An introduction to network analysis and modeling with applications to social contagion processes

Part II

Juan G. Restrepo
University of Colorado at Boulder

June 28, 2022
ICERM

Talk outline

Part 2

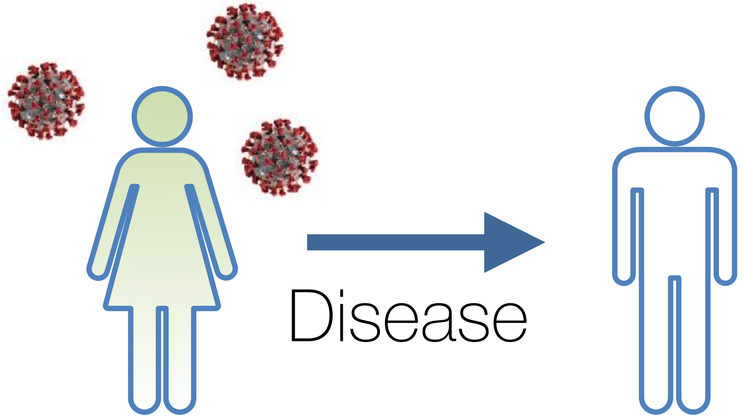
- Modeling processes on networks: SIS example
- Creating, importing and visualizing networks in matlab and python
- Example: creating model of **diversity** and productivity

Talk outline

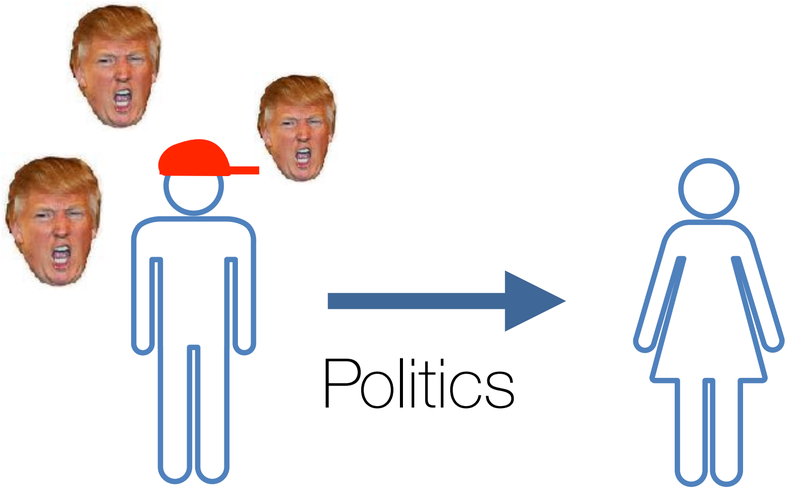
Part 2

- Modeling processes on networks: SIS example
- Creating, importing and visualizing networks in matlab and python
- Example: creating model of **diversity** and productivity

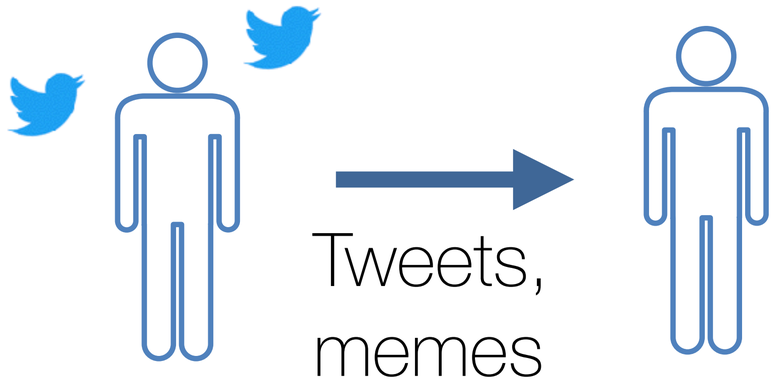
Social contagion: more than disease



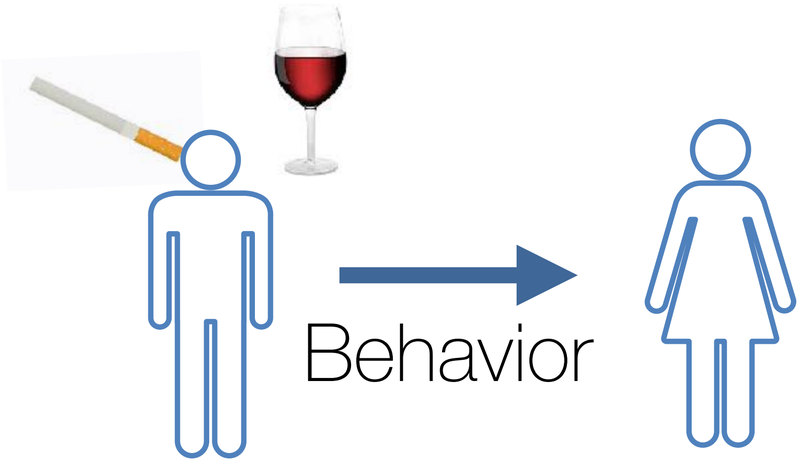
Disease



Politics



Tweets,
memes



Behavior

Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

Define dynamics: How can the states of the nodes change?

Simulate dynamics: On network models or relevant network datasets

Analyze dynamics: Usually involves approximations…

Interpret results: For example, how can a desirable property be maximized?

Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

Define dynamics: How can the states of the nodes change?

Simulate dynamics: On network models or relevant network datasets

Analyze dynamics: Usually involves approximations…

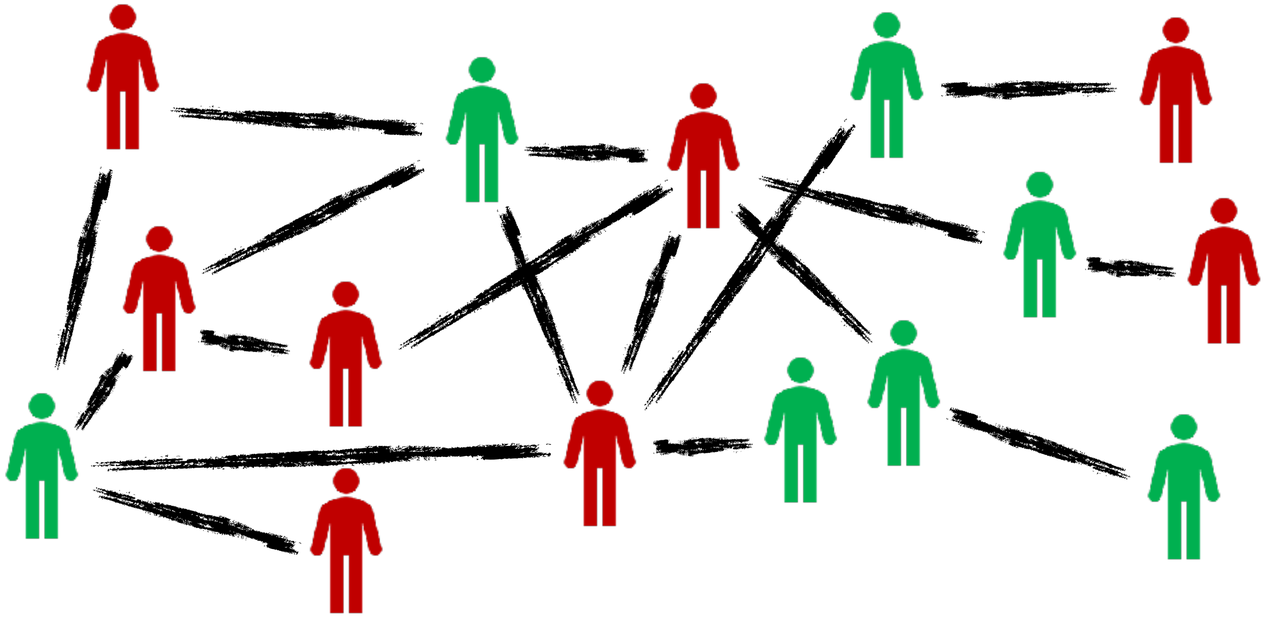
Interpret results: For example, how can a desirable property be maximized?

Simplest model of social contagion: SIS

Susceptible-Infected-Susceptible (**SIS**):

Nodes represent individuals, links represent interactions.

Each individual can be in two states: susceptible (S) or infected (I).



Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

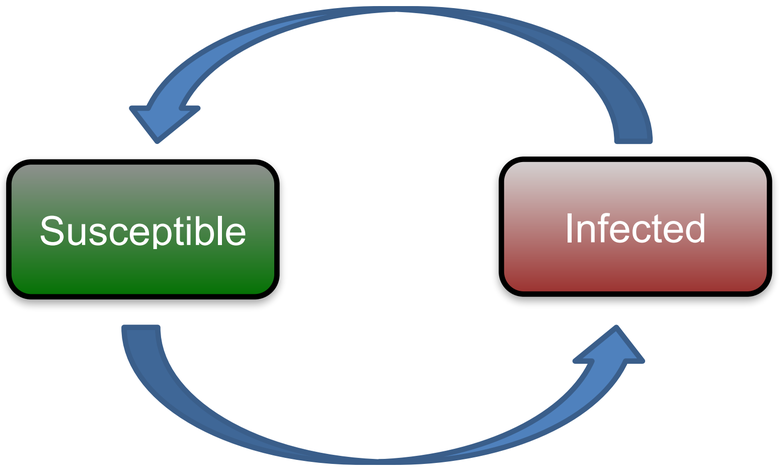
Define dynamics: How can the states of the nodes change?

Simulate dynamics: On network models or relevant network datasets

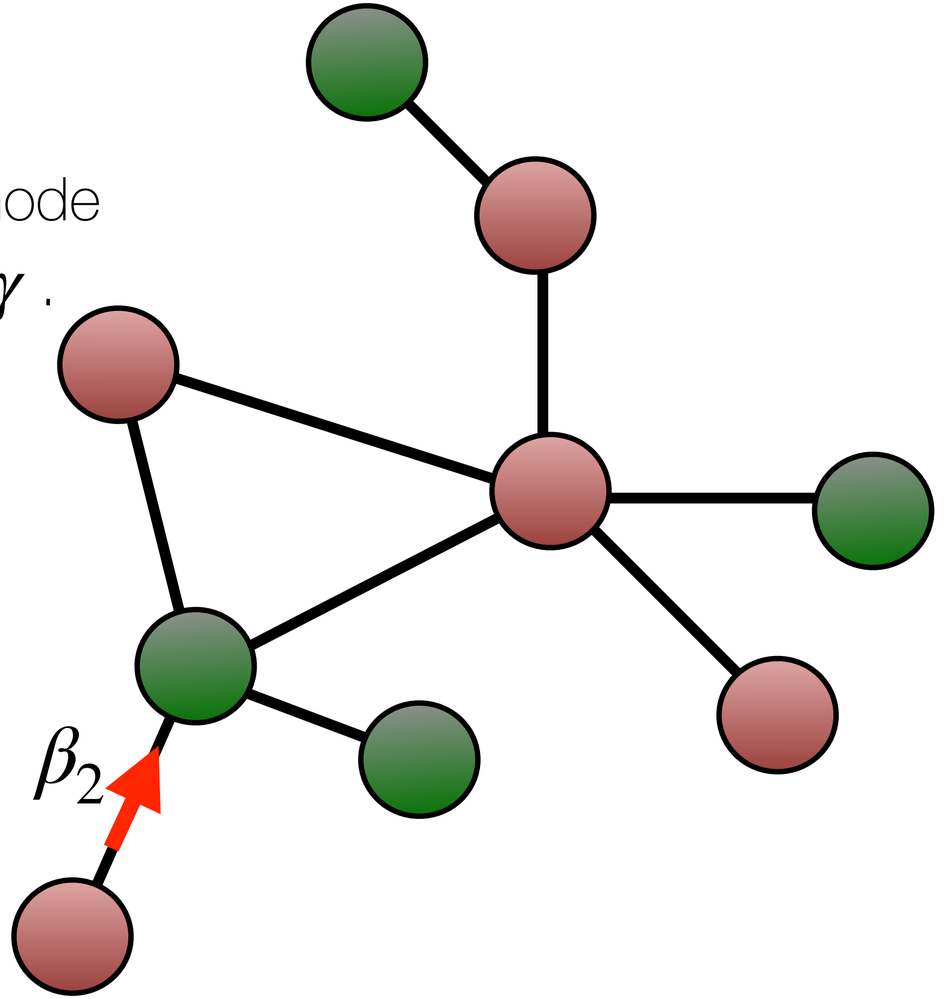
Analyze dynamics: Usually involves approximations…

Interpret results: For example, how can a desirable property be maximized?

Network SIS model



An infected node heals at rate γ .



A healthy node can get infected by an infected neighbor at rate β_2.

Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

Define dynamics: How can the states of the nodes change?

Simulate dynamics: On network models or relevant network datasets

Analyze dynamics: Usually involves approximations…

Interpret results: For example, how can a desirable property be maximized?

Network SIS model

Loop through (small) time intervals Δt:

Loop through all nodes $i = 1\ldots, N$

If node i is infected, heal with probability $\gamma \Delta t$

If node i is susceptible,

Find the number of infected network neighbors, M

Infect with probability $1 - (1 - \beta \Delta t)^M \approx M \beta \Delta t$

End

End

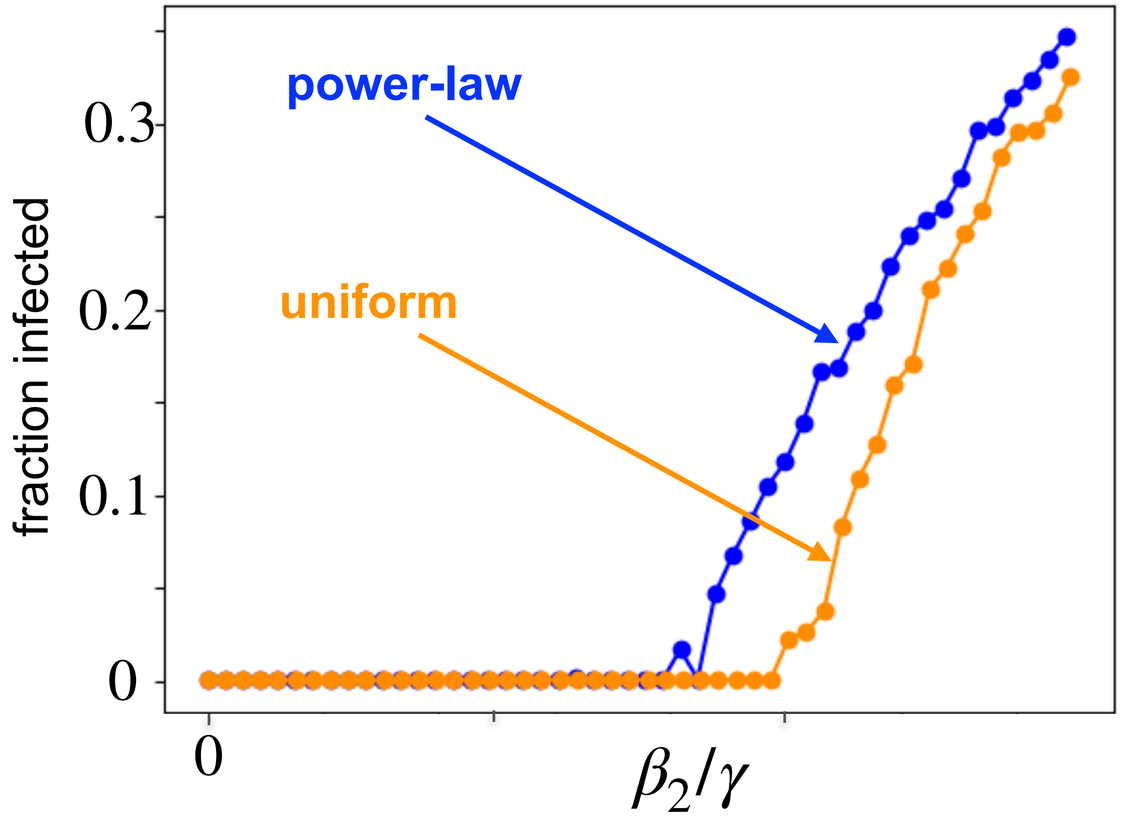
End

Epidemic threshold

Network SIS model simulated on two degree-based models **with the same mean degree**:

Power law degree distribution

Degree distribution uniform in $[50,100]$



Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

Define dynamics: How can the states of the nodes change?

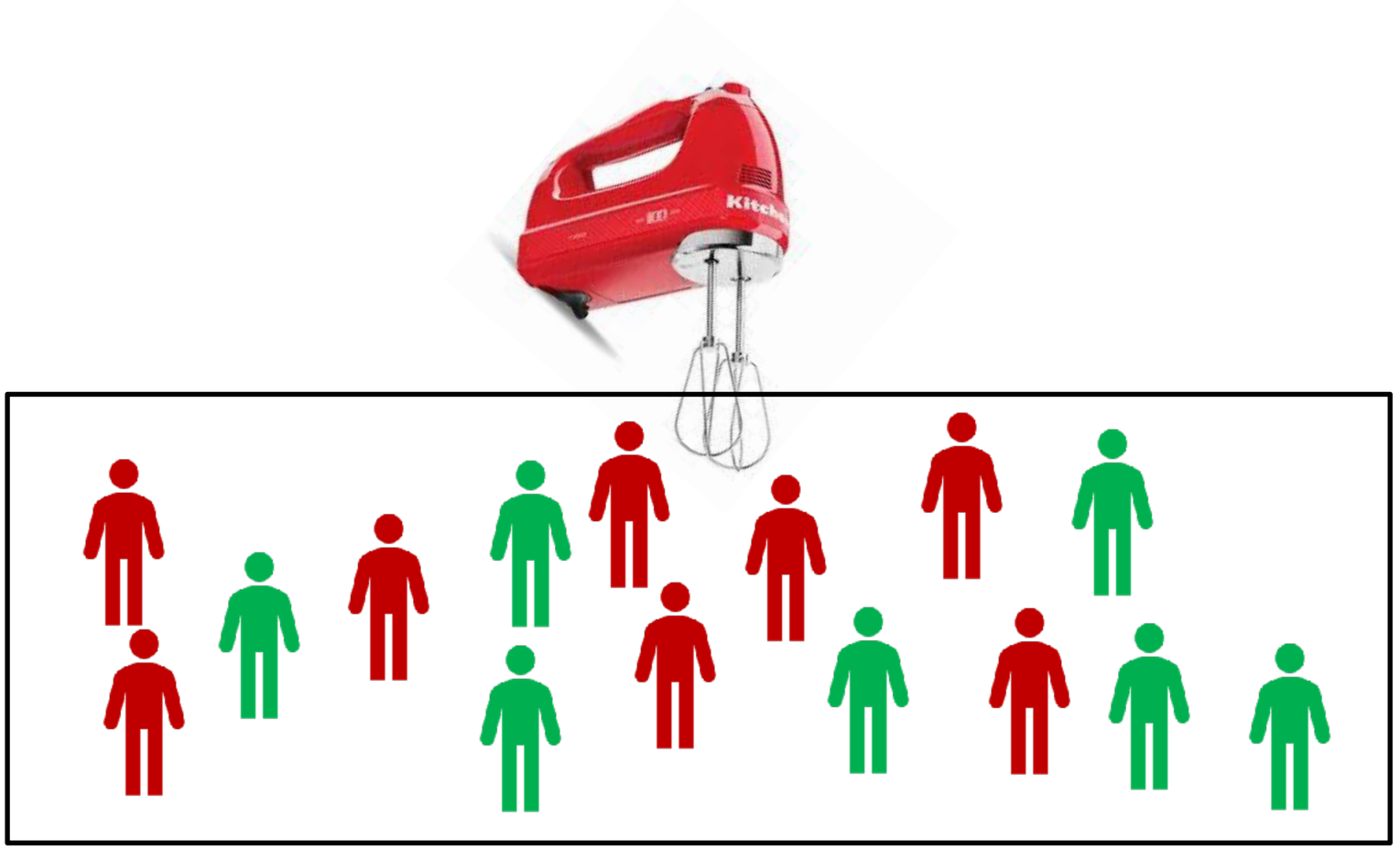
Simulate dynamics: On network models or relevant network datasets

Analyze dynamics: Usually involves approximations…

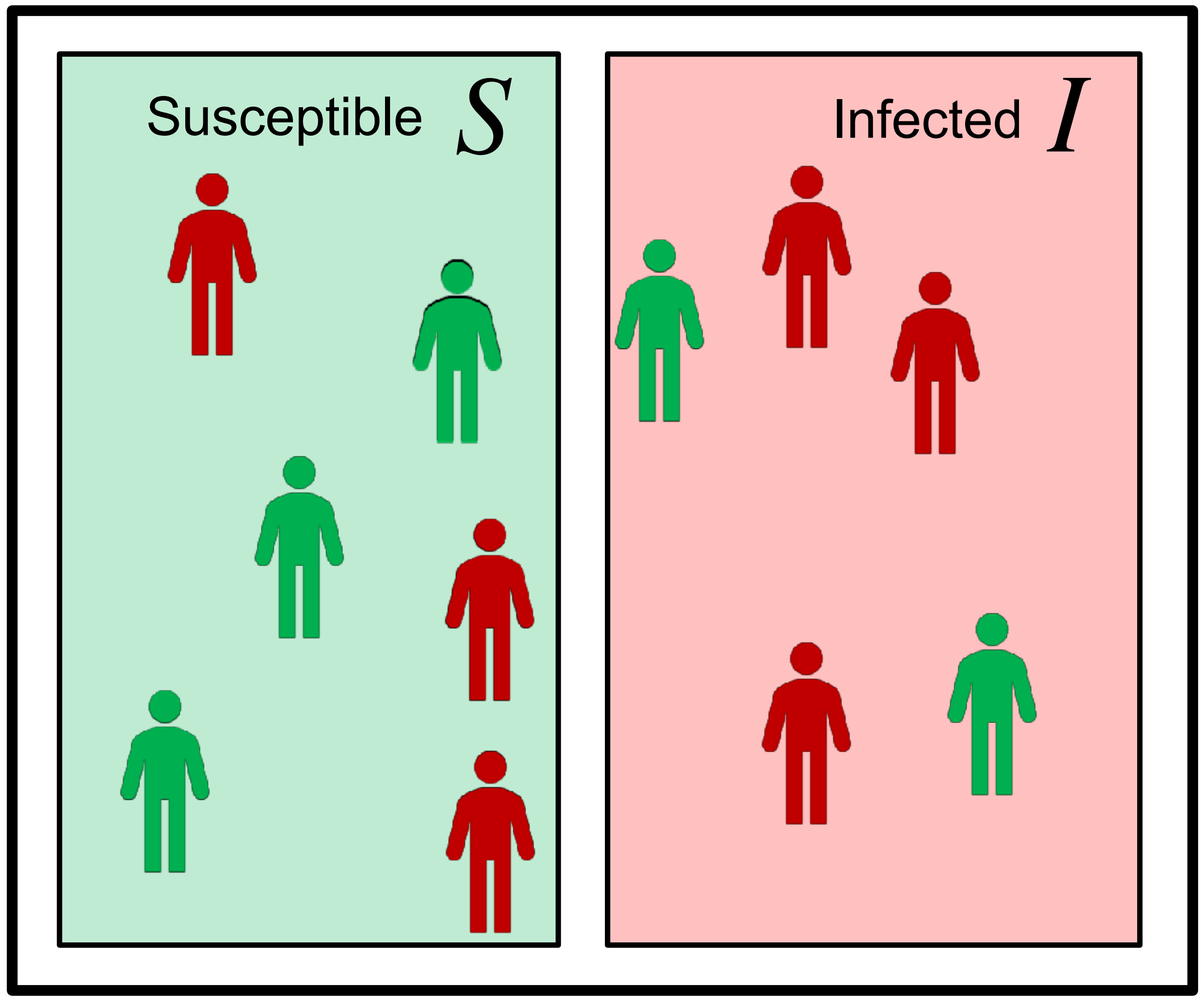
Interpret results: For example, how can a desirable property be maximized?

Fully mixed contagion

This models assumes that a population of individuals is fully mixed and that everyone influences each other equally



Compartmental model for SIS



$$S + I = 1$$

$$\frac{dI}{dt} = -\gamma I + \beta_2 NSI$$

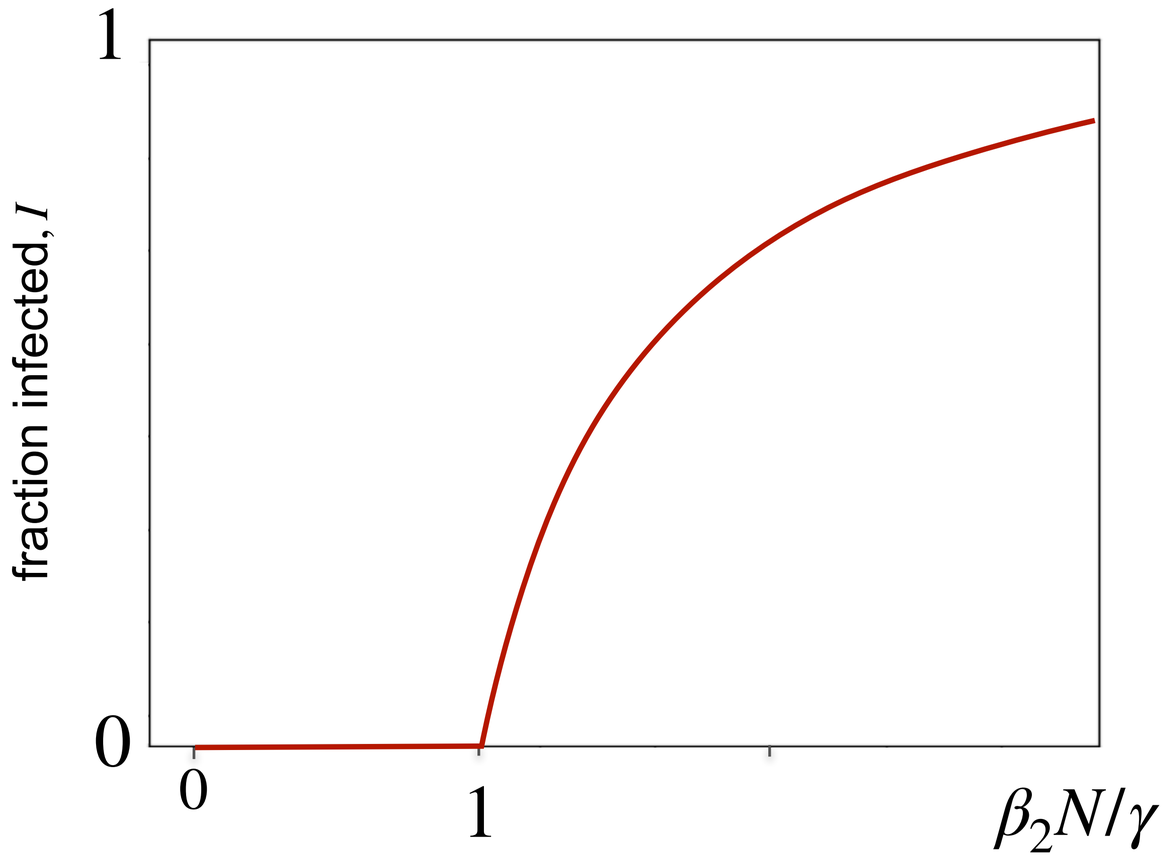
$\gamma =$ rate of healing

$\beta_2 =$ rate of infection

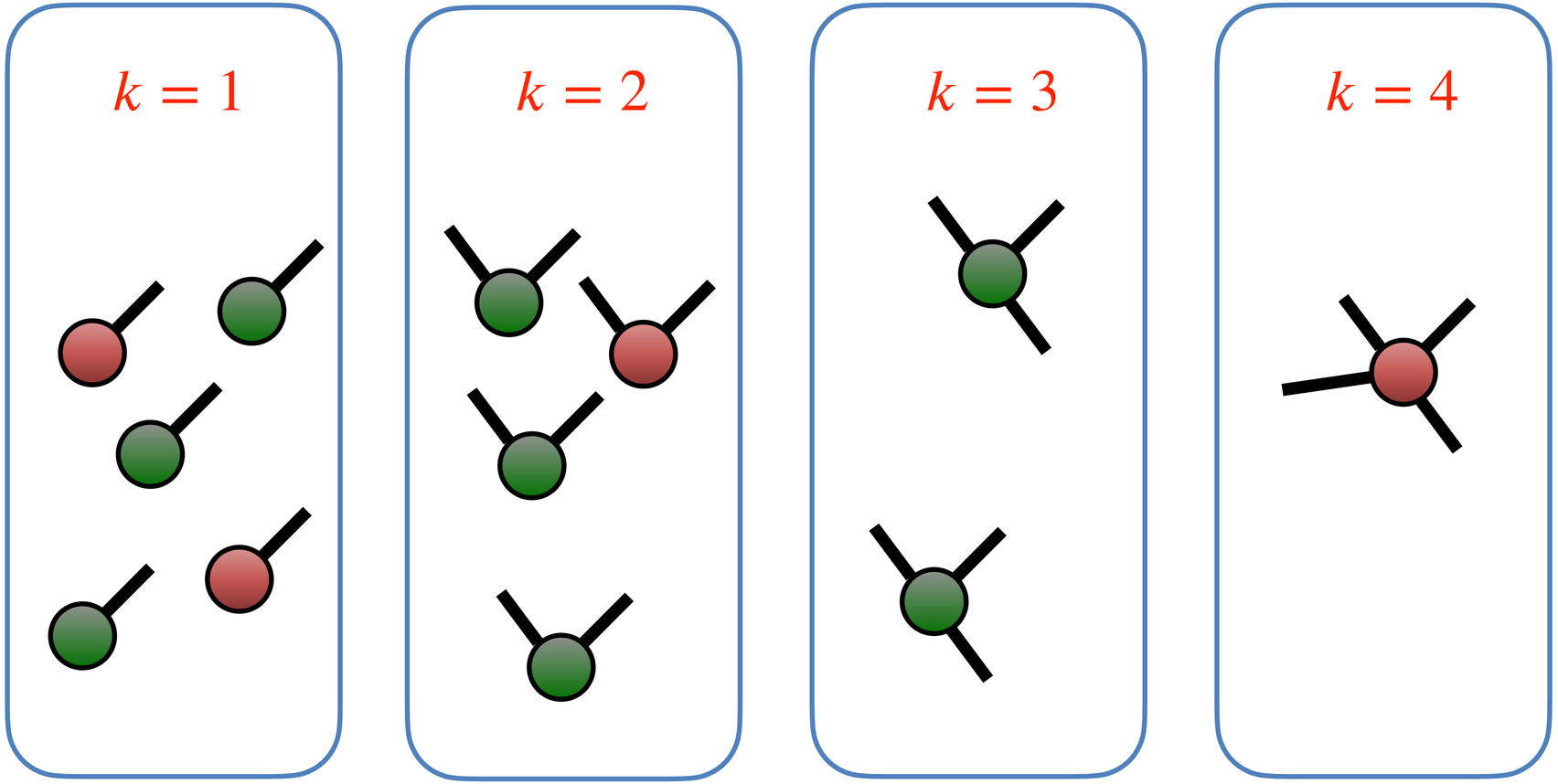
Compartmental model for SIS

$$\frac{dI}{dt} = -\gamma I + \beta_2 N(1-I)I$$

Equilibrium solution of this equation:



Degree-based compartmental models



x_k= fraction of nodes that are infected

Pastor-Satorras, Romualdo, and Alessandro Vespignani. "Epidemic spreading in scale-free networks." Physical review letters 86.14 (2001): 3200.

Network SIS model

Degree-based compartmental model predicts that an epidemic occurs for $\beta_2 > \beta_2^c = \gamma \frac{\langle k \rangle}{\langle k^2 \rangle}$

where $\langle k^n \rangle$ is the nth moment of the degree distribution, $\langle k^n \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^n$

Pastor-Satorras, Romualdo, and Alessandro Vespignani. "Epidemic spreading in scale-free networks." Physical review letters 86.14 (2001): 3200.

A more refined analysis predicts that an epidemic occurs for $\beta_2 > \beta_2^c = \frac{\gamma}{\lambda}$

where λ is the largest eigenvalue of the adjacency matrix A

Wang, Yang, et al. "Epidemic spreading in real networks: An eigenvalue viewpoint." 22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.. IEEE, 2003.

For more sophisticated analyses, see

Kiss, István Z., Joel C. Miller, and Péter L. Simon. "Mathematics of epidemics on networks." Cham: Springer 598 (2017): 31.

Modeling dynamics on networks

Define nodes and links: What do they represent?

Define node variables: What properties do nodes have? What states can they be in?

Define dynamics: How can the states of the nodes change?

Simulate dynamics: On network models or relevant network datasets

Analyze dynamics: Usually involves approximations

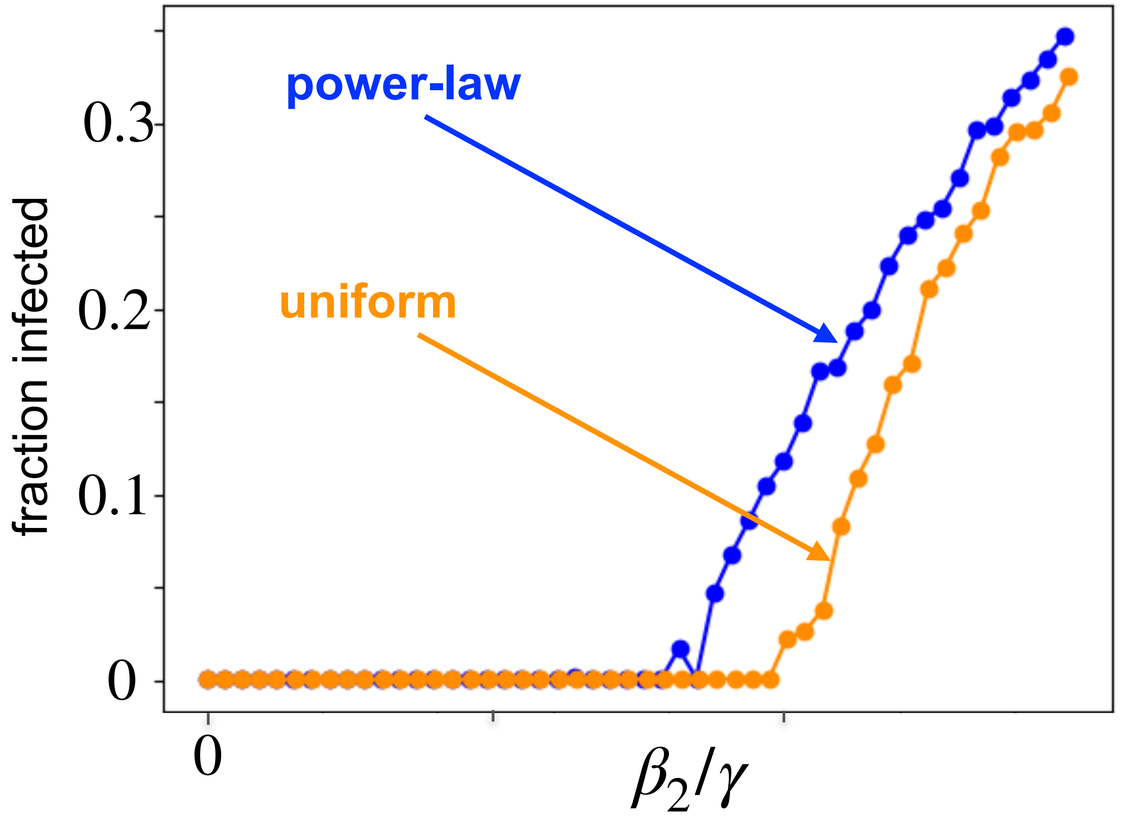
Interpret results: For example, how can a desirable property be maximized?

Epidemic threshold

$$\beta_2 > \beta_2^c = \gamma \frac{\langle k \rangle}{\langle k^2 \rangle}$$

The more heterogeneous the degree distribution is, the lower the epidemic threshold

Heterogeneity in the contact network promotes epidemics



Talk outline

Part 2

- ✓ Modeling processes on networks: SIS example
- Creating, importing and visualizing networks in matlab and python
- Example: creating model of **diversity** and productivity

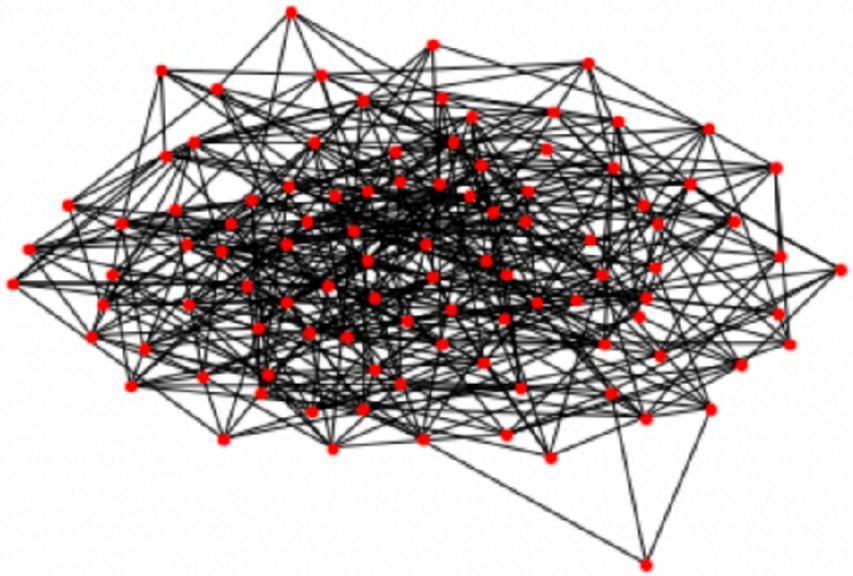
Python & networkx

The **networkx** library in **Python** has many prepackaged network generative models and tools

Erdos-Renyi network

```
n = 100;
G = nx.gnp_random_graph(n, 10/n);
```

```
pos = nx.spring_layout(G)
nx.draw(G,  node_color="r", alpha=1,node_size = 20)
```

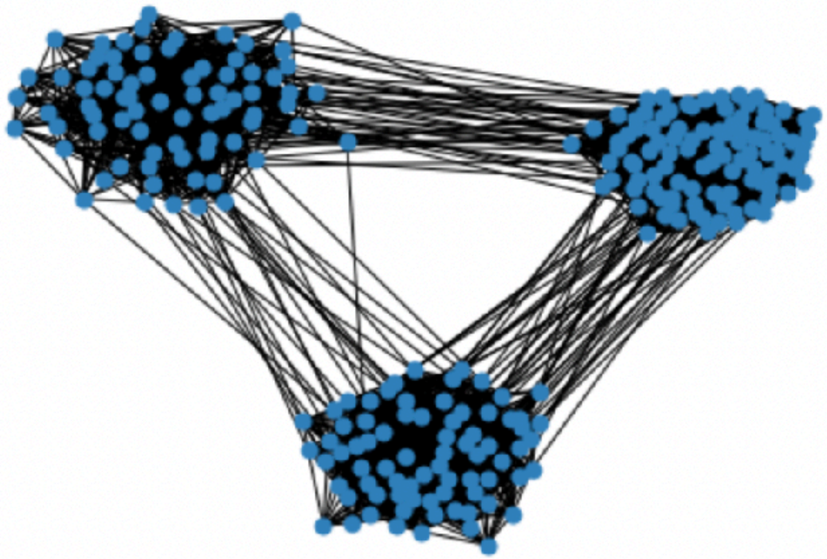


Python & networkx

Stochastic Block Model

```
sizes = [75, 75, 100]
probs = [[0.25, 0.005, 0.005], [0.005, 0.35, 0.005], [0.005, 0.005, 0.40]]
g = nx.stochastic_block_model(sizes, probs, seed=0)
```

```
pos = nx.spring_layout(g)
nx.draw(g, pos, node_color="#3182bd", alpha=1,node_size = 50)
```

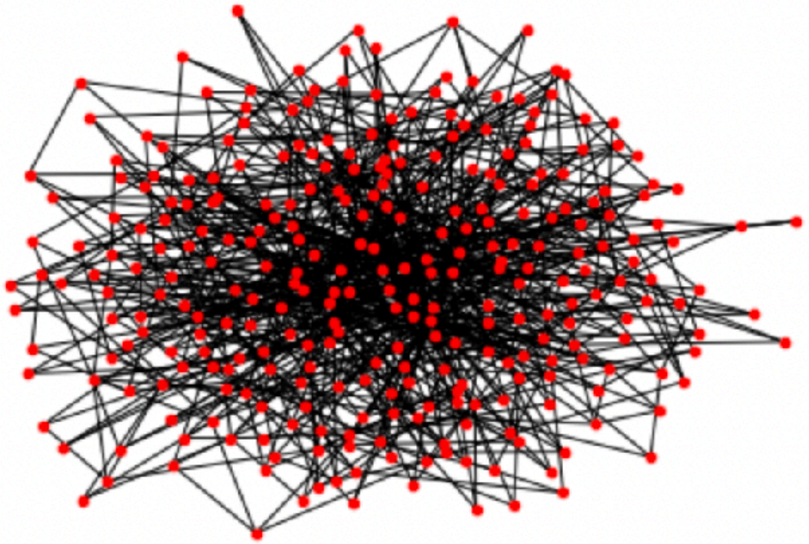


Python & networkx

Preferential attachment (Barabasi-Albert model)

In [6]:
```
n = 300;
G = nx.barabasi_albert_graph(n, 3, seed=None, initial_graph=None)
```

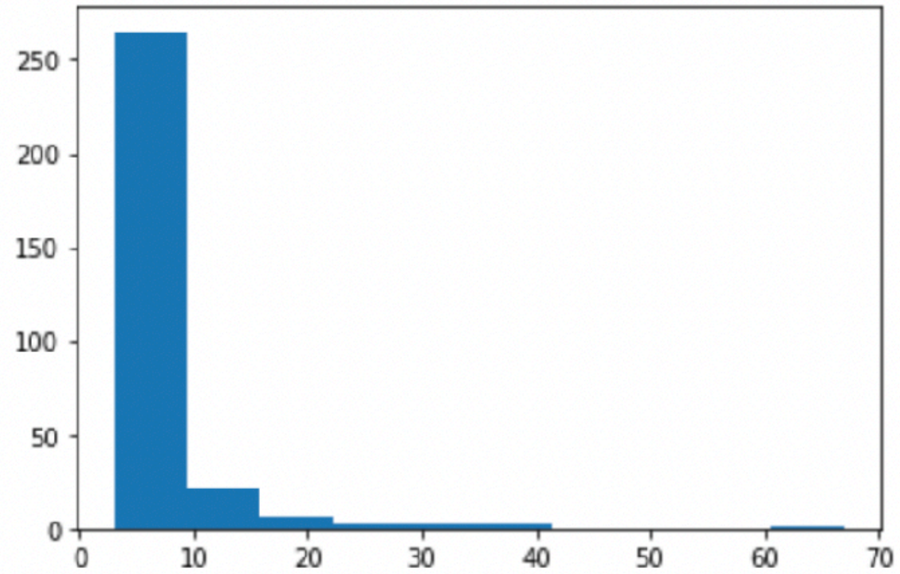
In [7]:
```
pos = nx.spring_layout(G)
nx.draw(G, pos,  node_color="r", alpha=1,node_size = 20)
```



Python & networkx

Degree distribution

```
k = nx.degree_sequence = sorted((d for n, d in G.degree()), reverse=True)
pyplot.hist(k);
```



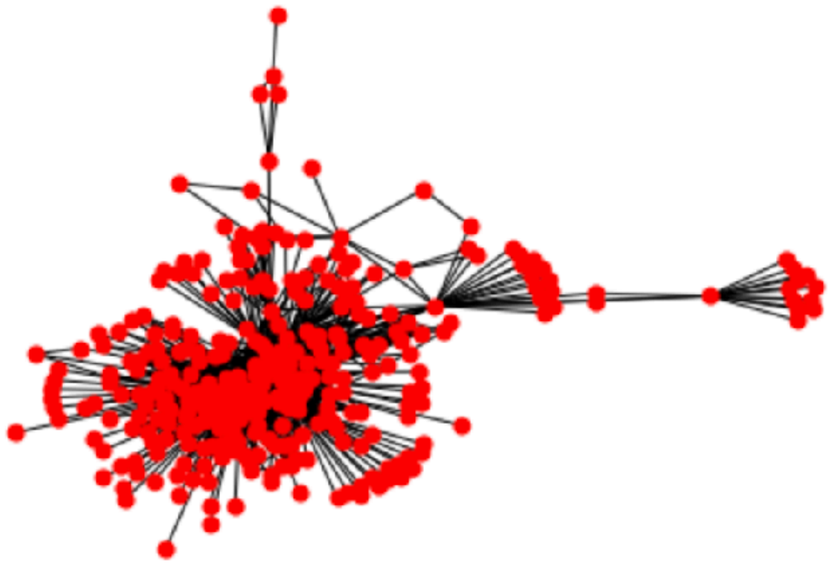
Python & networkx

Import edgelist file, create graph, and plot

In [8]:
```
edgelist = pd.read_csv('/Users/restrepj/Dropbox/TalksDropbox/ICERM/airports_links.csv')
```

In [9]:
```
# Create empty graph
g = nx.Graph()
# Add edges and edge attributes
for i, elrow in edgelist.iterrows():
    g.add_edge(elrow[0], elrow[1], attr_dict=elrow[2:].to_dict())
```

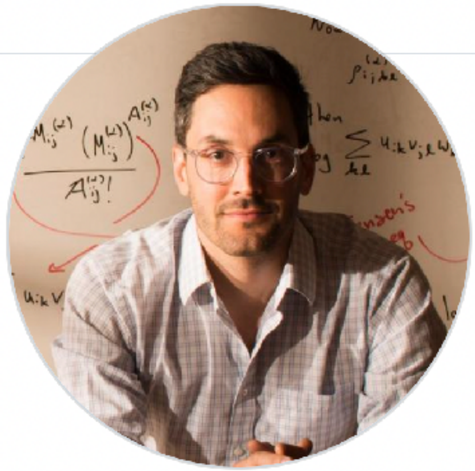
In [121]:
```
pos = nx.spring_layout(G)
nx.draw(g,  node_color="r", alpha=1,node_size = 50)
```



matlab & Python: webweb

Visualization can be misleading, but it can also be useful for dissemination

webweb is an & easy-to-use visualization tool for matlab/python



Daniel Larremore
dblarremore

webweb

Search webweb

webweb on GitHub

webweb
getting started
examples
documentation
webweb in the wild
changelog
about

webweb

webweb is a tool for creating, displaying, and sharing interactive network visualizations on the web, designed for simplicity and ease of use. With just a few lines of `python`, `networkx`, or `matlab`, webweb will build and launch a visualization in your browser.

Here's an example of webweb's style and functionality. The network itself comes from Hunter Wapman et al.'s analysis of character co-occurrences in the novel *Infinite Jest*.

matlab

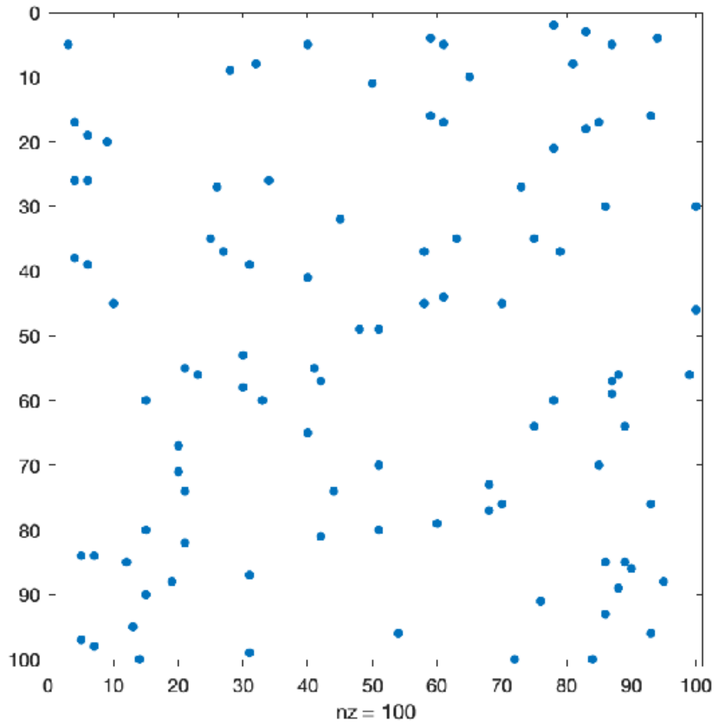
Erdös-Renyi network

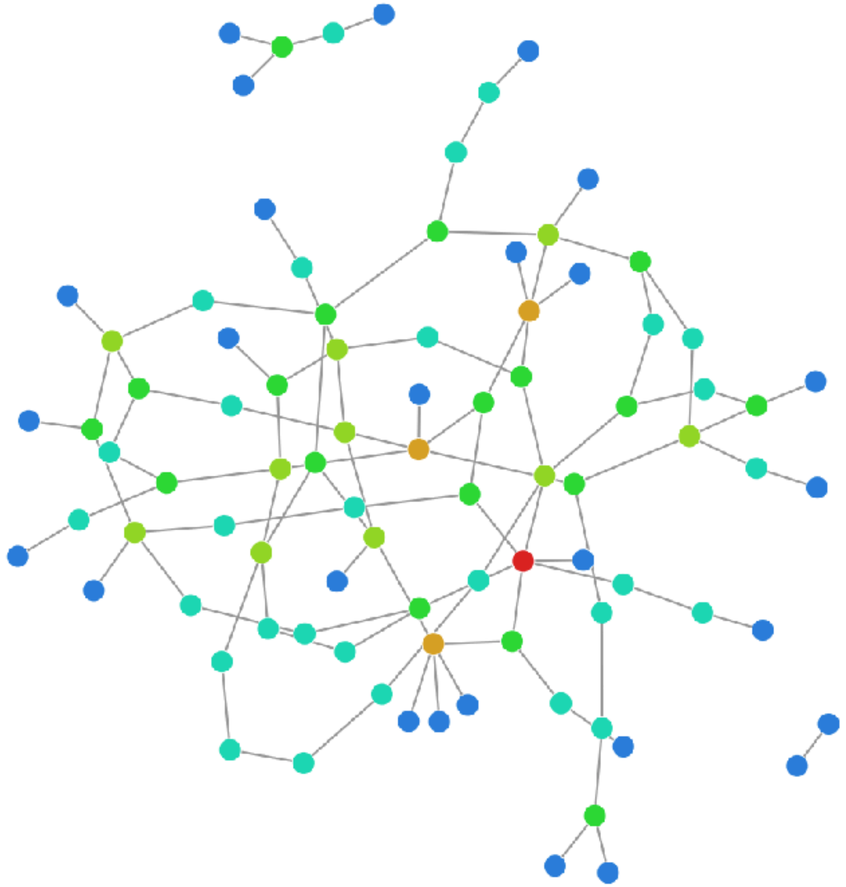
Editor – /Users/restrepj/Dropbox

erdos.m | diversity.m

```
1   n = 50;
2   A = sprand(n,n,1/n);
3   spy(A)
4   webweb(A)
```

Generate matrix
Show matrix
Visualize matrix





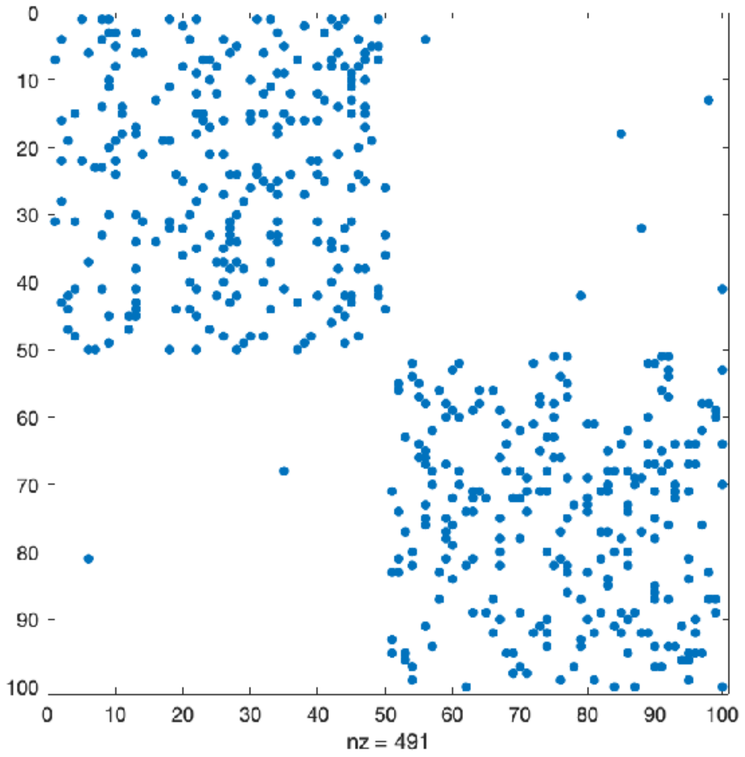
matlab

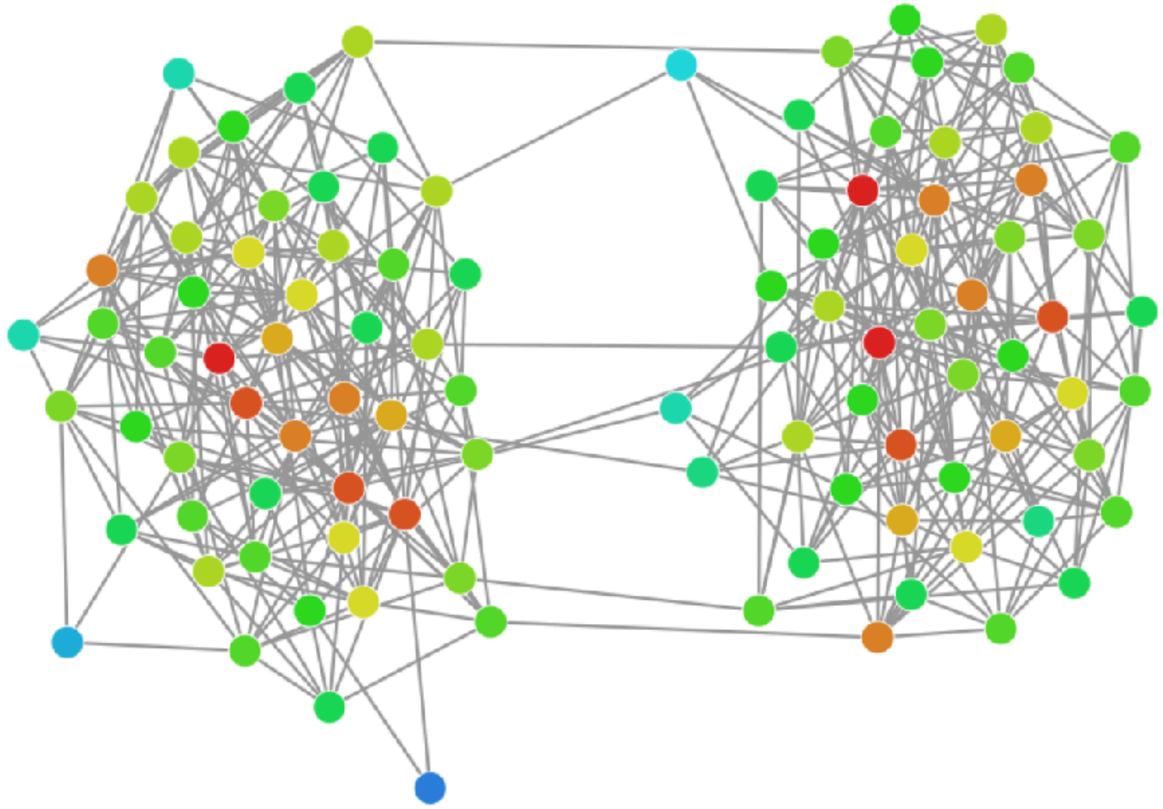
Stochastic Block Model

Editor - /Users/restrepj/Dropbox/TalksDropbox/ICERM/sbm.m *

erdos.m | diversity.m | sbm.m * | spread.m | problem2.m

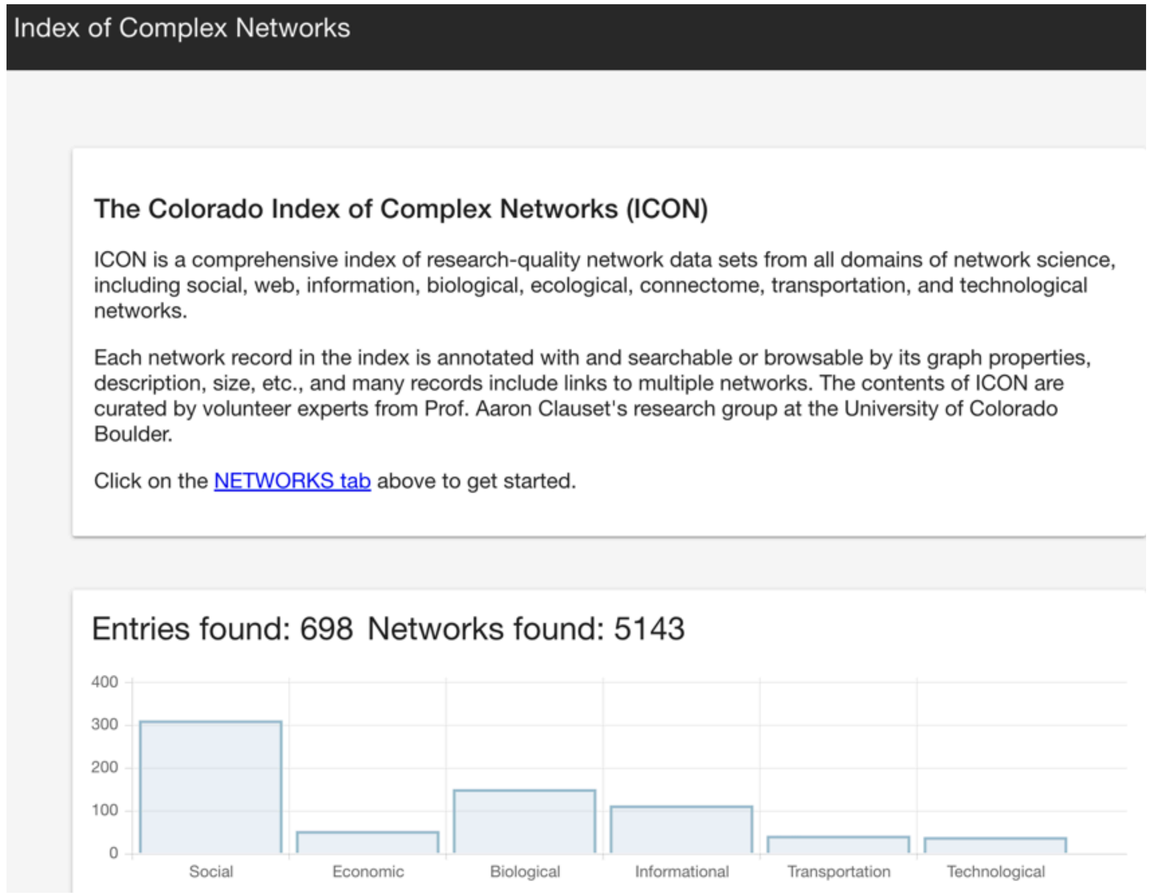
```
1   n = 100;
2   a11 = sprand(n/2,n/2,10/n);     a12 = sprand(n/2,n/2,0.2/n);
3   a21 = sprand(n/2,n/2,0.2/n);    a22 = sprand(n/2,n/2,10/n);
4   a = [a11 a12; a21 a22];
5   a = sparse(a);
6   spy(a)
7   webweb(a)
```

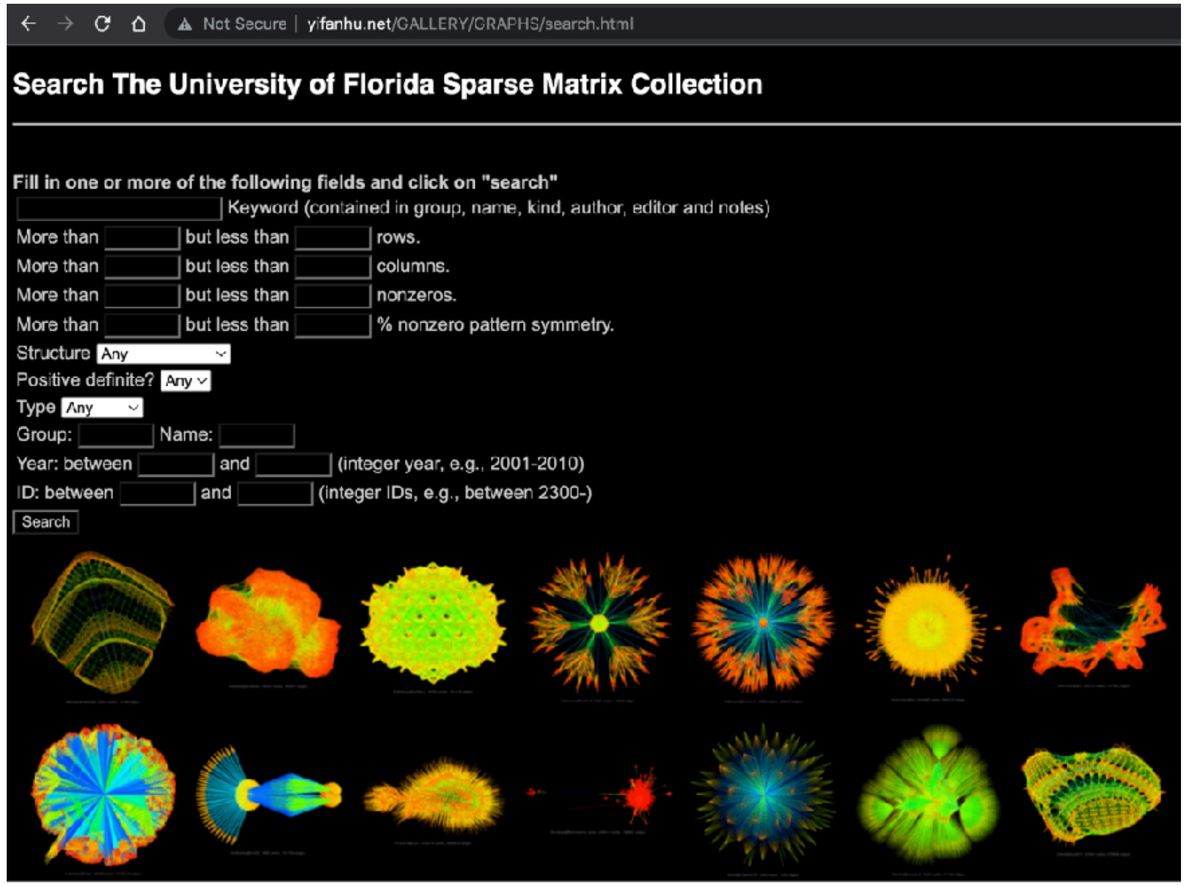




matlab: importing network datasets

Many searchable databases for networks





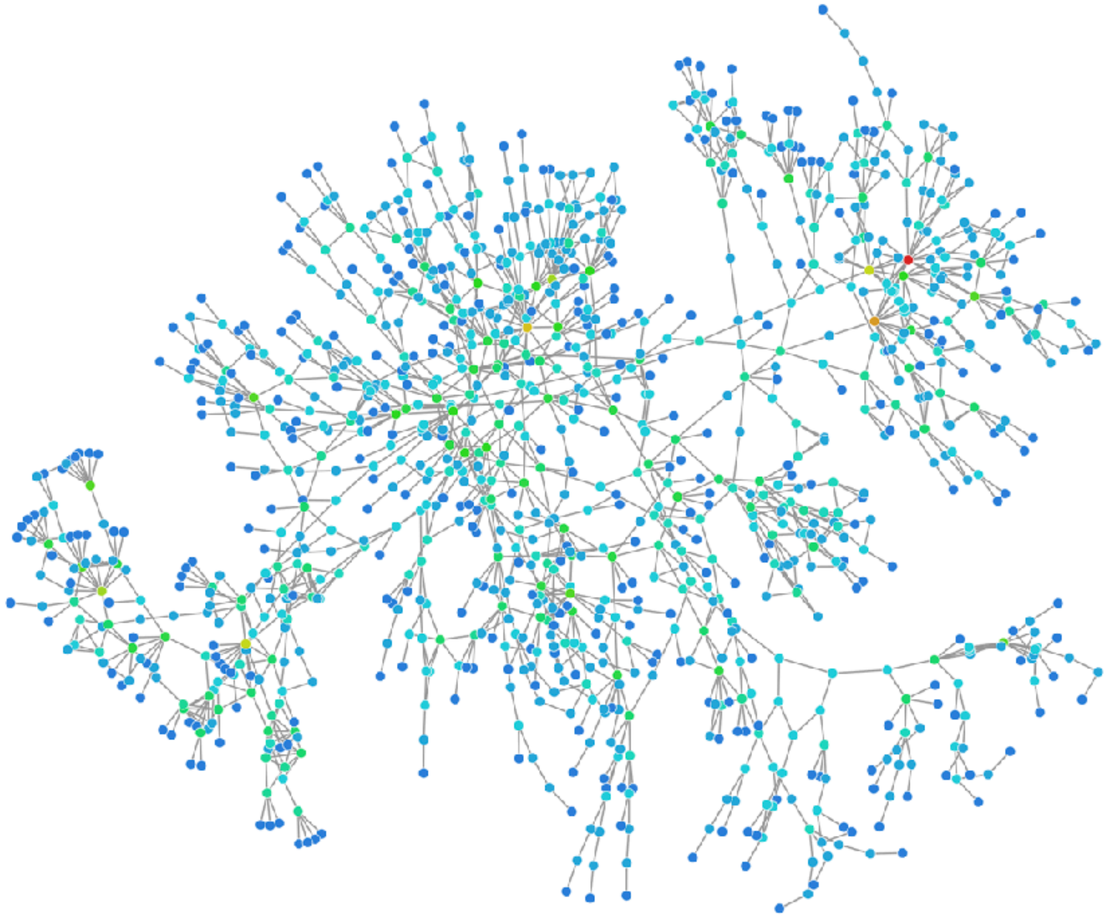
matlab

Importing power-grid network into matlab

Editor – /Users/juanga/Dropbox/TalksDropbox/ICERM/power.m*

problem2.m | diversity.m | erdos.m | sbm.m | power.m* | problem5.m | +

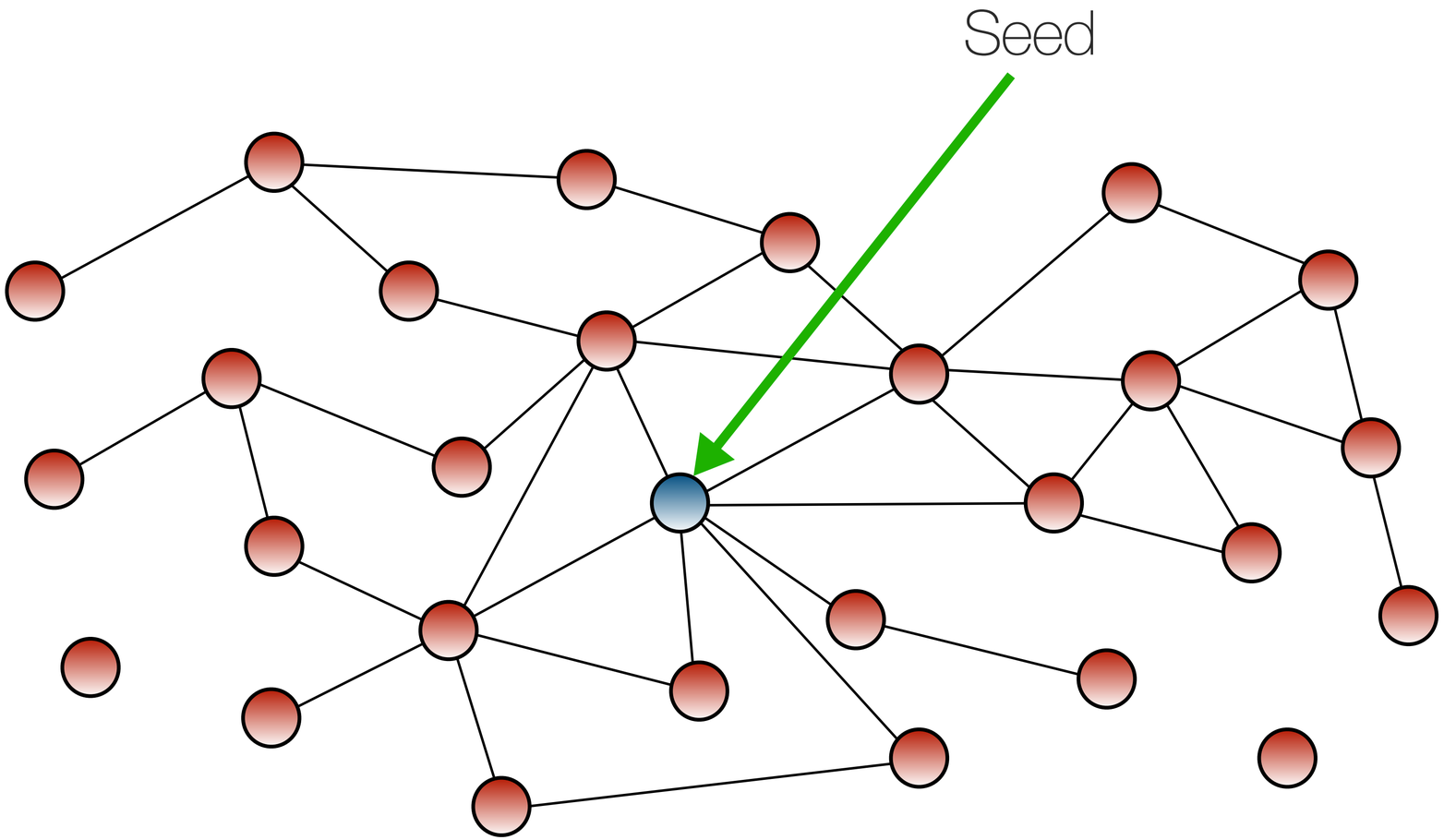
```
1 -     load('1138_bus.mat');     % load dataset
2 -     a = sign(abs(Problem.A));  %create adjacency matrix, make unweighted
3 -     n = length(a);
4 -     webweb(a)
```

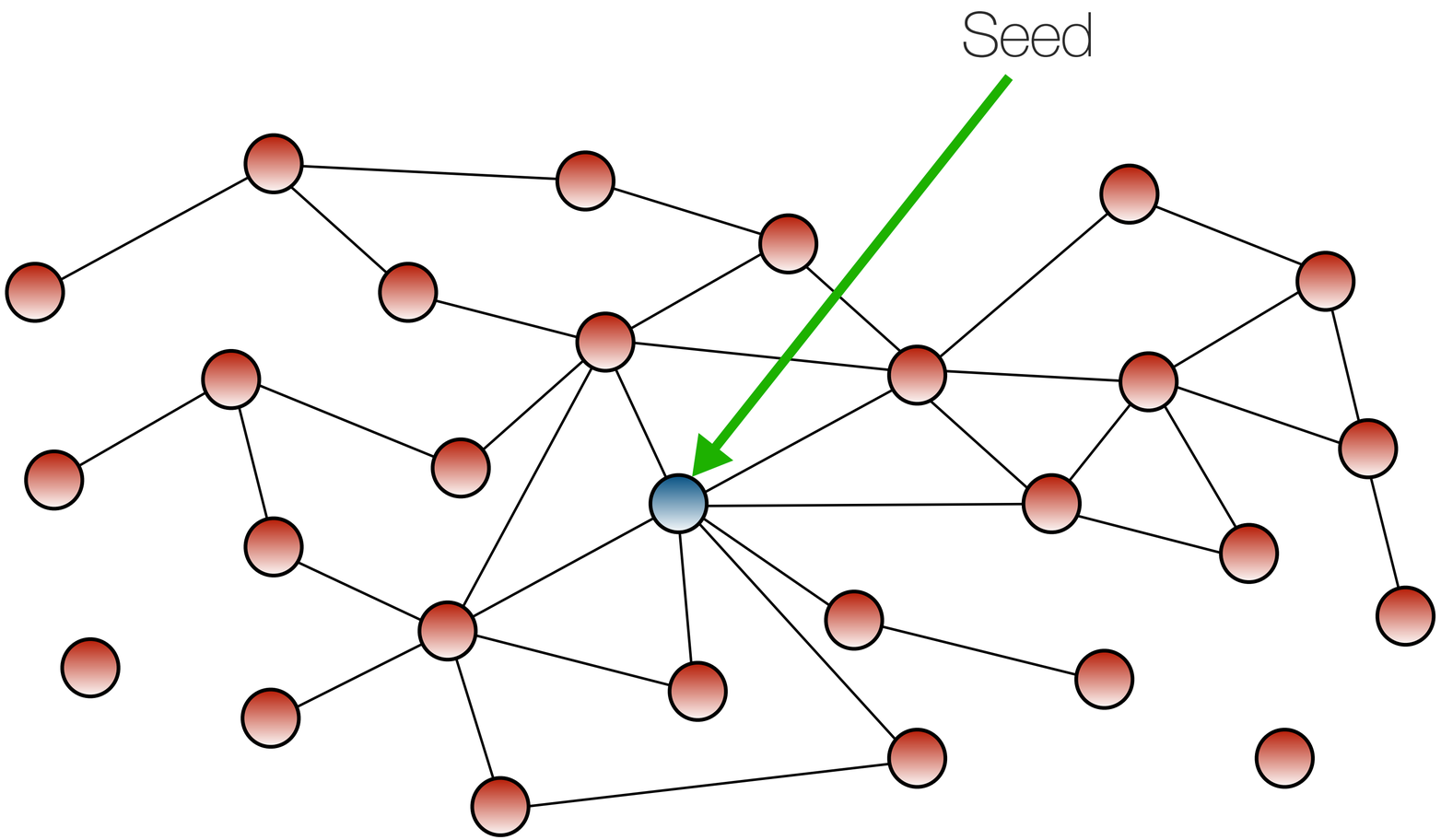


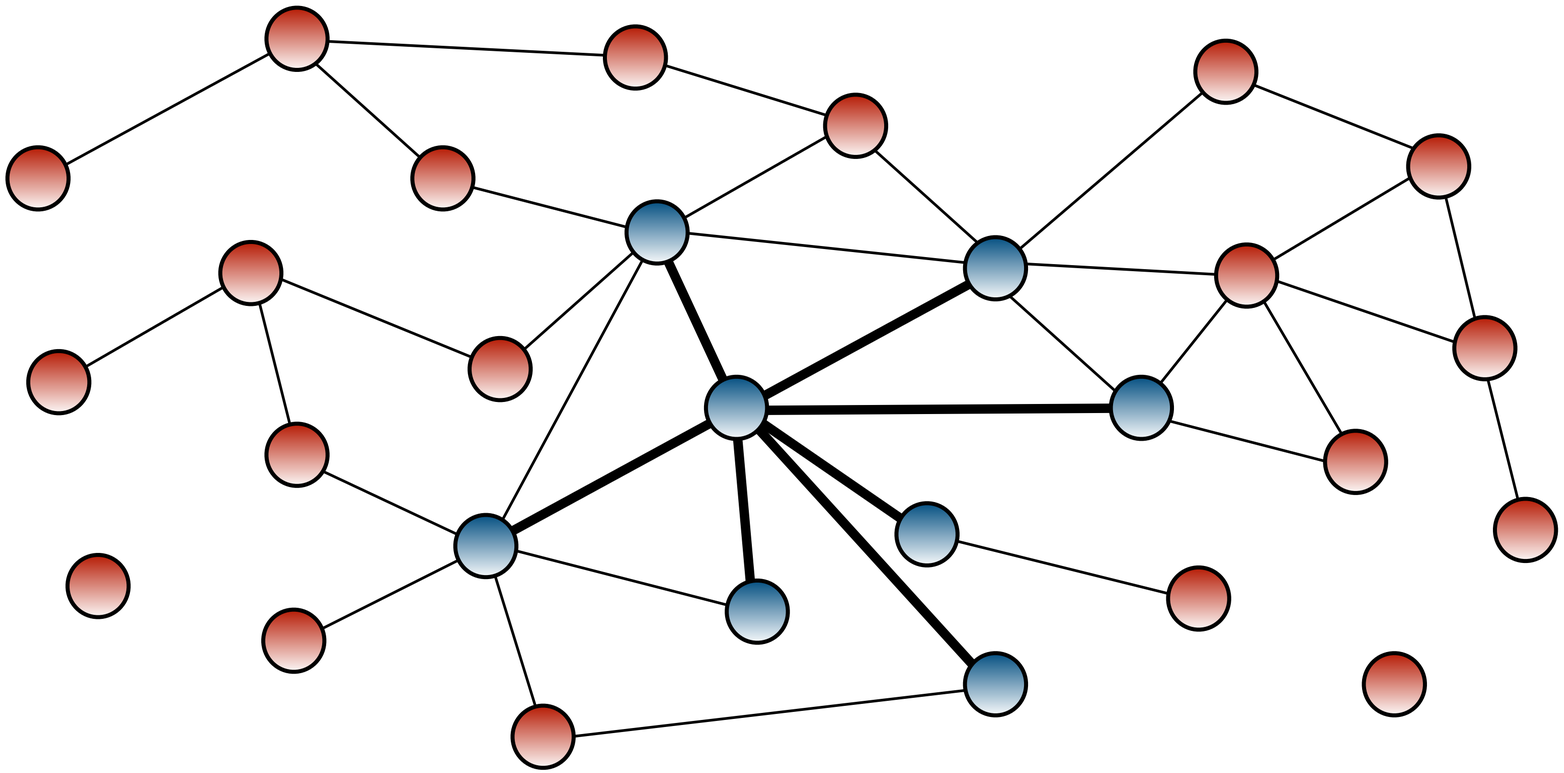
matlab

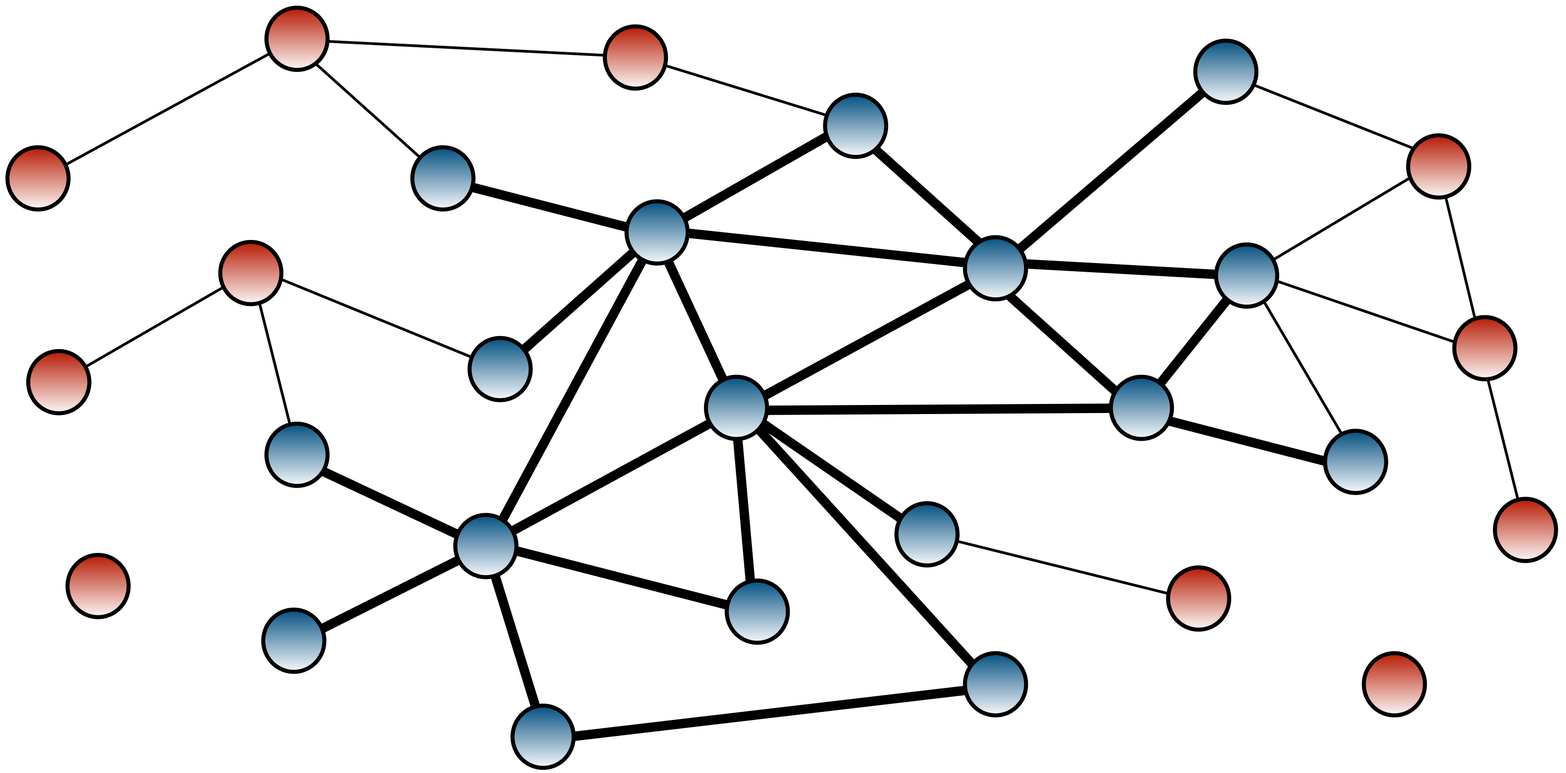
Simulating the simplest spreading process: if a neighbor is infected, you become infected

$x^t(i) = 0$ if node i is not infected at time t, $x^t(i) = 1$ if node i is infected at time t




Seed





matlab

Simulating the simplest spreading process: if a neighbor is infected, you become infected

$x^t(i) = 0$ if node i is not infected at time t, $x^t(i) = 1$ if node i is infected at time t

$$x^{t+1}(i) = \text{sign}\left(\sum_{j=1}^{N} A_{ij}x^t(j)\right) = (\text{sign}(A\mathbf{x}))_i$$

Seed **one** node at $t = 0$ and iterate until $t = 5$

Editor - /Users/juanga/Dropbox/TalksDropbox/ICERM/spread.m*

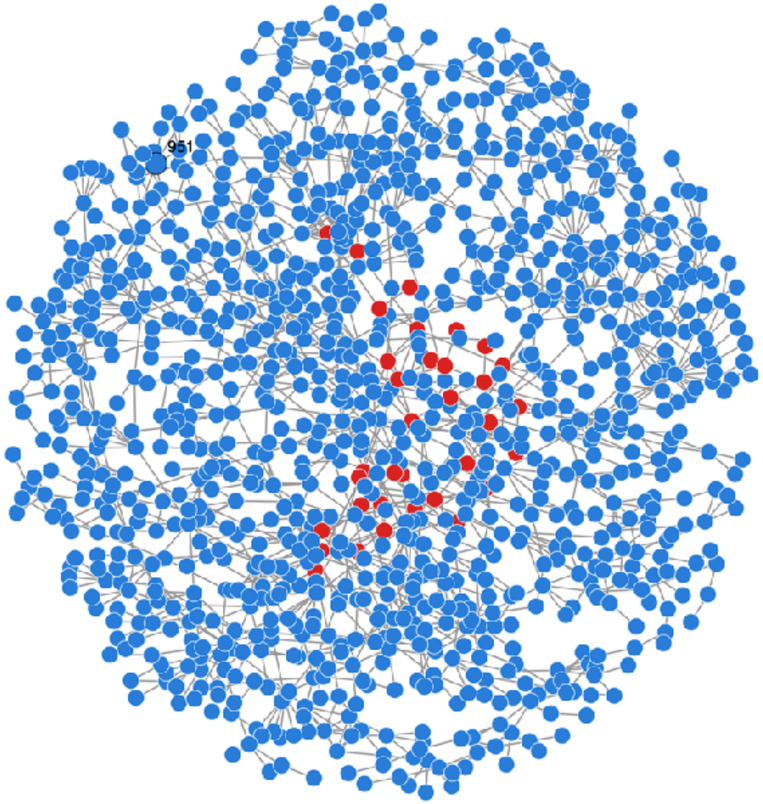
erdos.m | sbm.m | power.m | problem2.m | diversity.m | spread.m*

```
1  load('1138_bus.mat');    % load dataset
2  a = sign(abs(Problem.A));  %create adjacency matrix, make unweighted
3  n = length(a);
4
5  [u lam] = eigs(a,1);
6
7  [centrality i] = max(abs(u));
8
9  x = zeros(n,1);
10 x(i) = 1;
11
12 for t = 1:5
13     x = sign(a*x);
14 end
15
16 x = zeros(n,1);
17 x(1) = 1;
18
19 for t = 1:5
20     x = sign(a*x);
21 end
```

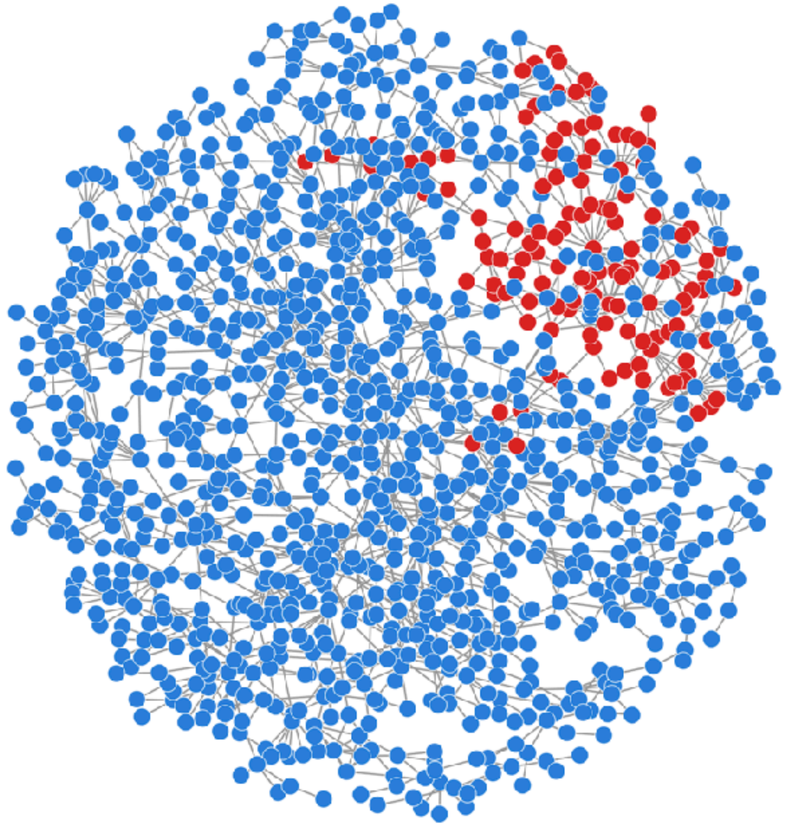
(Not optimized)

matlab

Simulating the simplest spreading process



low eigenvector centrality seed



high eigenvector centrality seed

Talk outline

Part 2

- ✓ Modeling processes on networks: SIS example
- ✓ Creating, importing and visualizing networks in matlab and python
- ▸ Example: modeling **diversity** and productivity

Diversity-productivity modeling example

Diversity and Productivity in Production Teams

Barton H. Hamilton, Jack A. Nickerson, Hideo Owan
Advances in the Economic Analysis of Participatory and Labor-Managed Firms
ISBN: 978-1-78190-220-2, eISBN: 978-1-78190-221-9
ISSN: 0885-3339
Publication date: 11 September 2012 Reprints & Permissions

Abstract

The popular press often touts workforce demographic diversity as profit enhancing because it may reduce the firm's communication costs with particular segments of customers or yield greater team problem-solving abilities. On the other hand, diversity also may raise communication costs within teams, thereby retarding problem solving and lowering productivity. Unfortunately, there is little empirical research that disentangles the above countervailing effects. Diversity in ability enhances the team productivity if there is significant mutual learning and collaboration within the team, while demographic diversity may harm productivity by making learning and peer pressure less effective and increasing team-member turnover. We evaluate these propositions using a novel panel data from a garment plant that shifted from individual piece rate to group piece rate production over three years. Because we observe individual productivity data, we are able to econometrically distinguish between the impacts of diversity in worker abilities and demographic diversity. Teams with more heterogeneous worker abilities are more productive at the plant. Holding the distribution of team ability constant, teams composed of only one ethnicity (Hispanic workers in our case) are more productive, but this finding does not hold for marginal changes in team composition. We find little evidence that workers prefer to be segregated; demographically diverse teams are no more likely to dissolve, holding team productivity (and hence pay) constant, than homogeneous teams.

Academy of Management Journal, Vol. 64, No. 5 | Articles

The Effects of Racial Diversity Congruence between Upper Management and Lower Management on Firm Productivity

is companion of

Orlando C. Richard, María del Carmen Triana and Mingxiang Li

Impact of Workplace Diversity

Elizabeth Foma
University of Guam
School of Business and Public Administration
P. O. Box 5169 UOG Station
Mangilao, Guam 96923
efoma@uguam.uog.edu; drmayuk@yahoo.com

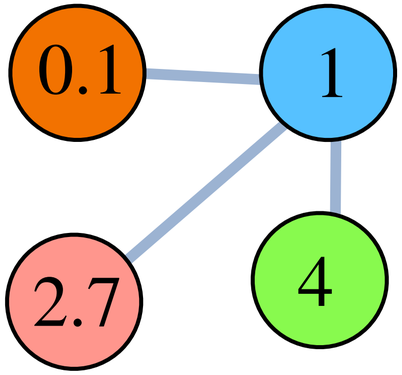
Review of Integrative Business & Economics Research

Diversity-productivity modeling example

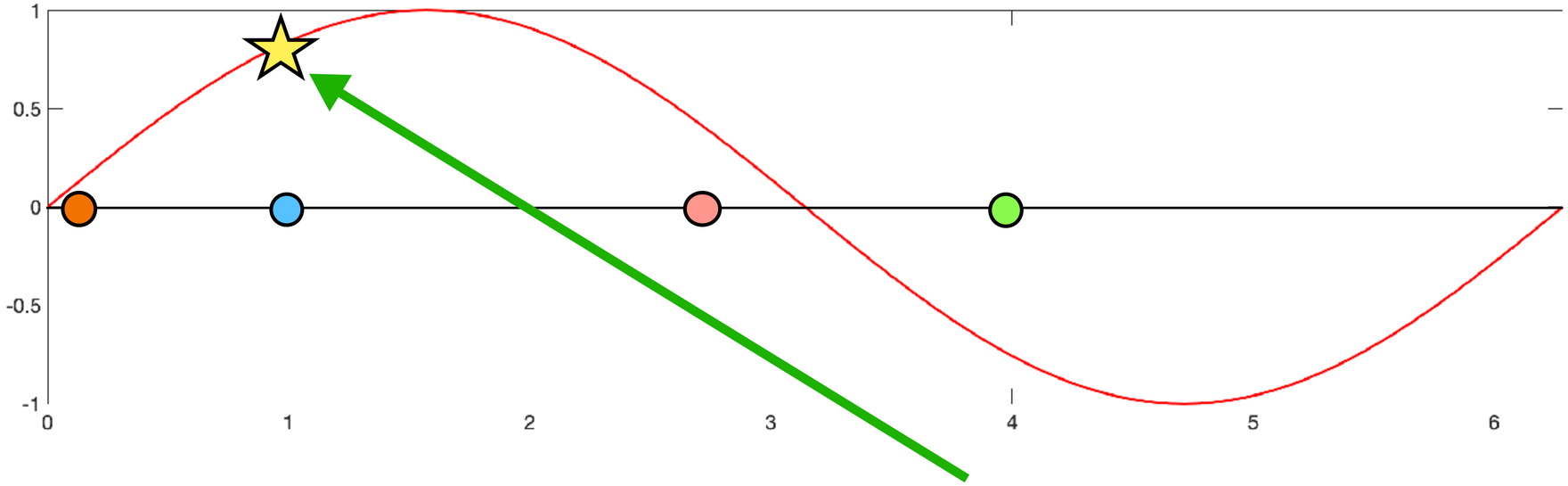
Suppose we have N workers (nodes), and they collaborate (links) in projects.

Each worker has a characteristic variable x (could represent race, political ideology, gender, skills, etc.)

Teams of workers try to find the solution to a particular problem. We will model the "problem" as maximizing the function $\sin(x - b)$

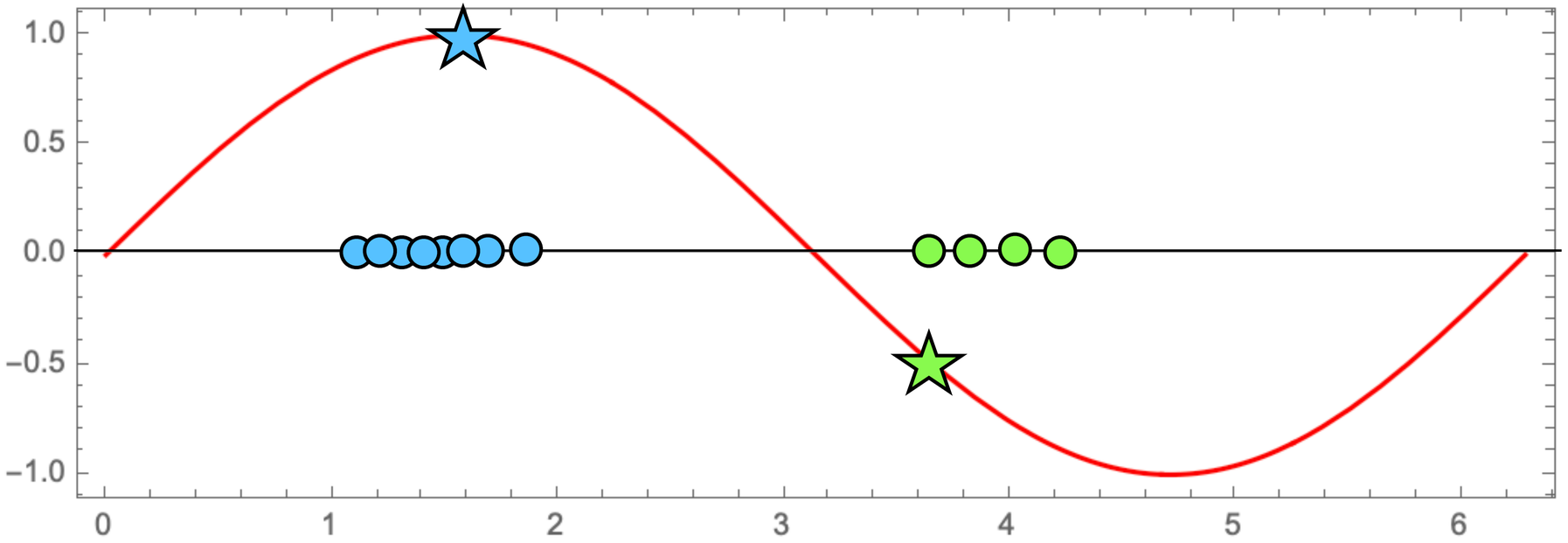


Team



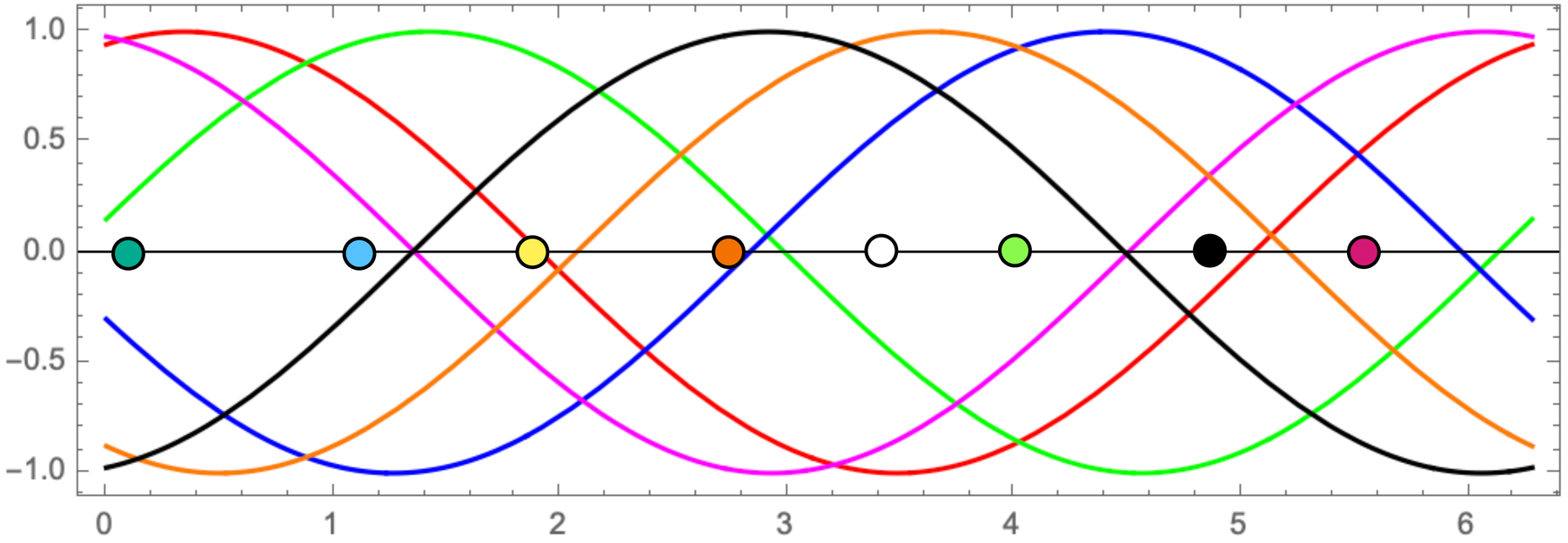
Team's "solution" = **productivity**

Modeling diverse team adaptation



Single problem, non-diverse teams:

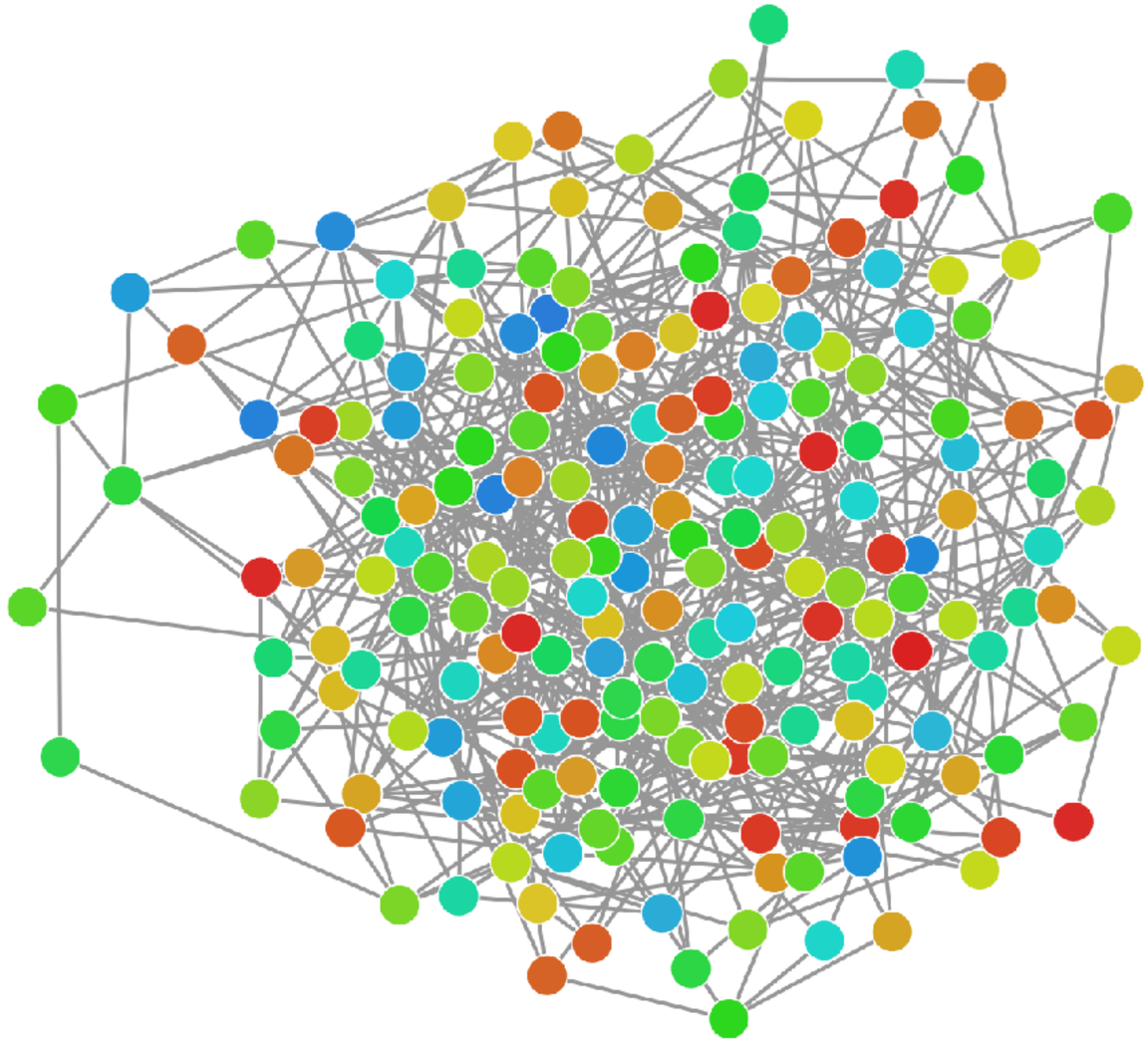
good or bad solution, depends on problem



Multiple problems, diverse team:

consistently good solutions

Diverse workspace

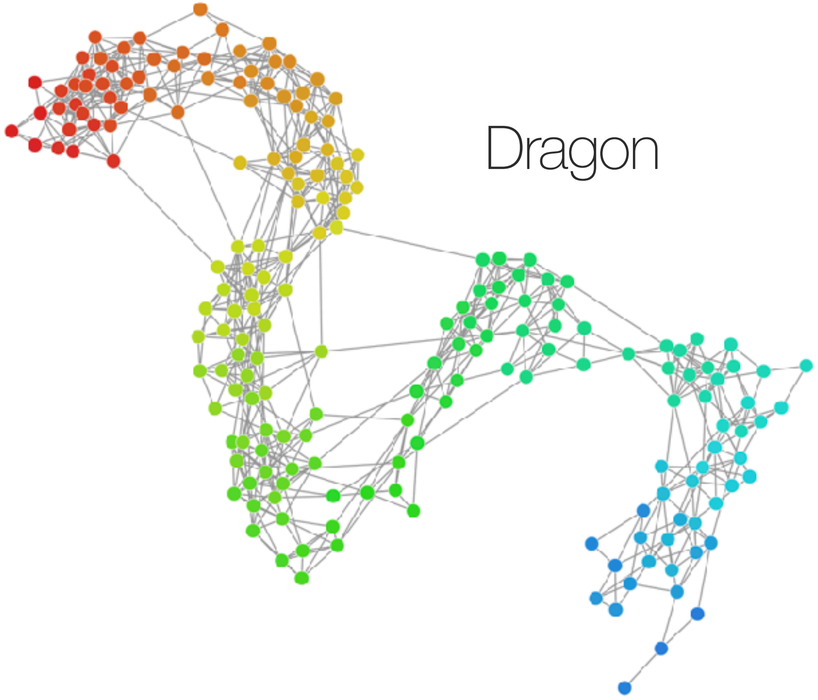


Non-diverse workspaces

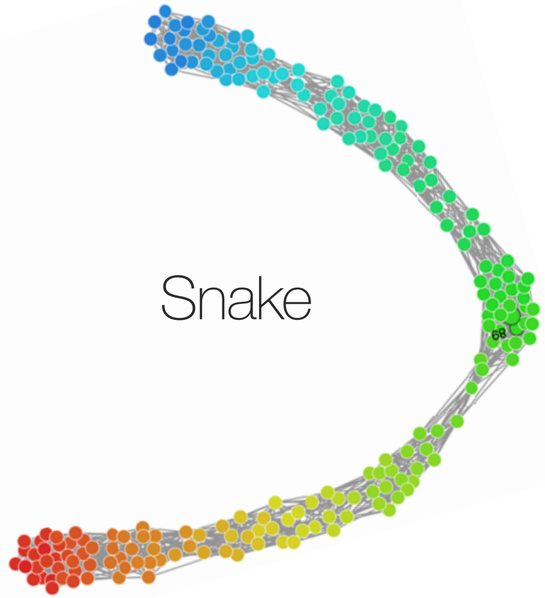
Connect nodes i, j with probability $p_{i,j} = f(|x^0(i) - x^0(j)|)$

Communities

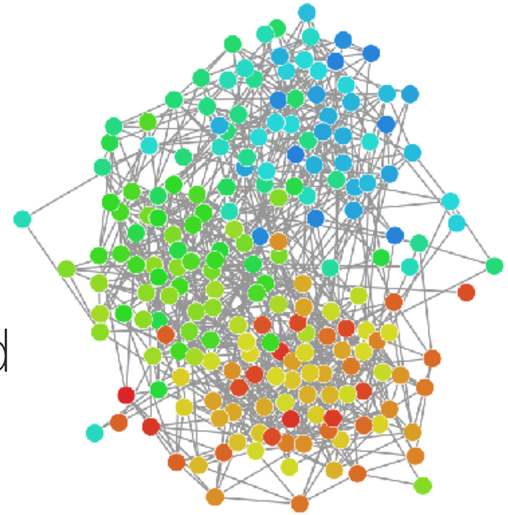
Dragon



Snake



Biased



Diversity modeling example

Dynamics

Every time step:

Choose a random node i

Choose a random "problem", that is, a random b for the function $h(x) = \sin(x - b)$

Find the team's productivity $\max_{A_{ij} \neq 0} \{h(x_i)\}$

If productivity is less than -0.5, rewire one connection from node i at random

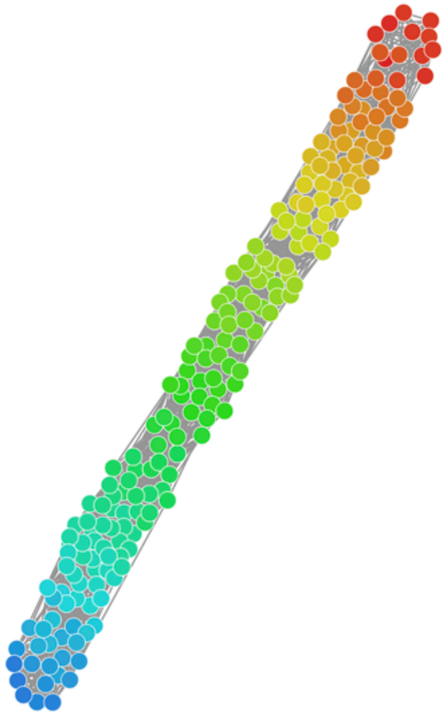
What if nodes still prefer to connect to similar nodes?

Pick another random node j

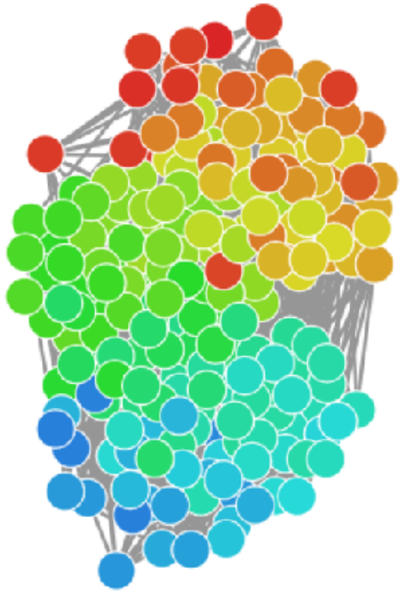
Node j chooses a neighbor k at random and rewires to random node m if $|x_j - x_m| < |x_j - x_k|$

Diversity modeling example

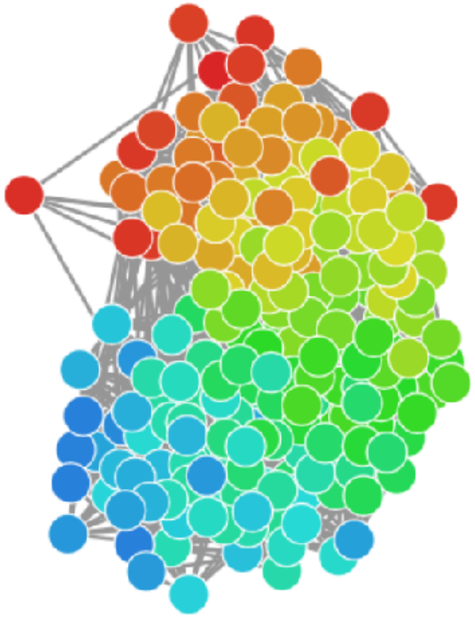
$t = 0$



$t = 12000$



$t = 30000$

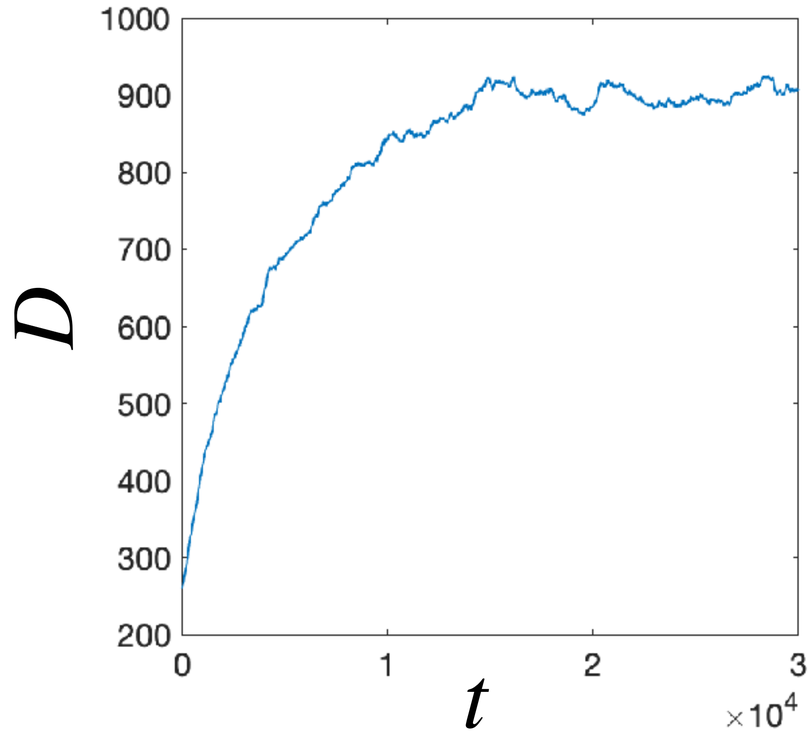


Diversity stagnates

Diversity modeling example

In our example, we can define diversity as $$D = \sum_{i,j} A_{ij} |x_i - x_j|$$

This accounts for **diversity across links**, not at the population level

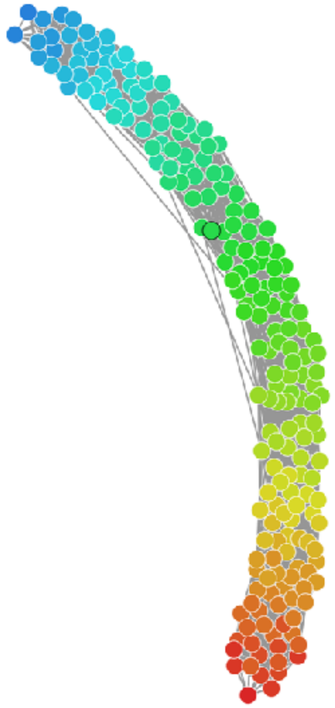


Diversity increases until the benefits of productivity increase are negated by the tendency to connect to similar people

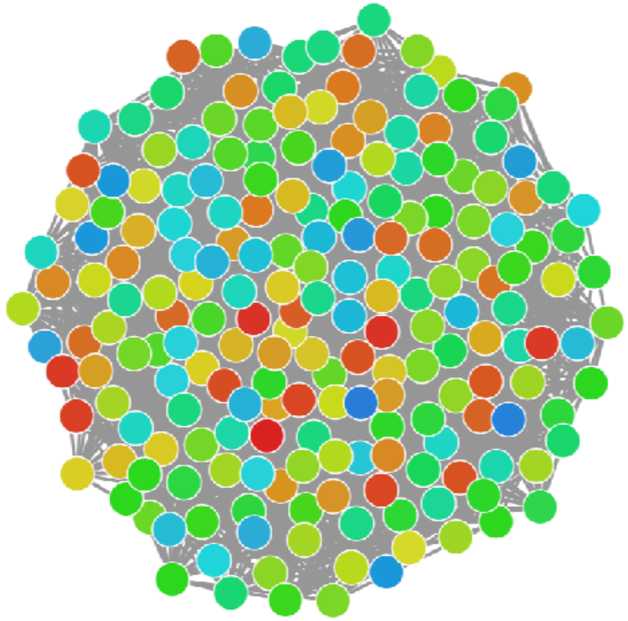
Diversity modeling example

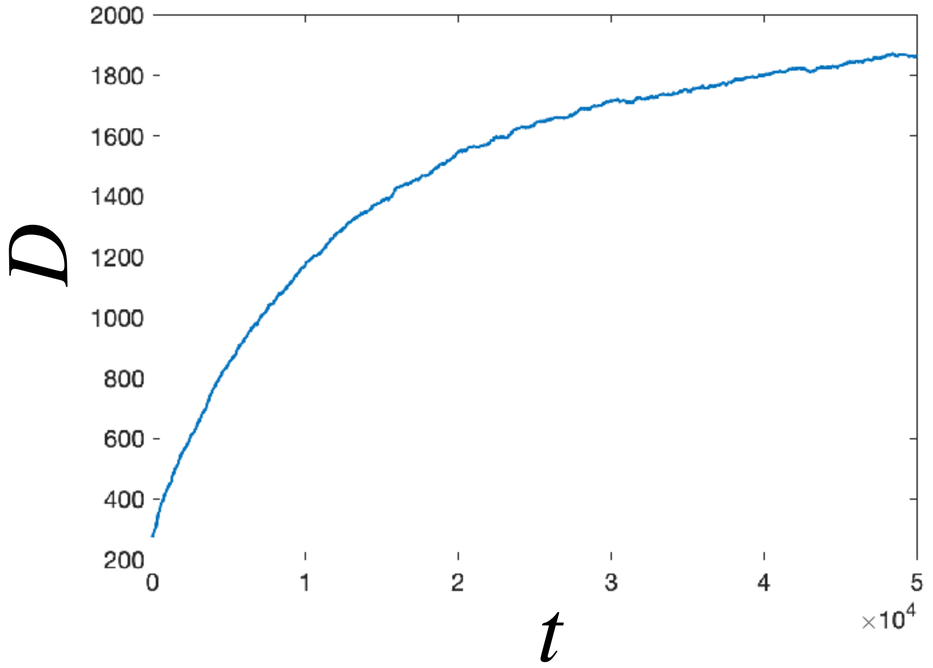
If nodes do not rewire seeking similar characteristics, diversity keeps increasing

$t = 0$



$t = 30000$





Diversity modeling example

This was a made-up example where we know what is going to happen.

If anyone is interested in collaborating or has ideas to make it more interesting, please let me know!

Talk outline

Part 2

- ✓ Modeling processes on networks: SIS example
- ✓ Creating, importing and visualizing networks in matlab and python
- ✓ Example: creating model of **diversity** and productivity

Q/A