of alternatives. The theory has roots in economics, psychology and sociology. Social exchange theory features many of the main assumptions found in rational choice theory and structuralism. It is also used quite frequently in the business world to imply a two-sided, mutually contingent and rewarding process involving transactions or simply exchange.

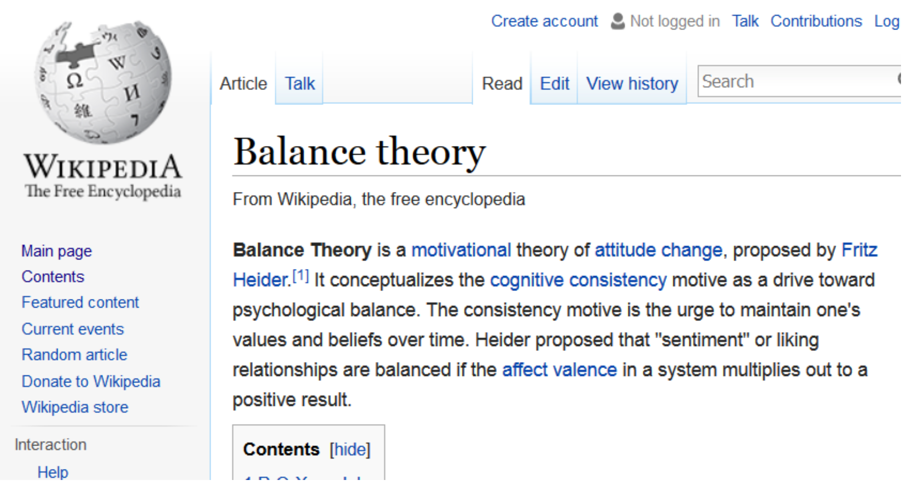
ERGMs dyadic level

Link metrics





Reciprocity and mutual exchange higher than expected by chance


Create account Not logged in Talk Contributions Log

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help

Article Talk Read Edit View history Search

Balance theory

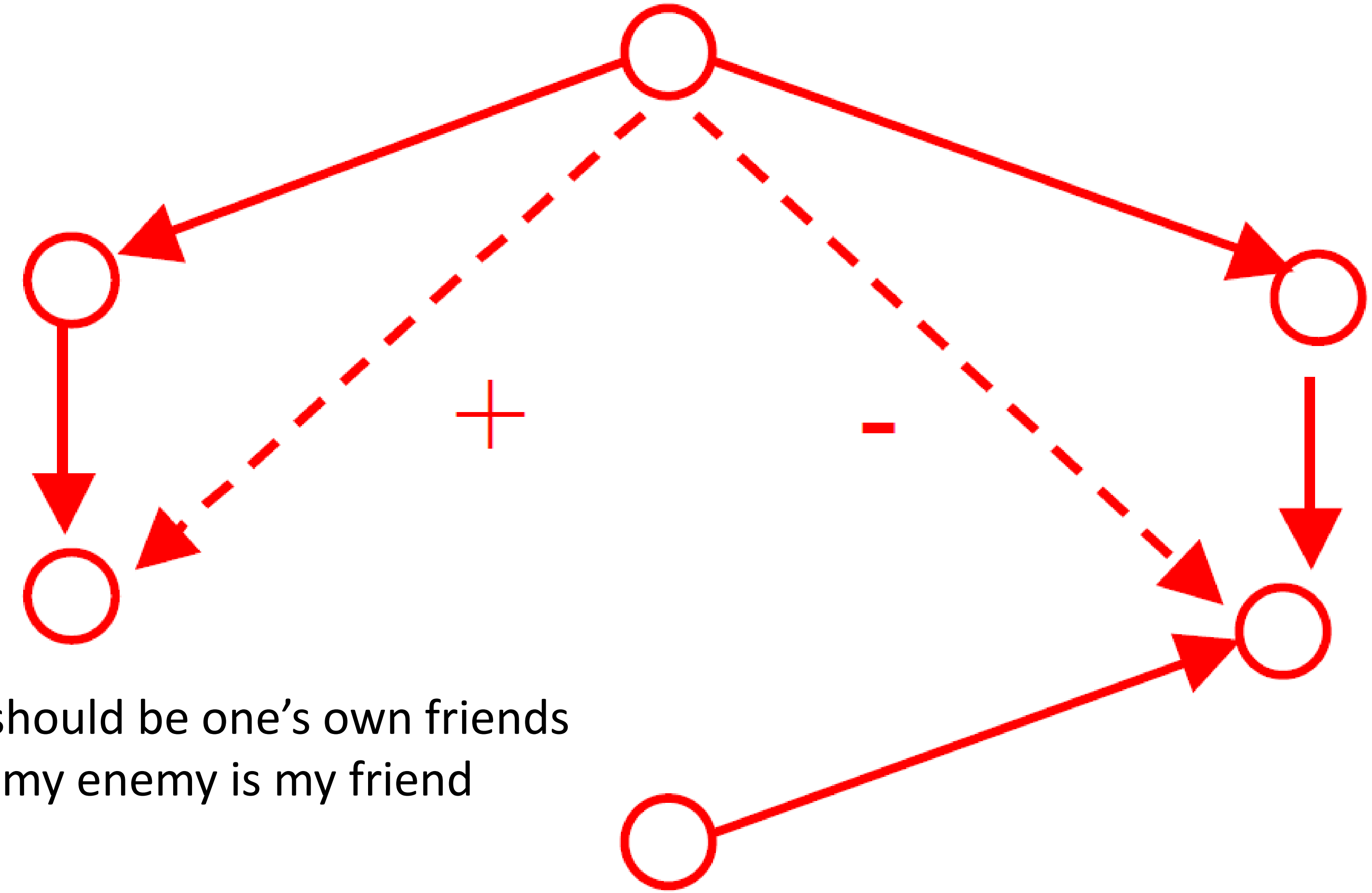
From Wikipedia, the free encyclopedia

Balance Theory is a motivational theory of attitude change, proposed by Fritz Heider.[1] It conceptualizes the cognitive consistency motive as a drive toward psychological balance. The consistency motive is the urge to maintain one's values and beliefs over time. Heider proposed that "sentiment" or liking relationships are balanced if the affect valence in a system multiplies out to a positive result.

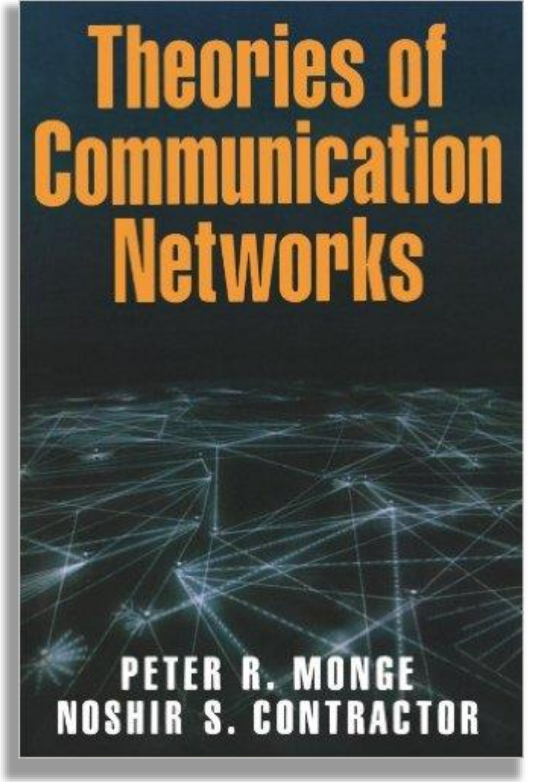
Contents [hide]

ERGMs triad level

Local metrics



Friends of friends should be one's own friends
The enemy of my enemy is my friend



Create account Not logged in Talk Contributions Log in

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal

Article Talk Read Edit View history Search

Preferential attachment

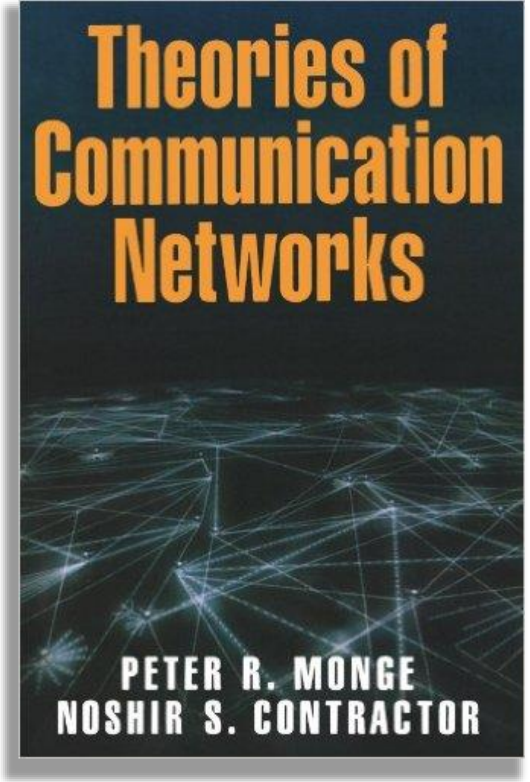
From Wikipedia, the free encyclopedia

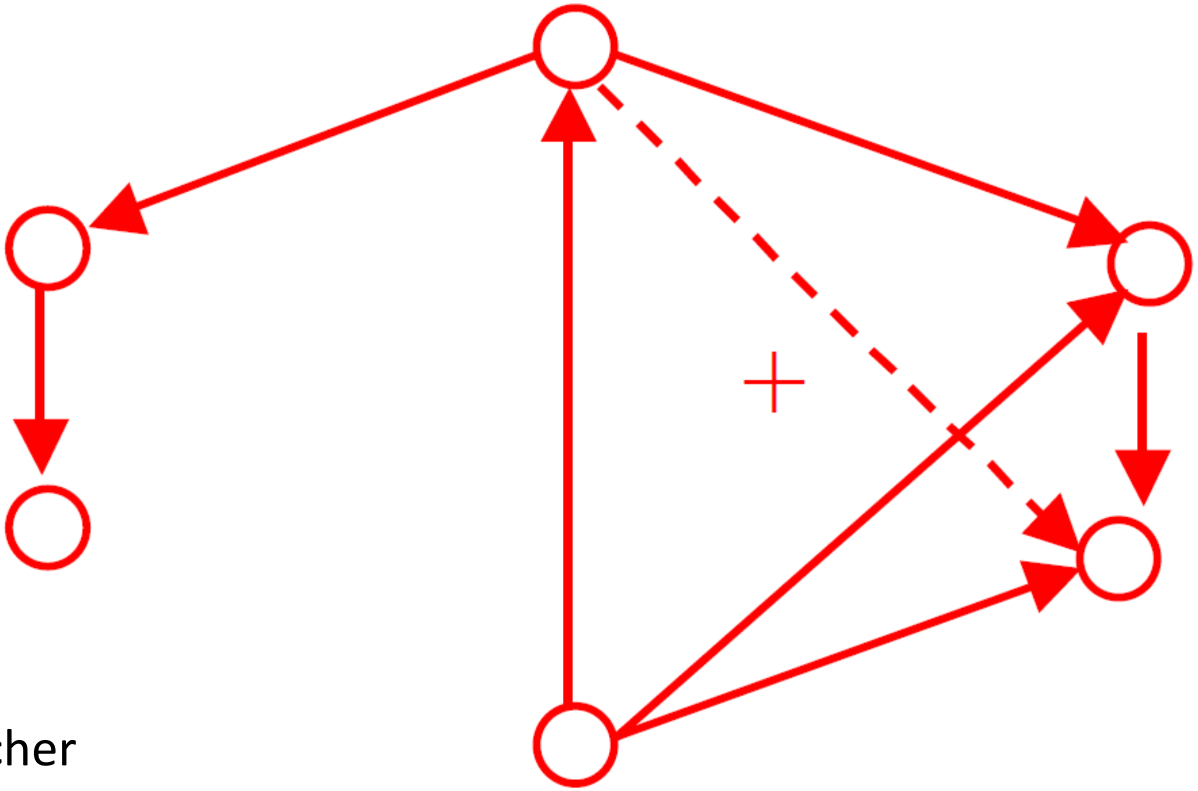
A **preferential attachment process** is any of a class of processes in which some quantity, typically some form of wealth or credit, is distributed among a number of individuals or objects according to how much they already have, so that those who are already wealthy receive more than those who are not. "Preferential attachment" is only the most recent of many names that have been given to such processes. They are also referred to under the names "Yule process", "cumulative advantage", "the rich get richer", and, less correctly, the "Matthew effect". They are also related to Gibrat's law. The principal reason for scientific interest in preferential attachment is that it can, under suitable circumstances, generate power law distributions.

Contents [hide]
1 Definition

ERGMs global level

Entire network metrics





The rich get richer

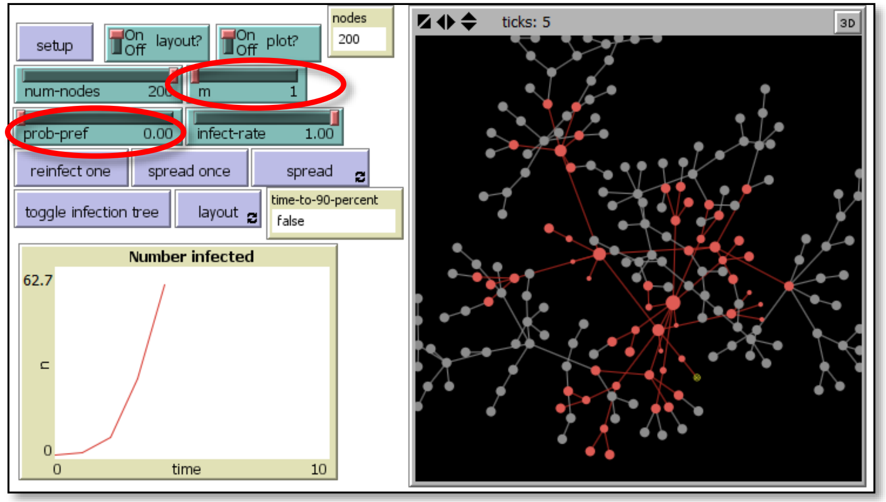
Using networks to make predictions

- **What will happen inside the network? Stable network with given structure**
  - How many clusters will form if the network is being attacked?
  - Diffusion: how does a rumor/ innovation/ opinion/ disease spread

- **What kind of network will form? Changing network structure (links & nodes)**
  - Will stable networks form?
  - Will efficient networks form?
  - How do networks move between stability and efficiency?
  - Can intervention help to improve efficiency/stability?

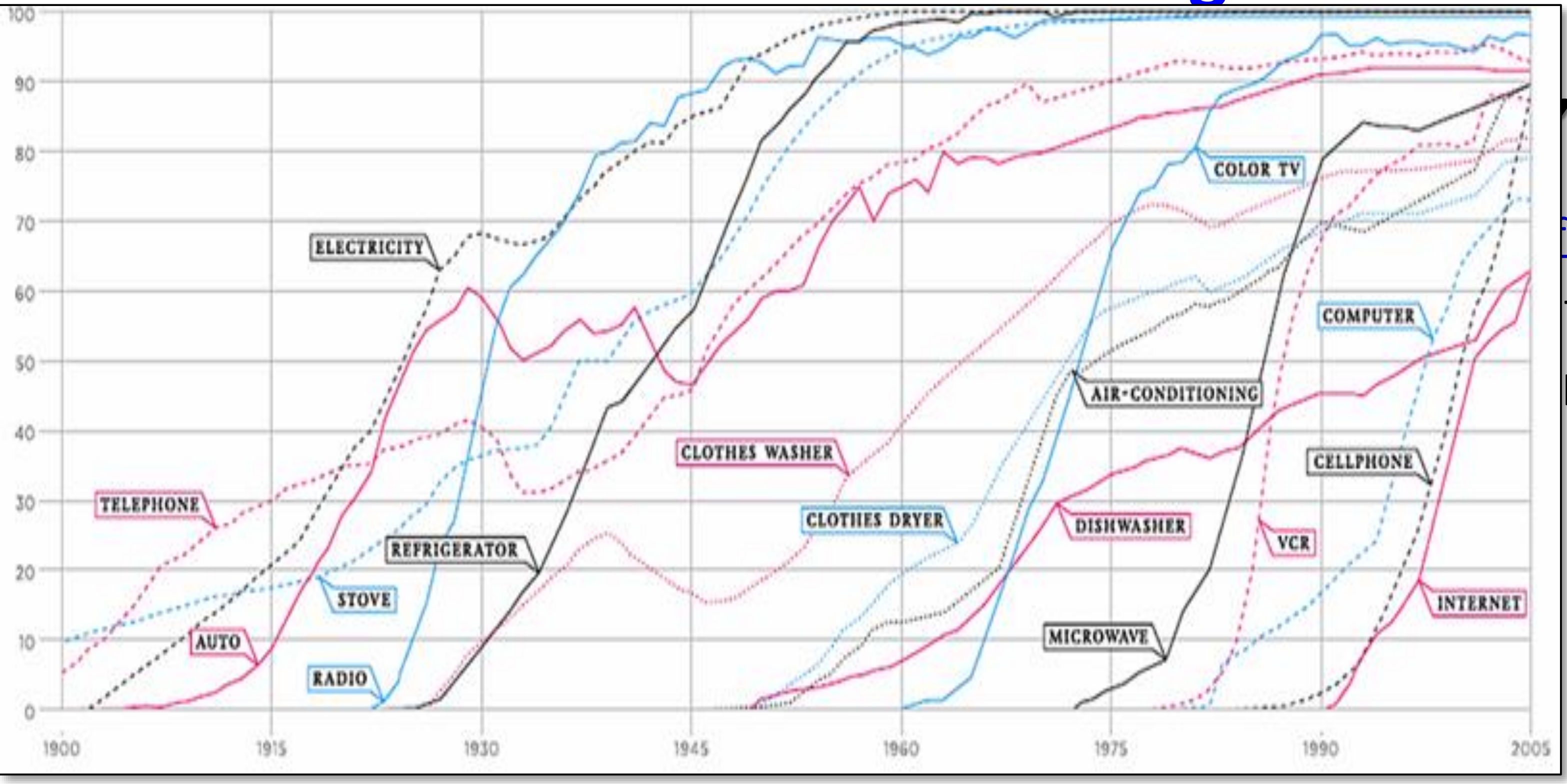
Predicting diffusion patterns

- **NetLogo model of diffusion on Erdos-Renyi , preferential attachment (and hybrid networks)**
  - http://www.LAdamic.com/netlearn/NetLogo501/BADiffusion.html
  - Erdos-Renyi (pro-pref = 0): how does the speed vary with m (# of degrees per node)?
  - What network leads to faster diffusion: Erdos-Renyi or preferential attachment?



Source: Adamic L. (2012), Social Network Analysis, Coursera

Predicting diffusion patterns



, preferential attachment
(and hybrid networks)

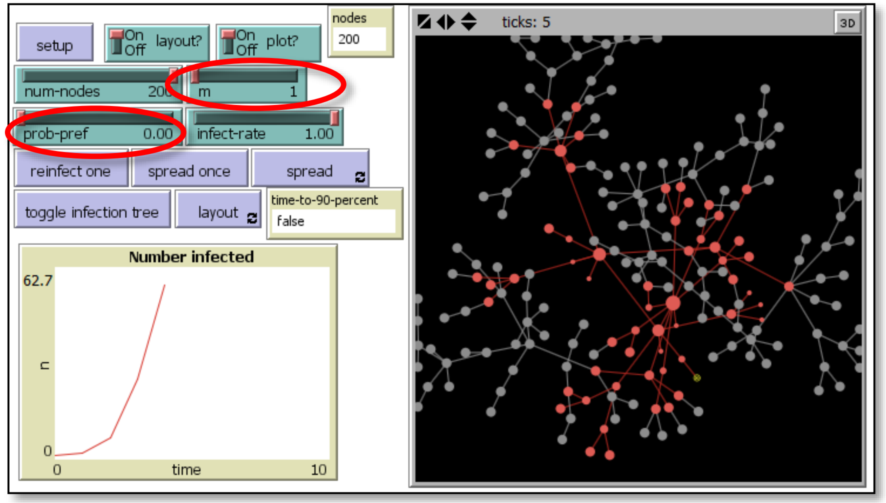
ffusion.html

th m (# of degrees per node)?

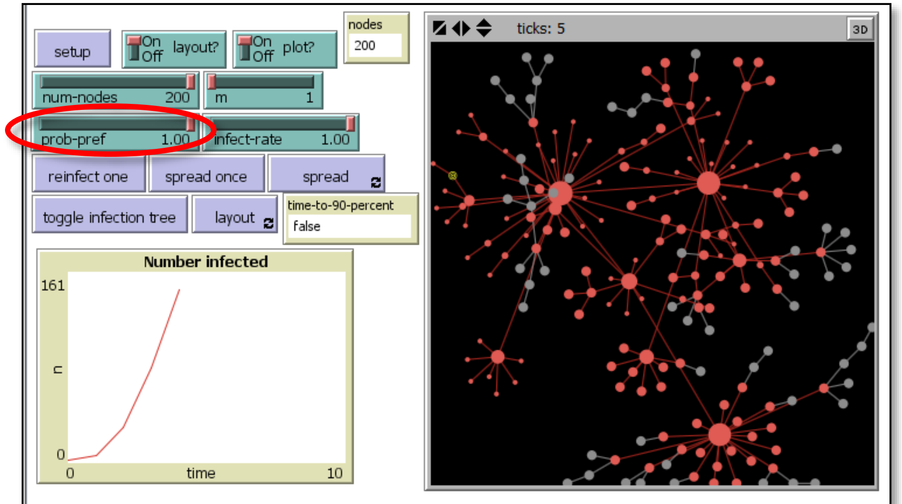
preferential attachment?

With many connections, diffusion is faster

BUT (i.e.: in diffusion of innovation): it can be shown that many connections also lead to adoption of suboptimal solutions! ...more fragmentation => more diversity!



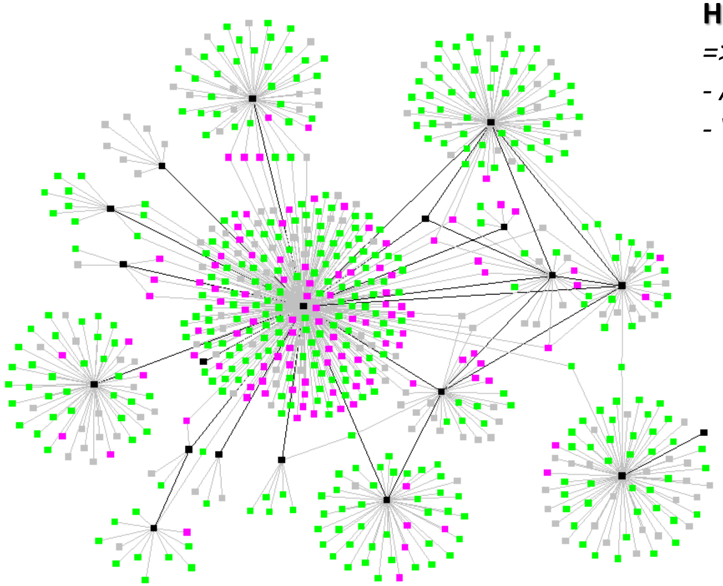
Source: Adamic L. (2012), Social Network Analysis, Coursera



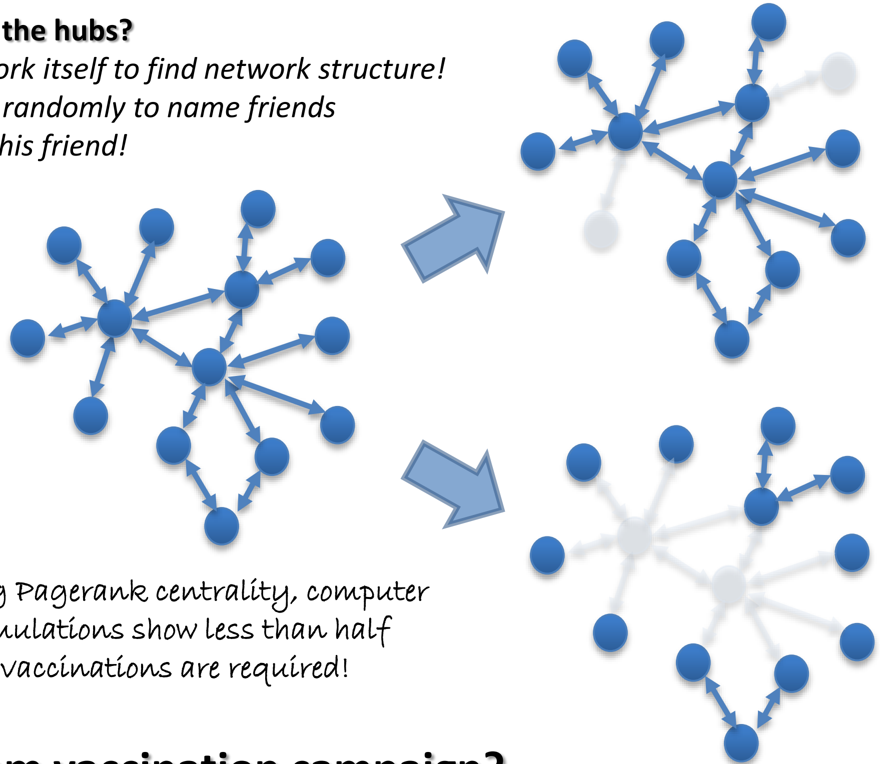
Predicting diffusion patterns

- **Hubs are important[2]**
  - Hubs are more likely to get disease / innovation / opinion
  - Hubs are more likely to pass on disease / innovation / opinion

  $\Rightarrow$ node with degree 1,000 => $1{,}000^2$ = 1,000,000 times more likely to be contagious



How to find the hubs?
=> use network itself to find network structure!
- Ask people randomly to name friends
- Vaccinate this friend!



Using Pagerank centrality, computer simulations show less than half vaccinations are required!

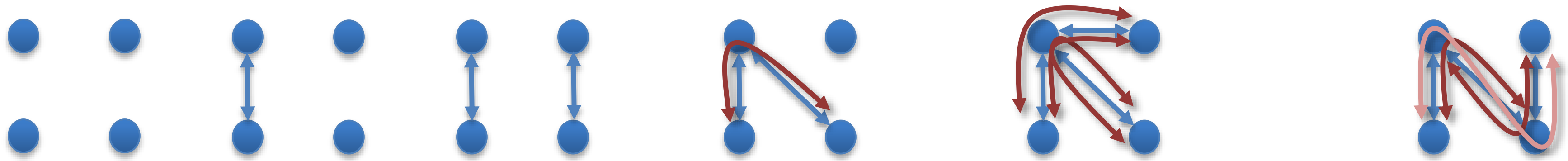
- **How to eliminate disease? Random vaccination campaign?**
- **How to accelerate diffusion of innovation?**

Source: Krebs, 2003; 2005; Mark Newman (2011); Using Networks To Make Predictions , Santa Fe Institute Ulam Lectures; http://www.youtube.com/watch?v=rwA-y-XwjuU ; Mac Hyman (2012). "Part 6 Simulation Models: TRANSISMS and EpiSims". Disaster Resillence leadership Academy" ; http://www.youtube.com/watch?v=pGftX_56X8g

Growing efficient networks until “equilibrium”

- **Logic from (economic) cost & gains** **(leading to Nash equilibrium, Pareto optimality, etc.)**
  - Grow / form / rearrange network until it’s in equilibrium = nothing changes
  - ⇒ No node or link can be added or deleted (pairwise stable equilibrium)
- **Efficient Networks in the Symmetric Connections Model**
  - Given certain **cost** and **benefits of links**: what’s the most efficient network structure?
    - **low cost** (benefit – benefit^2 > cost) => **complete** network is uniquely efficient
    - **medium cost** => **star** networks with all agents are uniquely efficient   Why stars?
    - **high cost** (benefit + (n-2)*benefit^2 < cost) => **empty** network is uniquely efficient

0 benefit – 0 cost | 2 benefit – 2 cost | 4 benefit – 4 cost | 4 benefit + 2 benefit^2 – 4 cost | 6 benefit + 6 benefit^2 – 6 cost | 6 benefit + 4 benefit^2 + 2 benefit^3 – 6 cost



2*0.8 – 2*0.2 = 1.6 – 0.4 = **1.2**

4*0.8 – 4*0.2 = 3.2 – 0.8 = **2.4**

4*0.8 + 2*0.8^2 – 4*0.2 = 3.2 + 1.28 – 0.8 = **3.68**

Indirect benefits!

6*0.8 + 6*0.8^2 – 6*0.2 = 4.8 + 3.84 – 1.2 = **7.44**

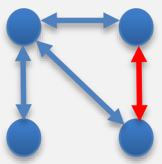
6*0.8 + 4*0.8^2 + 2*0.8^3 – 6*0.2 = 4.8 + 2.56 + 1.024 – 1.2 = **7.184**

normalized 0 ≤ benefit ≤ 1

benefit^3 < benefit^2

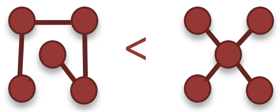
The cost and benefits of adding links

6 benefit + 6 benefit^2 – 6 cost



8 benefit + 4 benefit^2 – 8 cost

Better if: benefit + benefit^2 – cost > 0



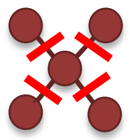
Direct links are more valuable than indirect links

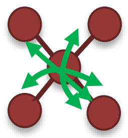
Source: Jackson, M. and A. Wolinsky (1996) “A Strategic Model of Social and Economic Networks,” Journal of Economic Theory 71(1):; Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera.

Growing efficient networks until "equilibrium"

➢ Efficient Networks in the Symmetric Connections Model

- Given certain **cost** and **benefits of links**: what's the most efficient network structure?
  - **low cost** (benefit – benefit^2 > cost) => **complete** network is uniquely efficient
  - **medium cost** =>
    - **Medium-low cost:** (benefit – benefit^2 < cost < benefit) => star (among others)
    - **Medium-high cost:** (benefit < cost < benefit + (n-2)*benefit^2) => since benefit < cost:

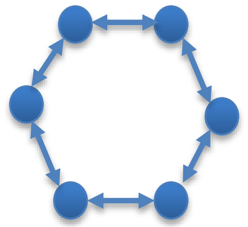




      - you only want to form links if you get indirect benefits!
      - Star center has no indirect benefits => wants to break the link…
      - But periphery has indirect benefits => wants to keep the links!

      ⇒ Social efficiency is higher as a star (most nodes benefit from star), but center doesn't want to play along (doesn't benefit from star)

      ⇒ **Social Efficiency** vs. **Social Instability**!
  - **high cost** (benefit + (n-2)*benefit^2 < cost) => **empty** network is uniquely efficient



For example:
for 6 nodes, and
[benefit] < cost < (benefit + benefit^2 + benefit^3)*(1- benefit^2)

Social efficiency: sum of values of everybody is high
Social stability: nobody can get better off by changing the network (i.e. adding or cutting links)

Source: Jackson, M. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," Journal of Economic Theory 71(1):; Bloch, F., and M.O. Jackson (2006) Definitions of Equilibrium in Network Formation Games, International Journal of Game Theory 34(3):305-318; Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera.

Dynamic Network Evolution

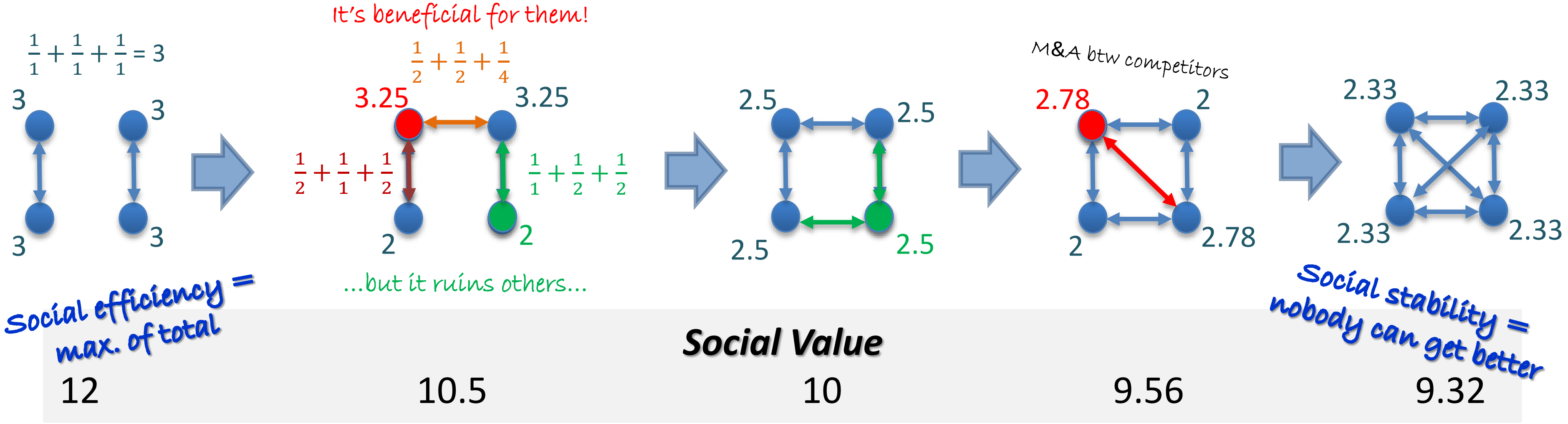
➢ Handling externalities

- Positive externalities: adding links provides additional benefit to same node; e.g. periphery nodes in star
- Negative externalities: adding links provides additional cost to other node; limited resource:
  - Time: if a person/researcher spends time with new friends => less time for the old friendships/teams
  - Money: if an investor links with new projects => less money for old investment connections

➢ How much do you get out of each link with negative externalities?

- Depends on the number of links I have, and the number of links friends have, whereas the value goes down as number of links go up ("co-author model" JW(96))
- Value = SUM of $\left[\frac{1}{\# \, my \, degrees}\right] + \left[\frac{1}{\# \, of \, friend's \, degrees}\right] + \left[synergy = \frac{1}{(\# \, my)*(\# \, friend) degrees}\right]$

Total is more than sum of parts

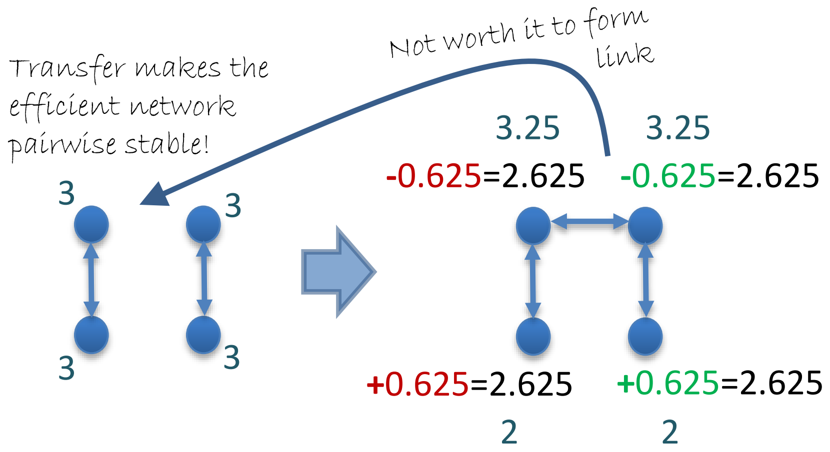


Source: Jackson, M. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," Journal of Economic Theory 71(1):; Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera.

Strategic Network Intervention

➢ Outside intervention / "transfers"

- Public Sector taxing or subsidizing
- Private Sector bargaining/favor among individuals
- e.g. use public or private funds to maintain a centralized Star (hub and spoke) network



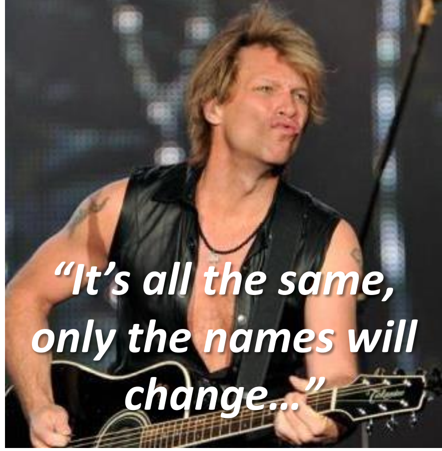
- But doesn't always work:
  - sometimes have to give to some without taking away from others

    ⇒ requires outside resources (take away from unaffected…)
  - sometimes have to take away from some without taking away from others

    ⇒ accumulate extra resources (give to unaffected…)

Source: Jackson, M. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," Journal of Economic Theory 71(1):; Bloch F, Jackson MO. 2007. The formation of networks with transfers among players. Journal of Economic Theory 133: 83 - 110; Jackson M. (2013); Social and Economic Networks: Models and Analysis; Coursera.

Network Science



Description

Find a real-world phenomenon and describe its properties

Online friend network with different global, local and position measures

Look for other properties

Analysis

Analyze these properties and look for patterns in the noise

Some have many more friends than others

Adjust theory

Science

Theory

The ones that play sports have more friends

Hypothesize a generative mechanism of the functionality / *modus operandi* that produces this phenomenon

?

Modeling / Testing

How to test if our ideas are correct or if we are just daydreaming?

Might it be just random chance?

Today's questions

I. How can we find out if our model of the network is correct?

II. How can we model network dynamics over time?

III. How can we use networks to make predictions?