

# Four Lectures on Networks

Aaron Clauset

 @aaronclauset

Assistant Professor of Computer Science

University of Colorado Boulder

External Faculty, Santa Fe Institute

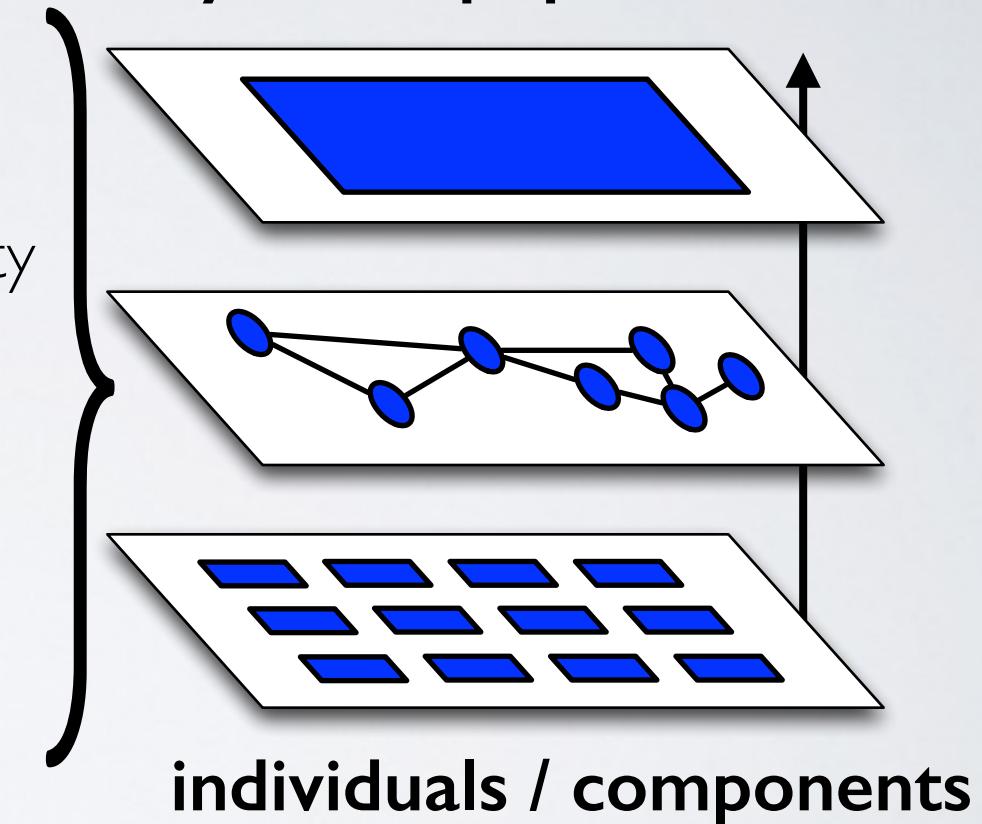
**lecture I: what are networks and how do we talk about them?**

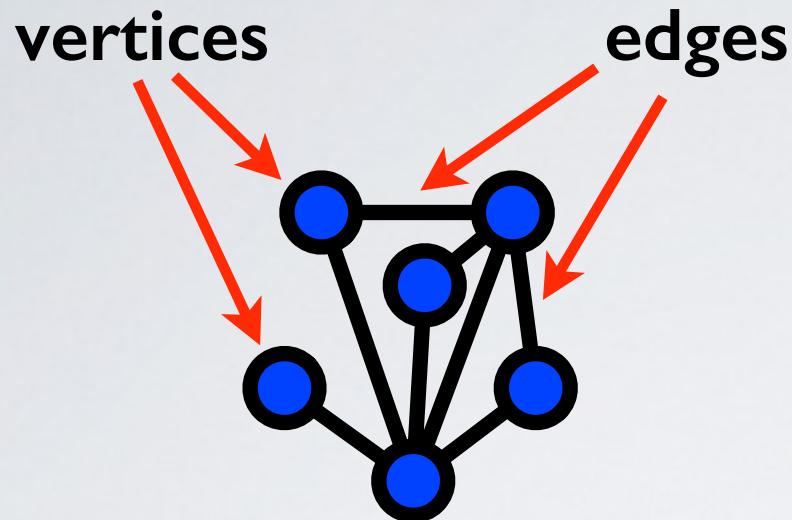
**what are networks?**

## what are networks?

- an approach
- a mathematical representation
- provide structure to complexity
- *structure above*  
individuals / components
- *structure below*  
system / population

**system / population**





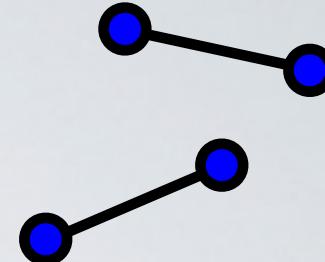
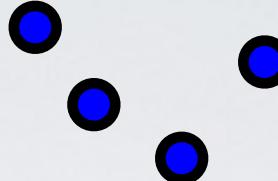
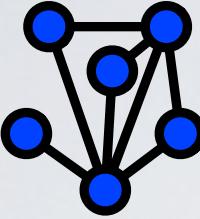
**what is a vertex?**

$V$  distinct objects (vertices / nodes / actors)

**when are two vertices connected?**

$$E \subseteq V \times V$$

pairwise relations (edges / links / ties)

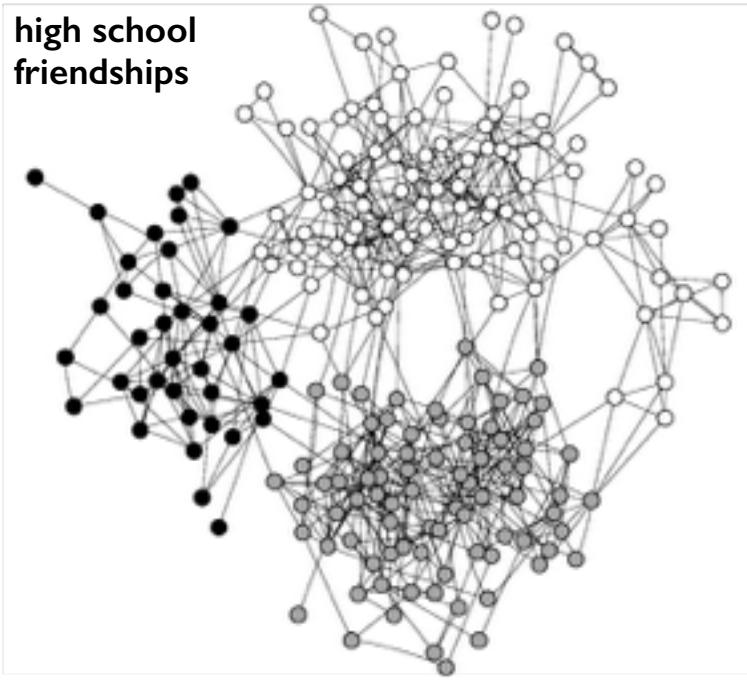


	<b>network</b>	<b>vertex</b>	<b>edge</b>
informational	Internet(1)	computer	IP network adjacency
	Internet(2)	autonomous system (ISP)	BGP connection
	software	function	function call
	World Wide Web	web page	hyperlink
	documents	article, patent, or legal case	citation
transportation	power grid transmission	generating or relay station	transmission line
	rail system	rail station	railroad tracks
	road network(1)	intersection	pavement
	road network(2)	named road	intersection
social	airport network	airport	non-stop flight
	friendship network	person	friendship
	sexual network	person	intercourse
biological	metabolic network	metabolite	metabolic reaction
	protein-interaction network	protein	bonding
	gene regulatory network	gene	regulatory effect
	neuronal network	neuron	synapse
	food web	species	predation or resource transfer

# social networks

**vertex:** a person

**edge:** friendship, collaborations, sexual contacts, communication, authority, exchange, etc.

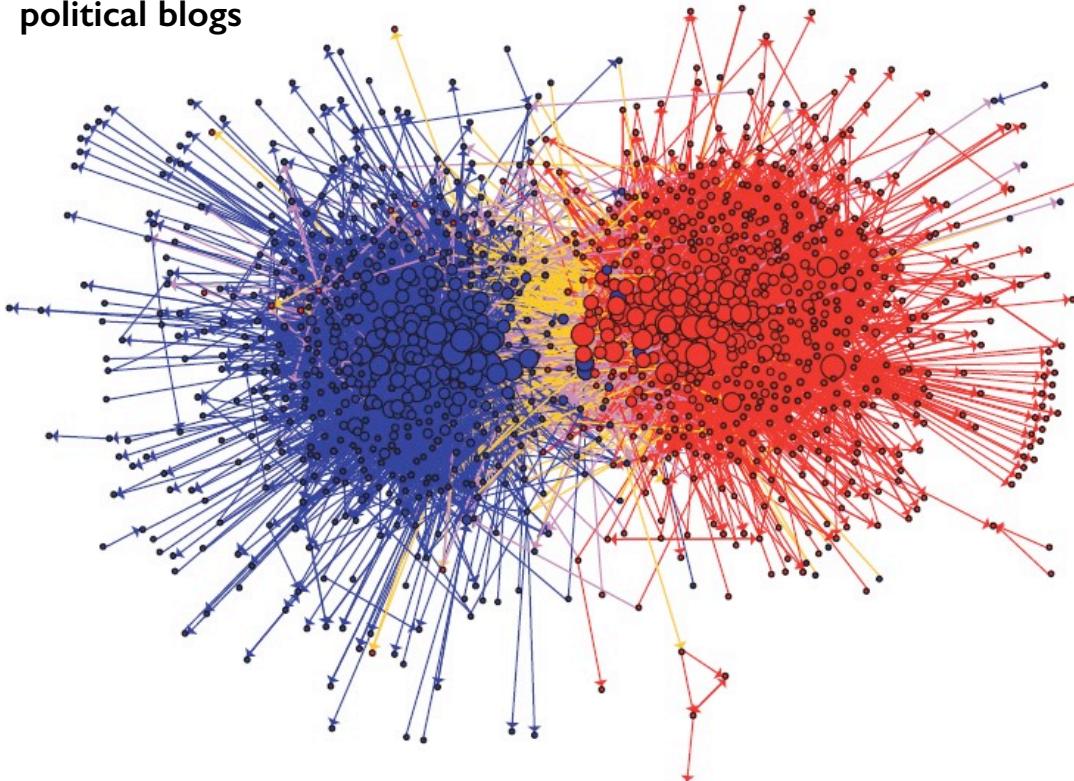


# information networks

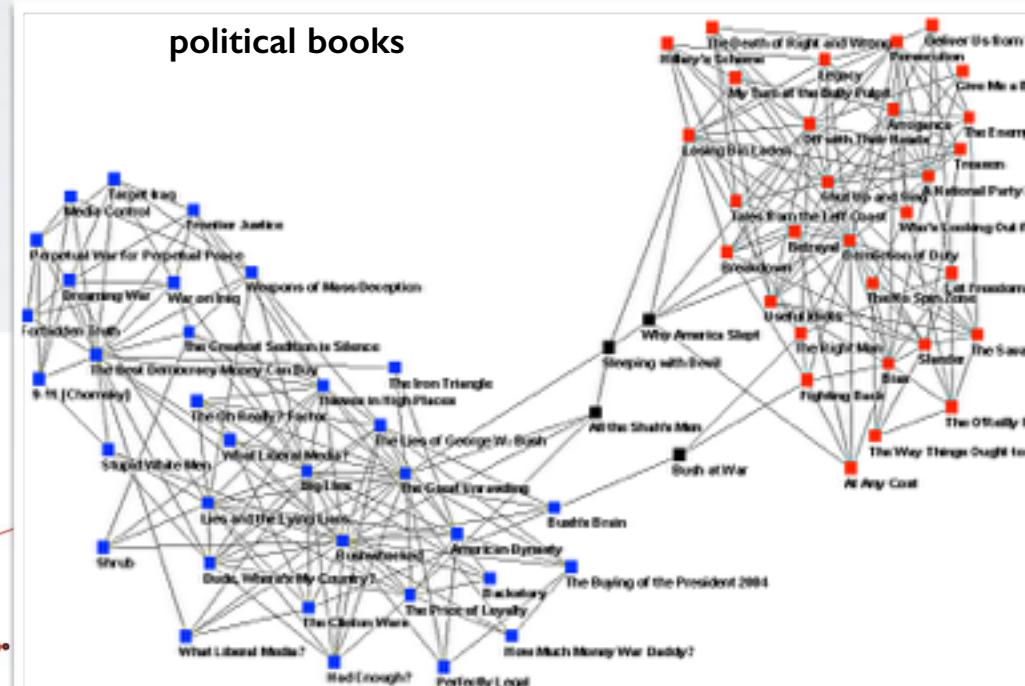
**vertex:** books, blogs, webpages, etc.

**edge:** citations, hyperlinks,  
recommendations, similarity, etc.

political blogs



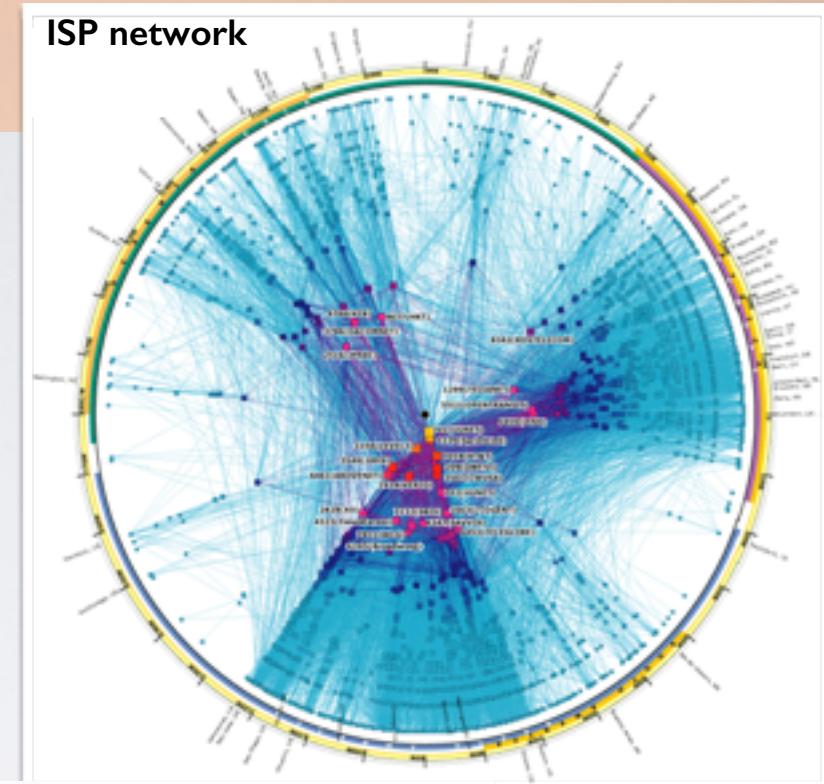
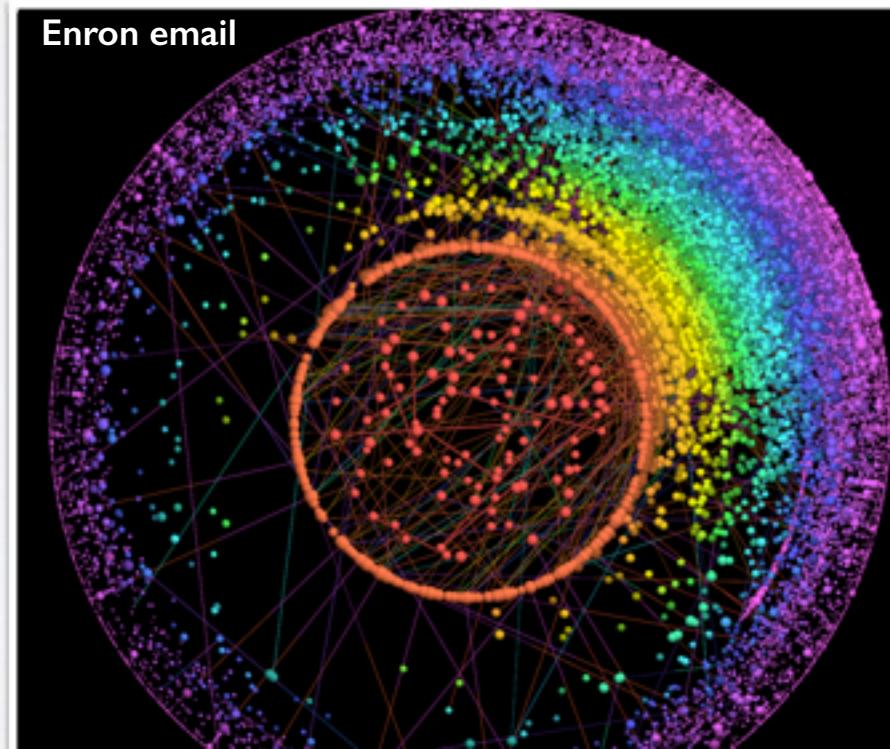
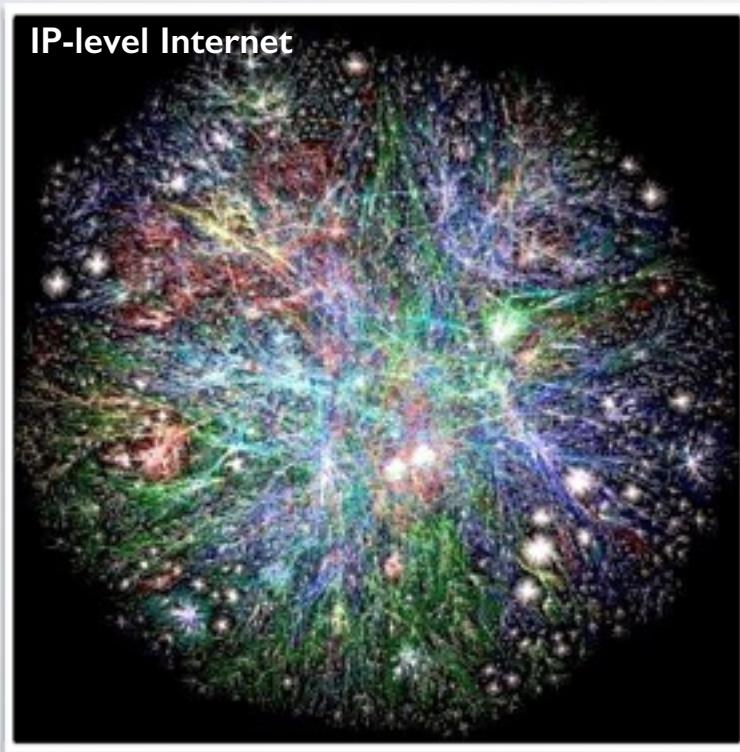
political books



# communication networks

**vertex:** network router, ISP, email address, mobile phone number, etc.

**edge:** exchange of information



# transportation networks

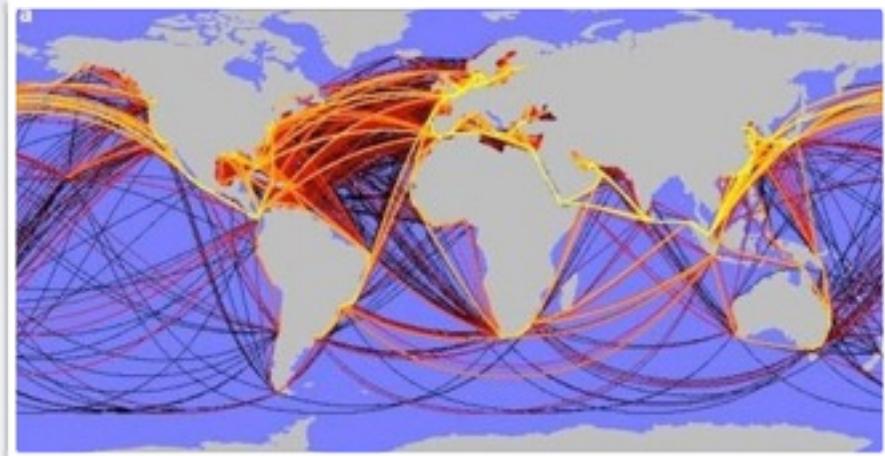
**vertex:** city, airport, junction, railway station, river confluence, etc.

**edge:** physical transportation of material



US Interstates

global shipping



global air traffic

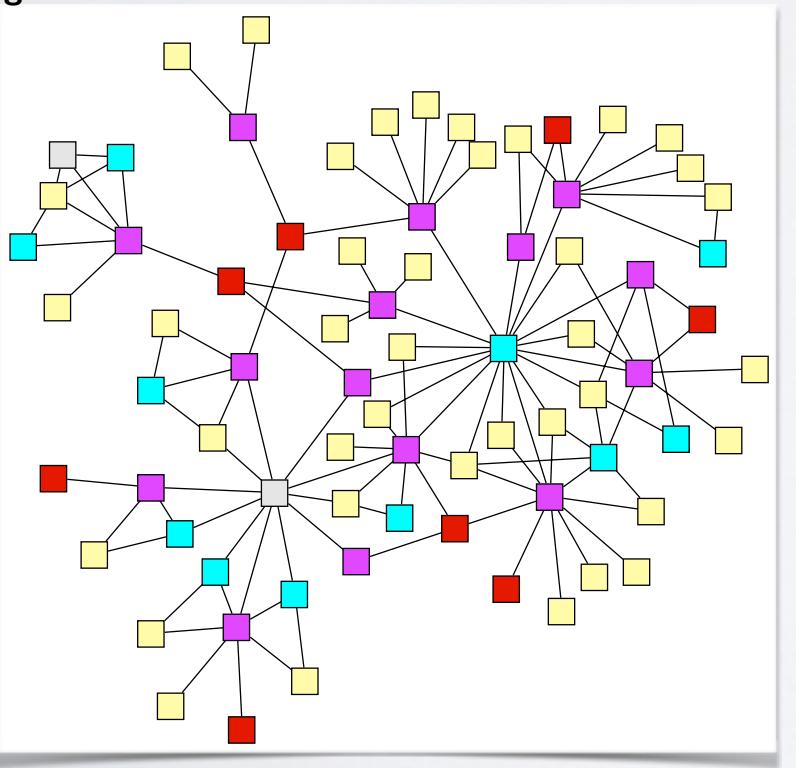


# biological networks

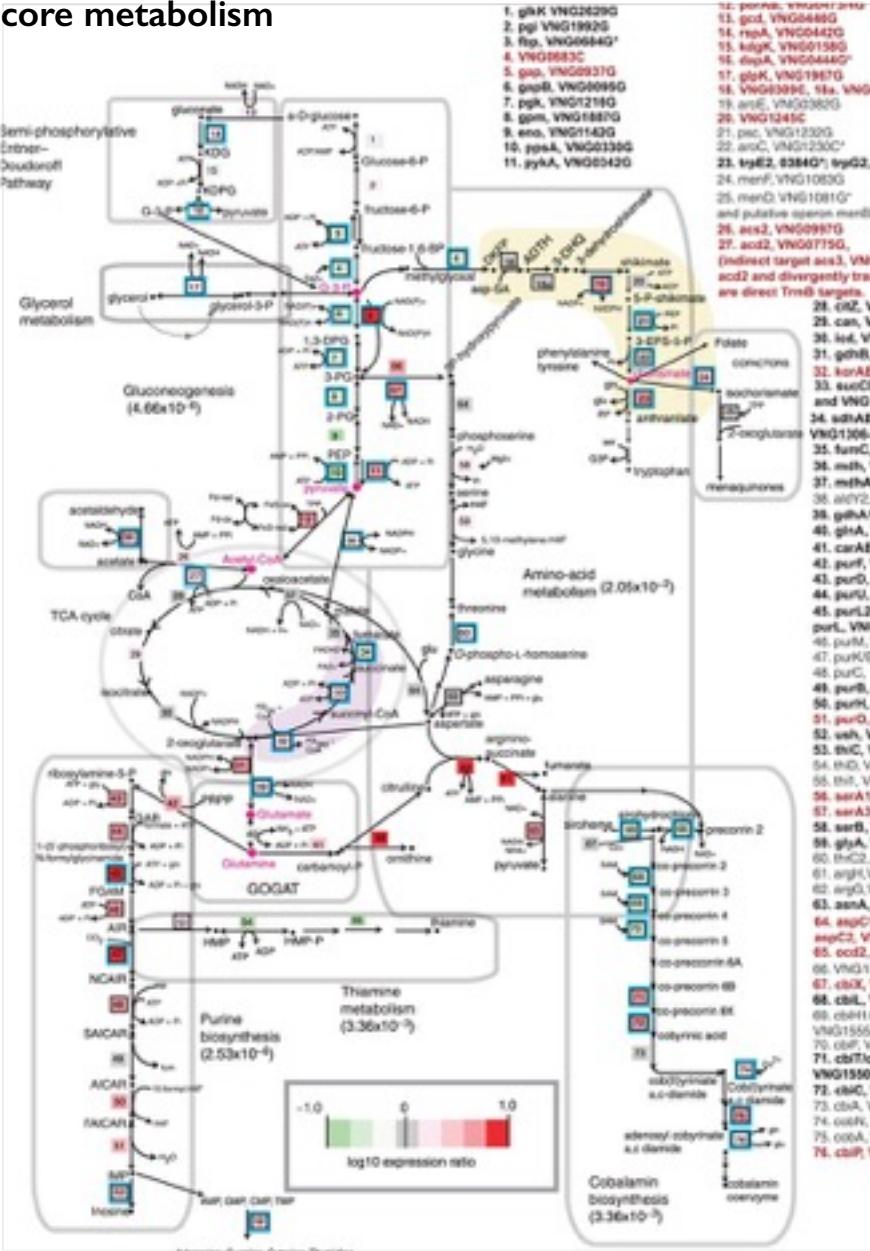
**vertex:** species, metabolic, protein, gene, neuron, etc.

**edge:** predation, chemical reaction, binding, regulation, activation, etc.

## grassland foodweb



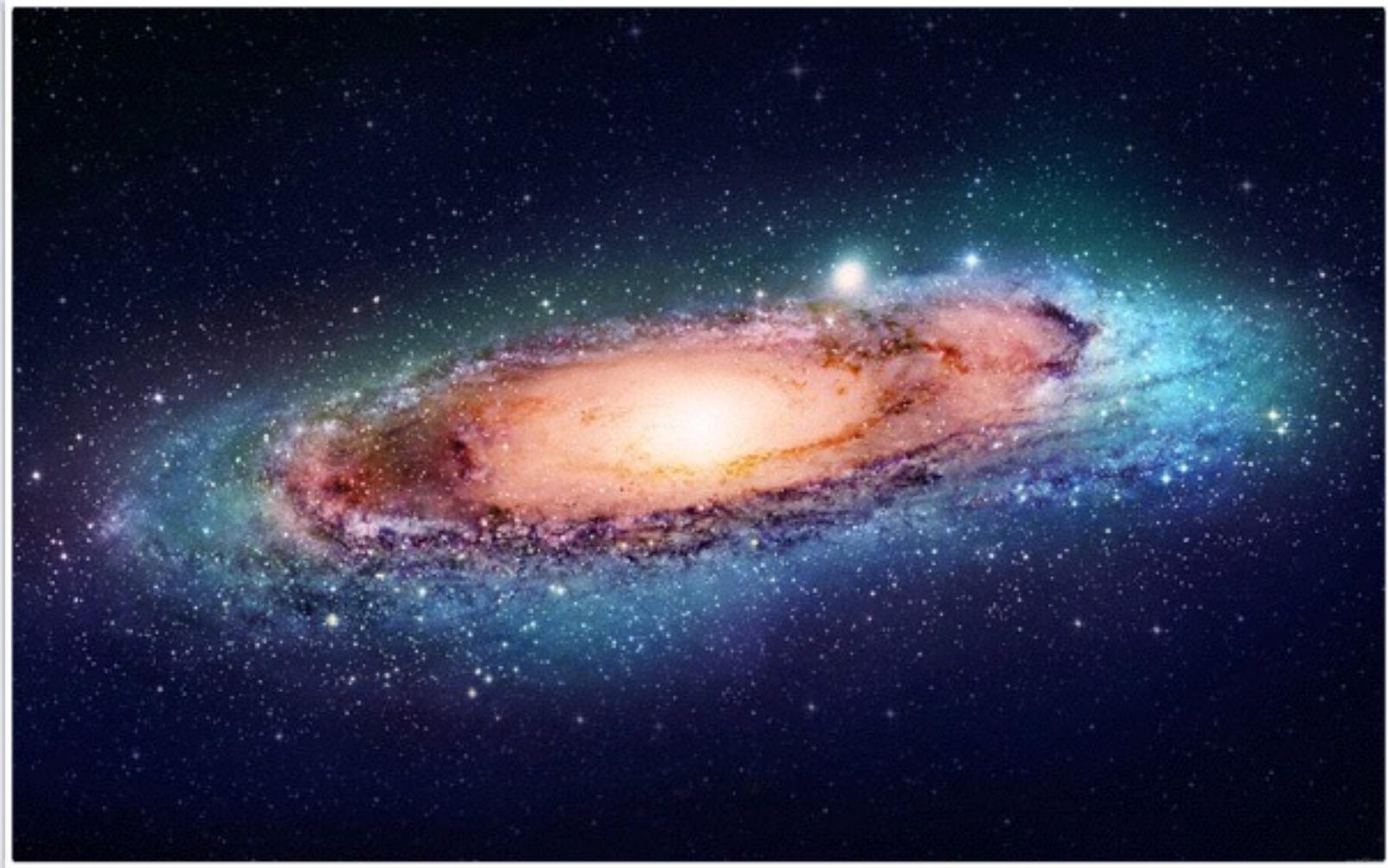
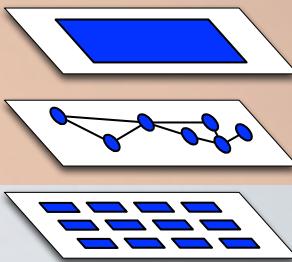
## core metabolism



**what's a network?**

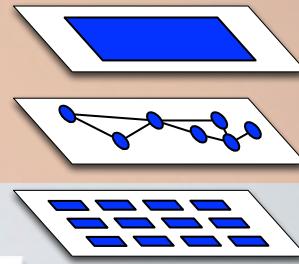
**pop quiz**

# what's a network?



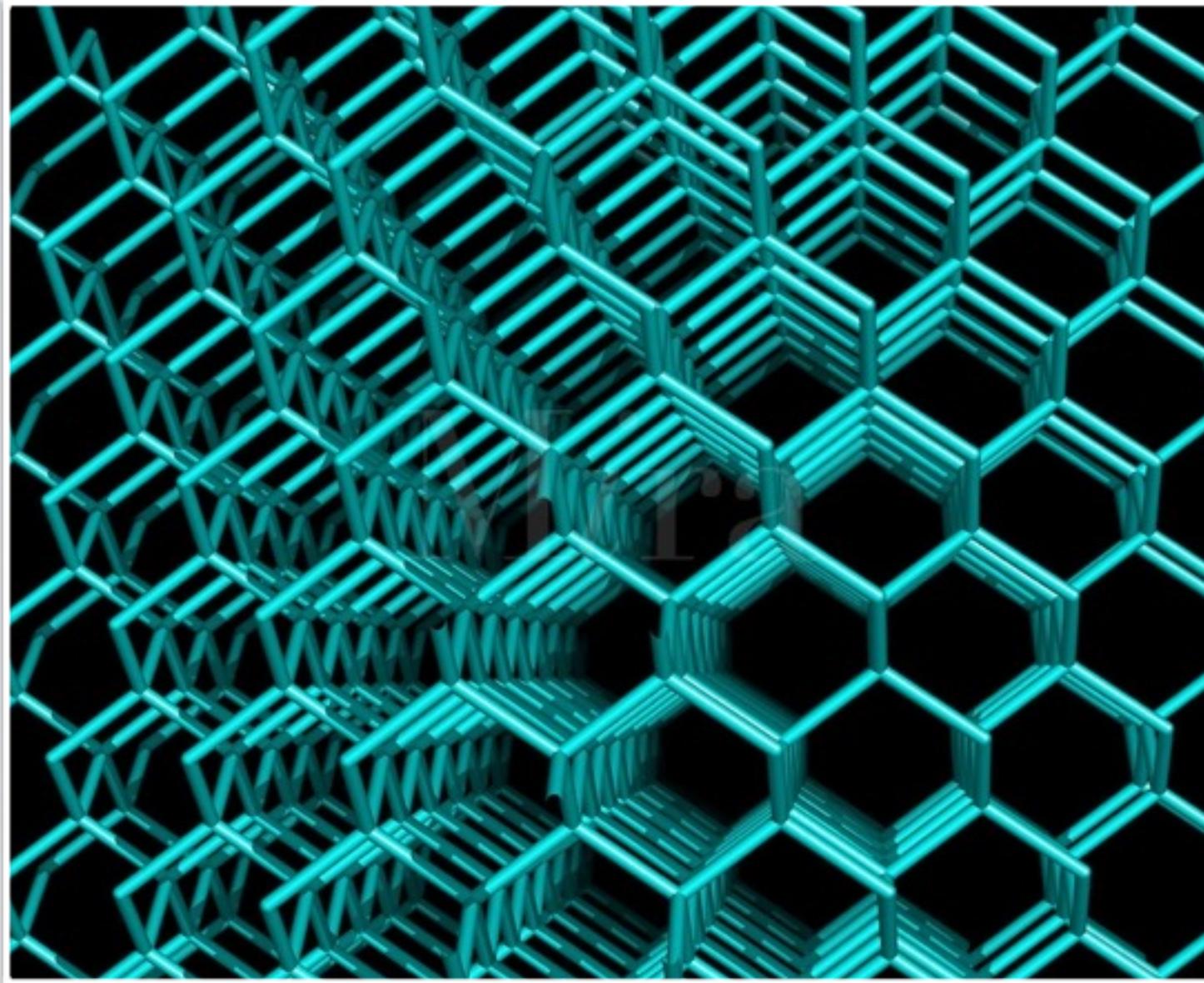
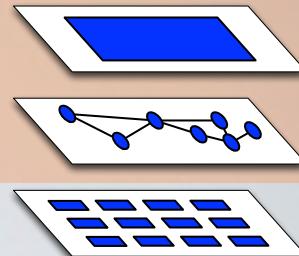
Andromeda galaxy

# what's a network?



cauliflower fractal

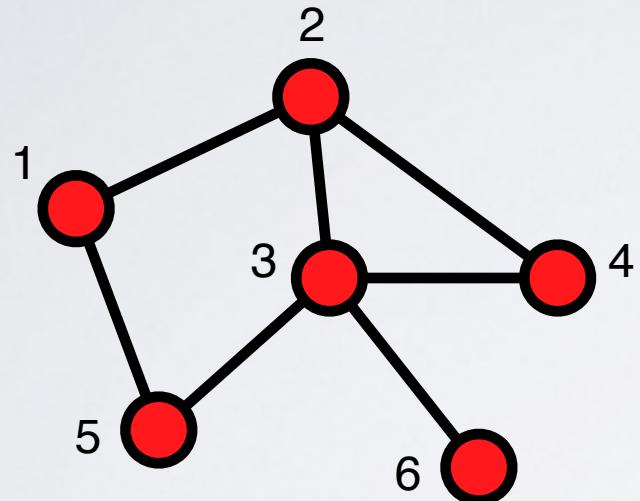
# what's a network?



diamond lattice

# **representing networks**

# *a simple network*

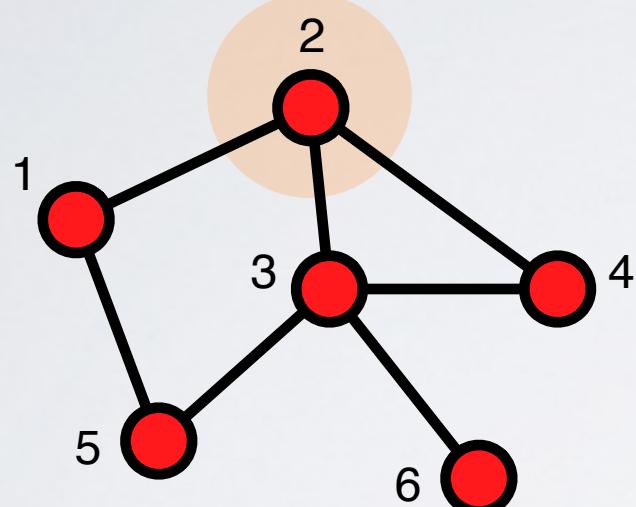


undirected

unweighted

no self-loops

# a simple network



undirected

unweighted

no self-loops

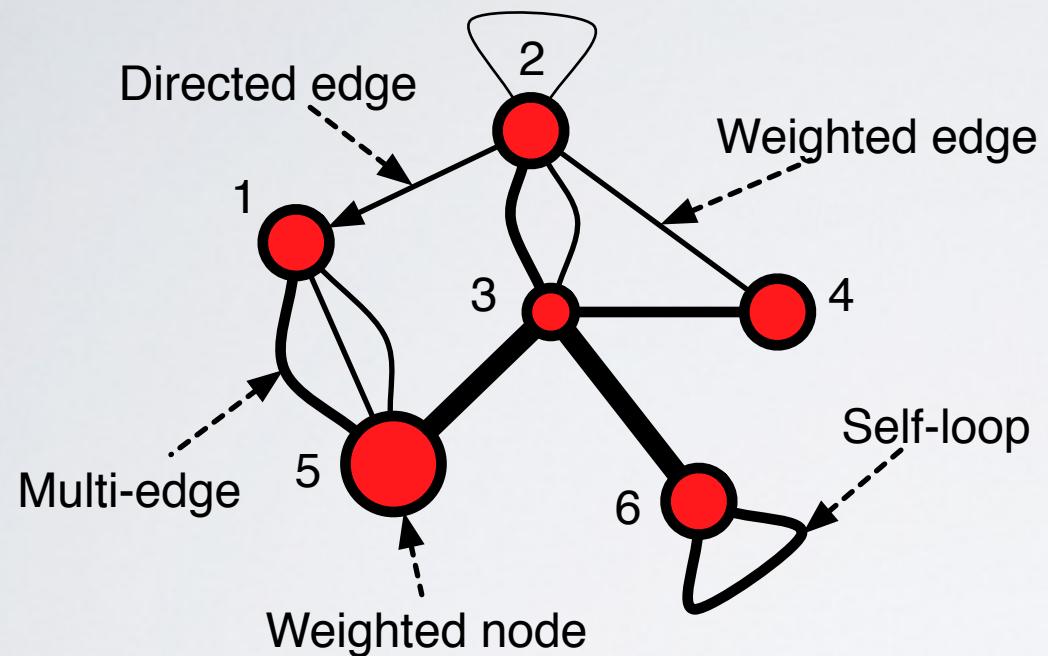
adjacency matrix

$A$	1	2	3	4	5	6
1	0	1	0	0	1	0
2	1	0	1	1	0	0
3	0	1	0	1	1	1
4	0	1	1	0	0	0
5	1	0	1	0	0	0
6	0	0	1	0	0	0

adjacency list

$A$
$1 \rightarrow \{2, 5\}$
$2 \rightarrow \{1, 3, 4\}$
$3 \rightarrow \{2, 4, 5, 6\}$
$4 \rightarrow \{2, 3\}$
$5 \rightarrow \{1, 3\}$
$6 \rightarrow \{3\}$

# a less simple network

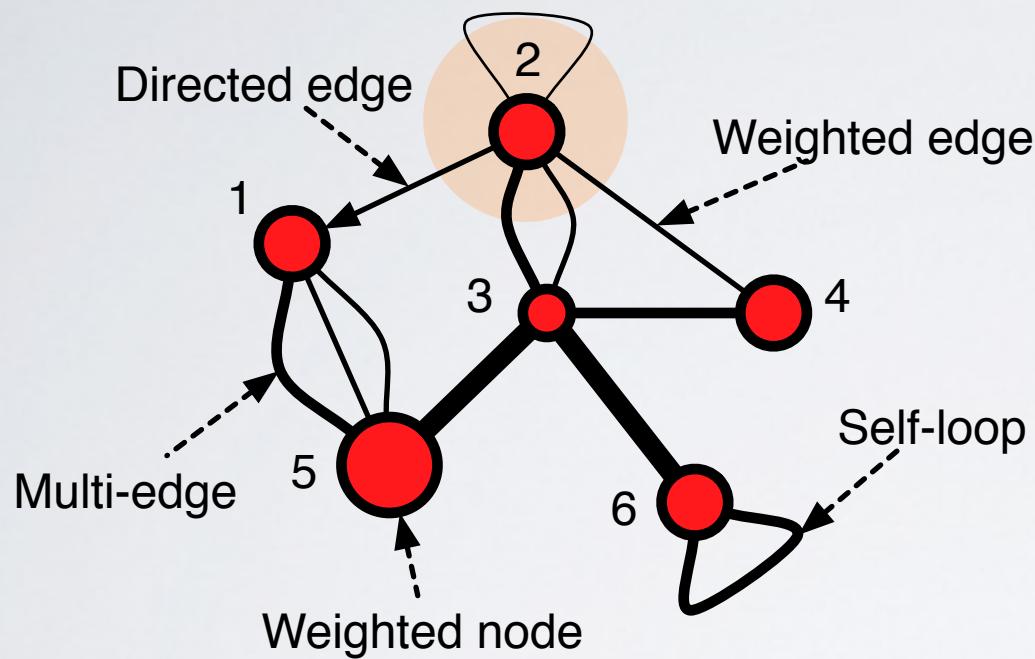


undirected

unweighted

no self loops

# a less simple network



adjacency matrix

$A$	1	2	3	4	5	6
1	0	0	0	0	{1, 1, 2}	0
2	1	$\frac{1}{2}$	{2, 1}	1	0	0
3	0	{2, 1}	0	2	4	4
4	0	1	2	0	0	0
5	{1, 1, 2}	0	4	0	0	0
6	0	0	4	0	0	2

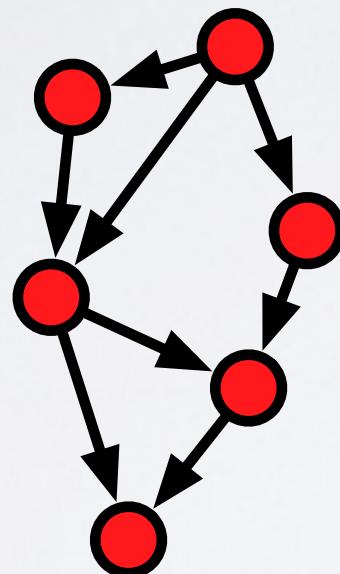
adjacency list

$A$
1 → {(5, 1), (5, 1), (5, 2)}
2 → {(1, 1), (2, $\frac{1}{2}$ ), (3, 2), (3, 1), (4, 1)}
3 → {(2, 2), (2, 1), (4, 2), (5, 4), (6, 4)}
4 → {(2, 1), (3, 2)}
5 → {(1, 1), (1, 1), (1, 2), (3, 4)}
6 → {(3, 4), (6, 2)}

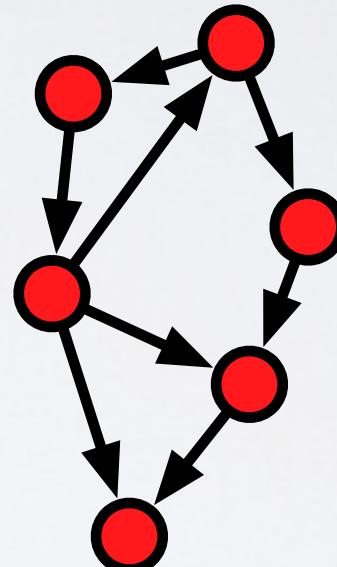
# directed networks

$$A_{ij} \neq A_{ji}$$

citation networks  
foodwebs\*  
epidemiological  
others?



directed acyclic graph

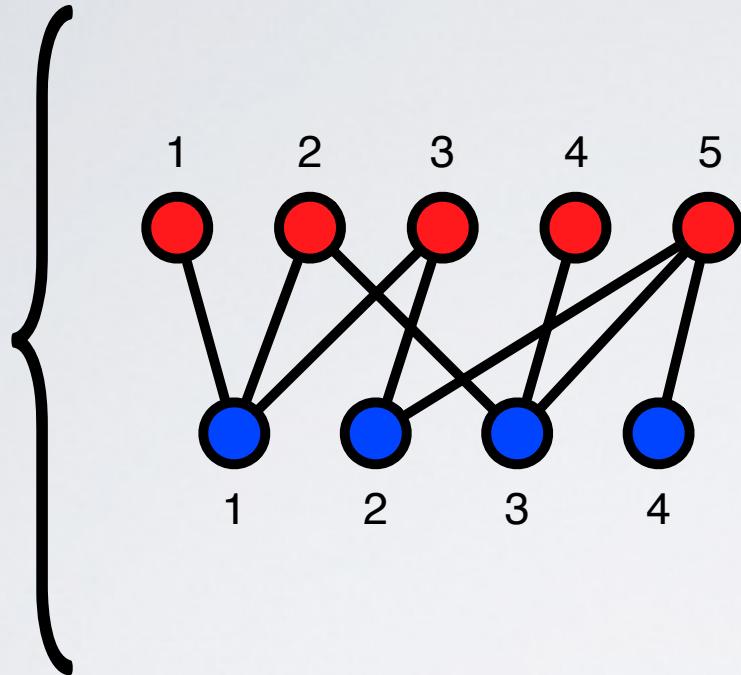


directed graph

WWW  
friendship?  
flows of goods,  
information  
economic exchange  
dominance  
neuronal  
transcription  
time travelers

# bipartite networks

bipartite  
network



**no within-type edges**

authors & papers

actors & movies/scenes

musicians & albums

people & online groups

people & corporate boards

people & locations (checkins)

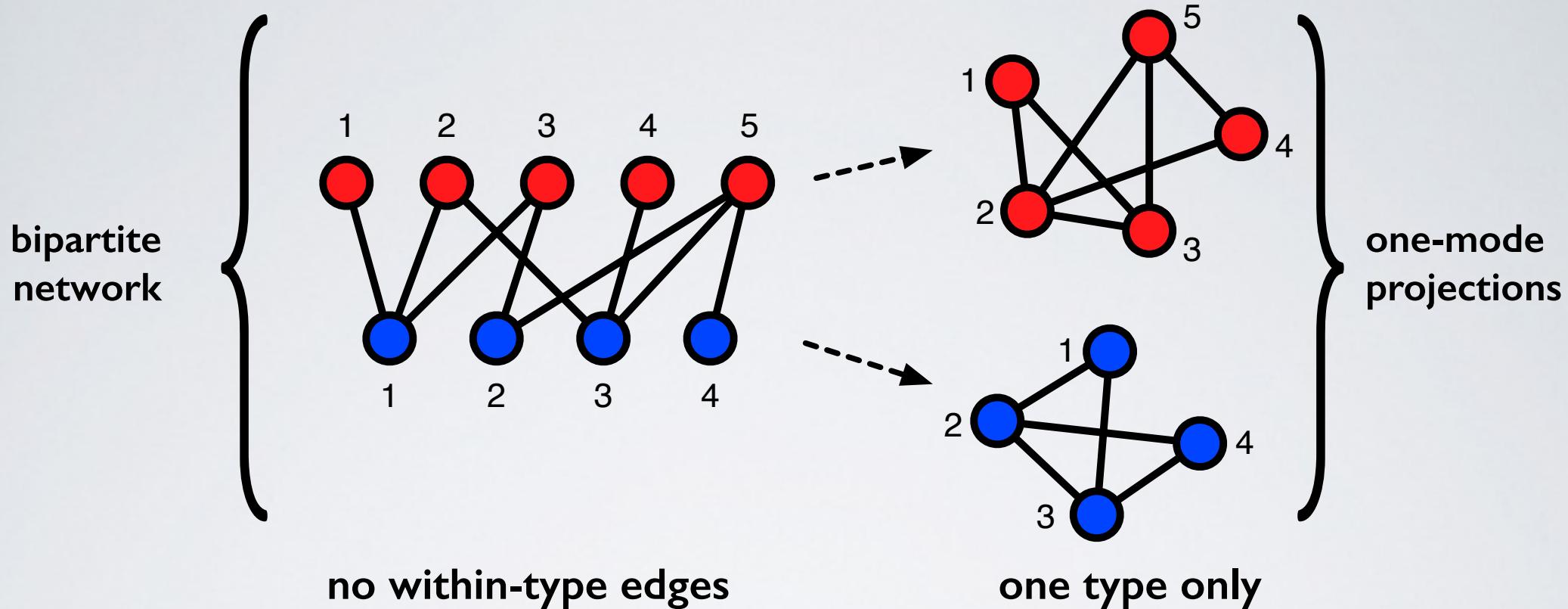
metabolites & reactions

genes & substrings

words & documents

plants & pollinators

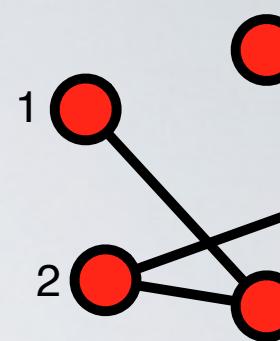
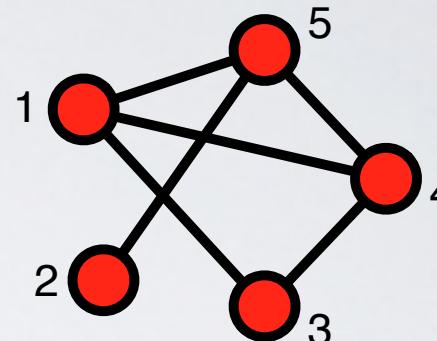
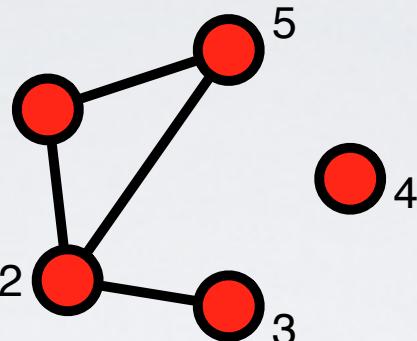
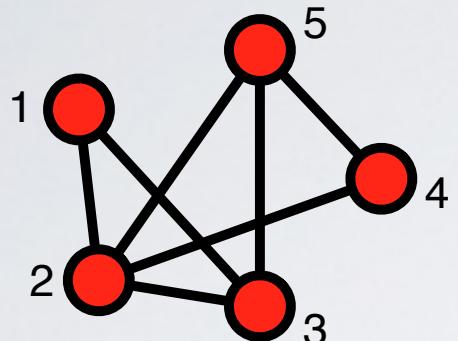
# bipartite networks



authors & papers  
actors & movies/scenes  
musicians & albums  
people & online groups  
people & corporate boards

people & locations (checkins)  
metabolites & reactions  
genes & substrings  
words & documents  
plants & pollinators

# temporal networks



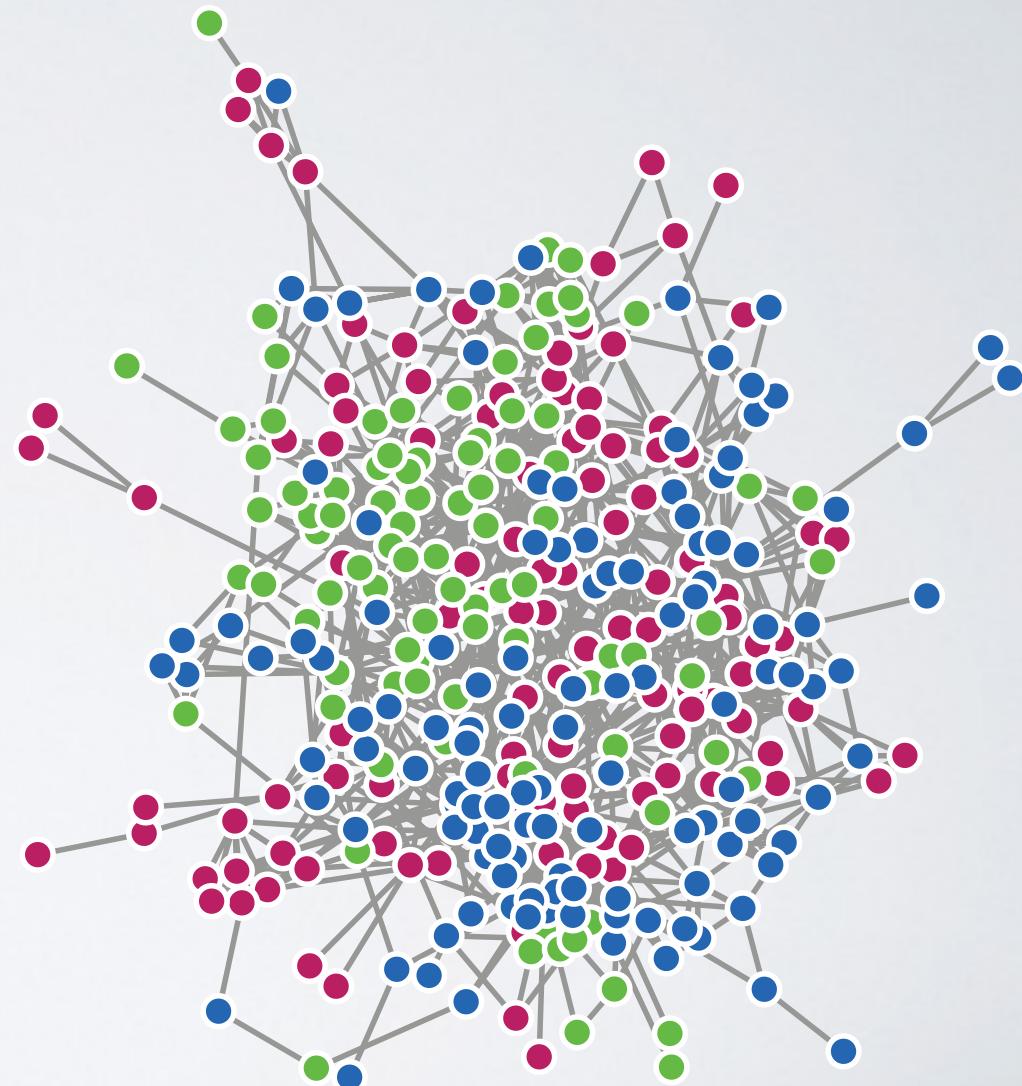
**any network over time**

discrete time (snapshots), edges  $(i, j, t)$

continuous time, edges  $(i, j, t_s, \Delta t)$

# describing networks

what networks look like



# describing networks

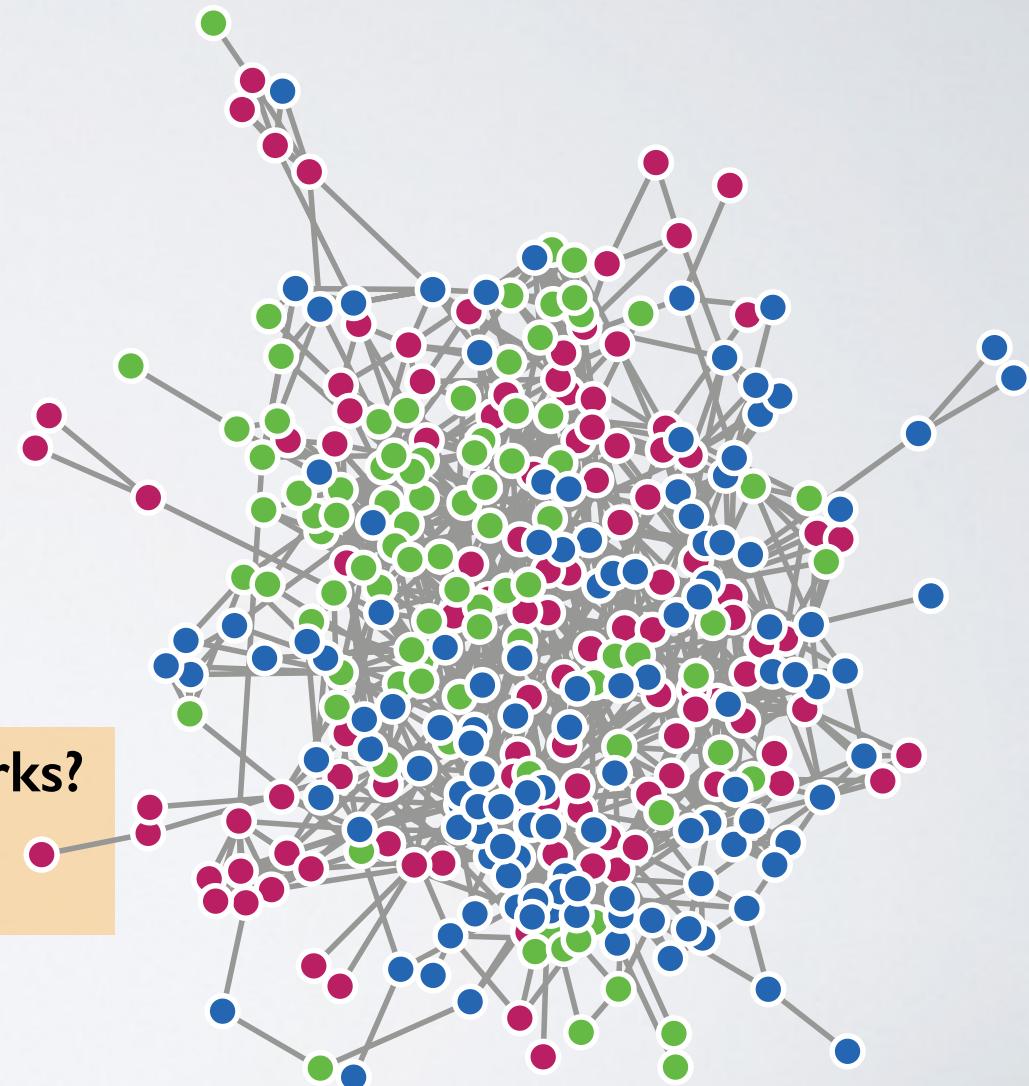
what networks look like

questions:

- **how are the edges organized?**
- **how do vertices differ?**
- **does network location matter?**
- **are there underlying patterns?**

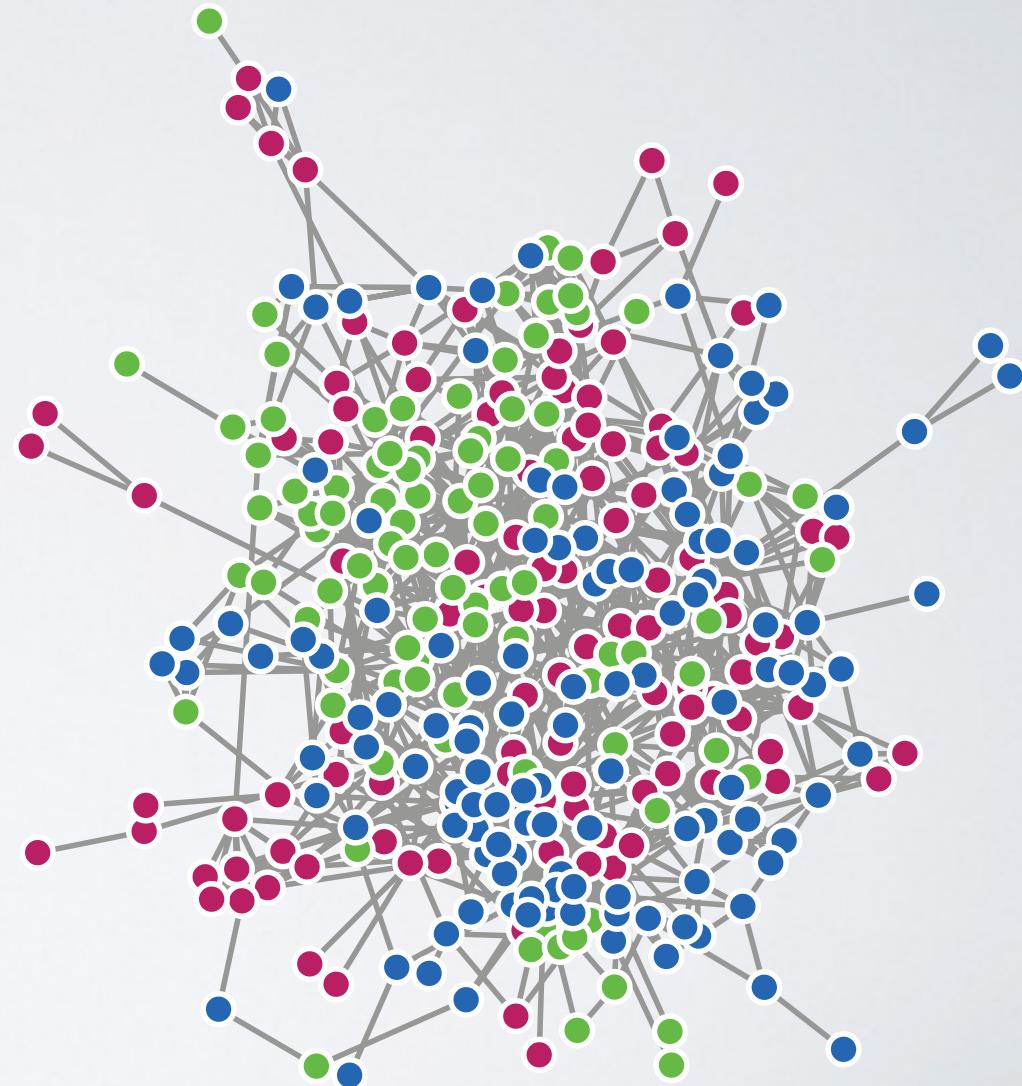
what we want to know

- **what processes shape these networks?**
- **how can we tell?**



# describing networks

a first step : **describe its features**

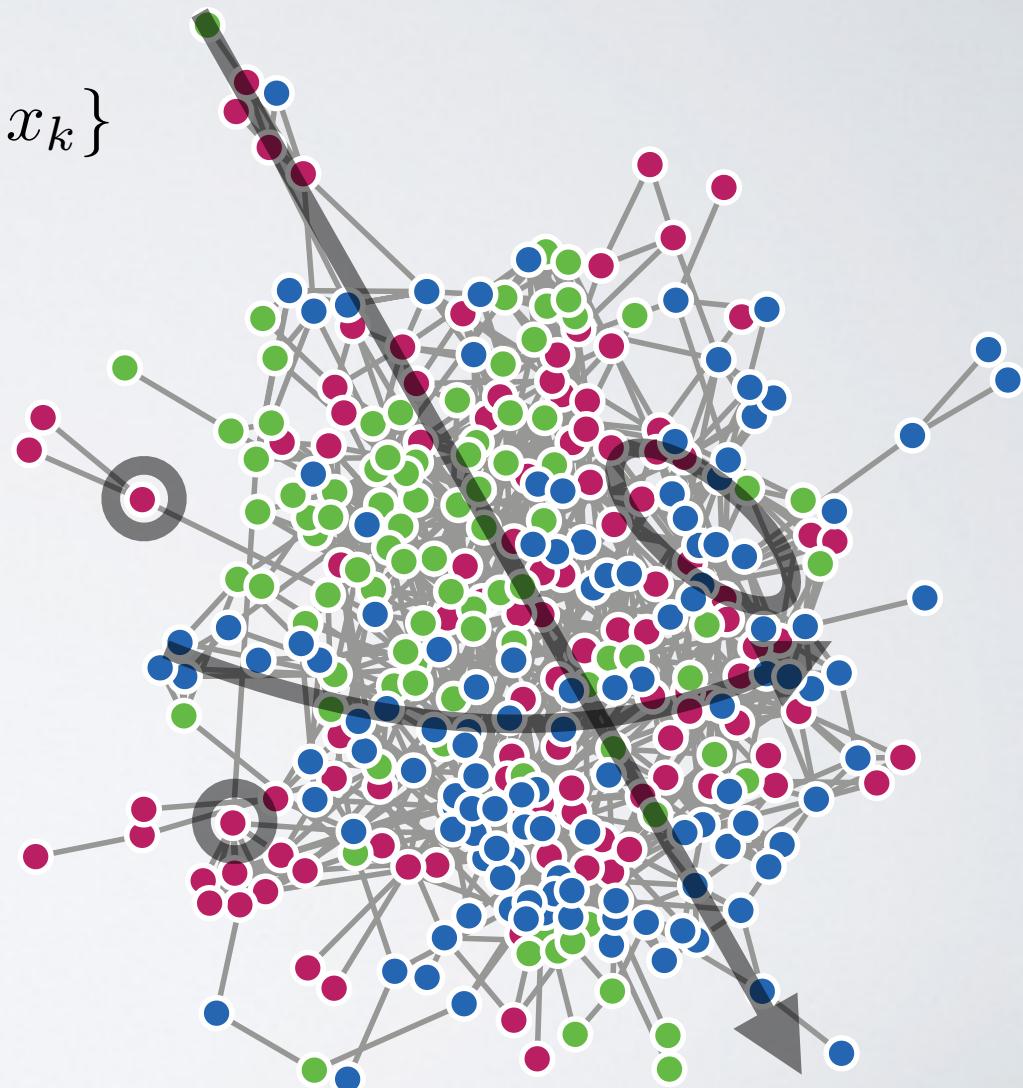


# describing networks

a first step : describe its features

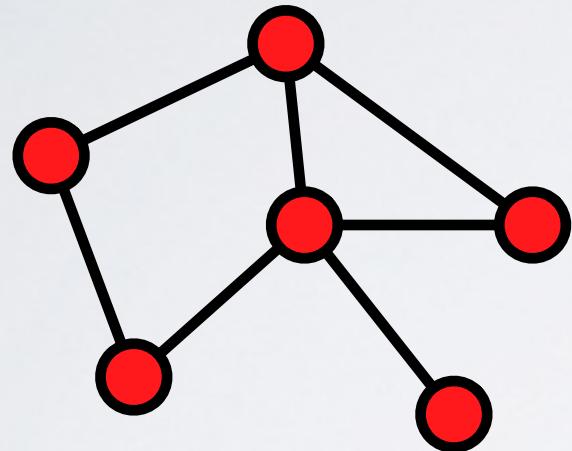
$$f : G \rightarrow \{x_1, \dots, x_k\}$$

- **degree distributions**
- **short-loop density (triangles, etc.)**
- **shortest paths (diameter, etc.)**
- **vertex positions**
- **correlations between these**

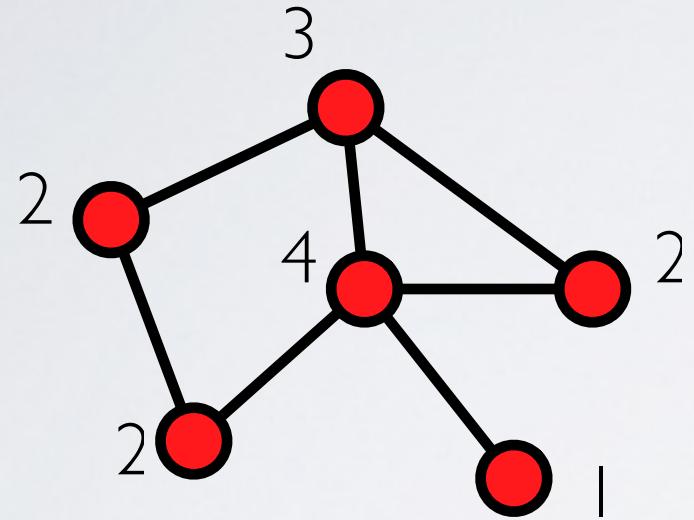


# describing networks

degree



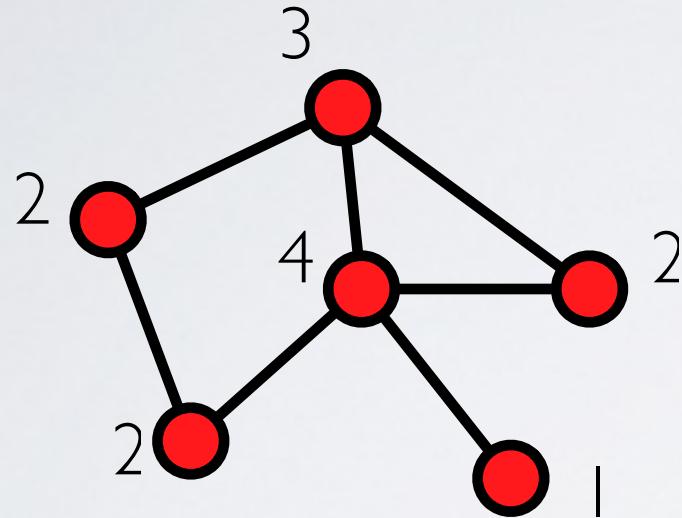
# describing networks



**degree:**  
number of connections  $k$

$$k_i = \sum_j A_{ij}$$

# describing networks



number of edges

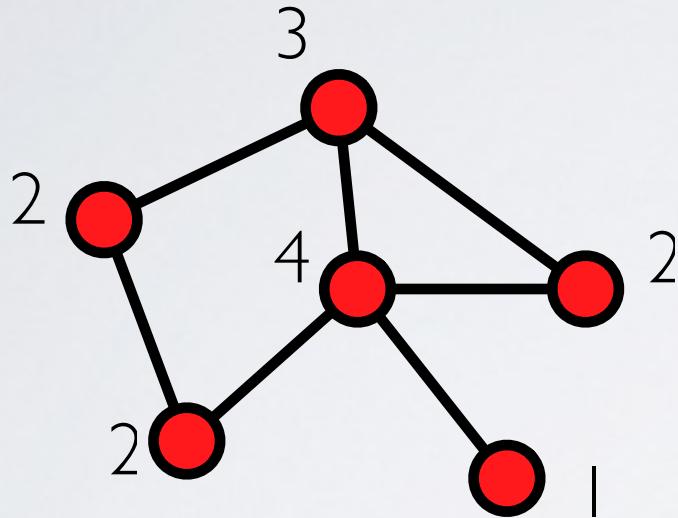
$$m = \frac{1}{2} \sum_{i=1}^n k_i = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n A_{ij} = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n A_{ji}$$

**degree:**

number of connections  $k$

$$k_i = \sum_j A_{ij}$$

# describing networks



**degree:**

number of connections  $k$

$$k_i = \sum_j A_{ij}$$

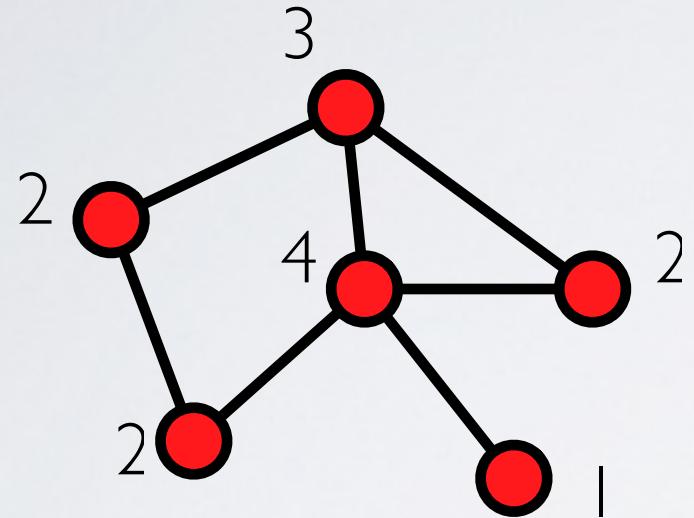
number of edges

$$m = \frac{1}{2} \sum_{i=1}^n k_i = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n A_{ij} = \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n A_{ji}$$

mean degree

$$\langle k \rangle = \frac{1}{n} \sum_{i=1}^n k_i = \frac{2m}{n}$$

# describing networks



**degree:**

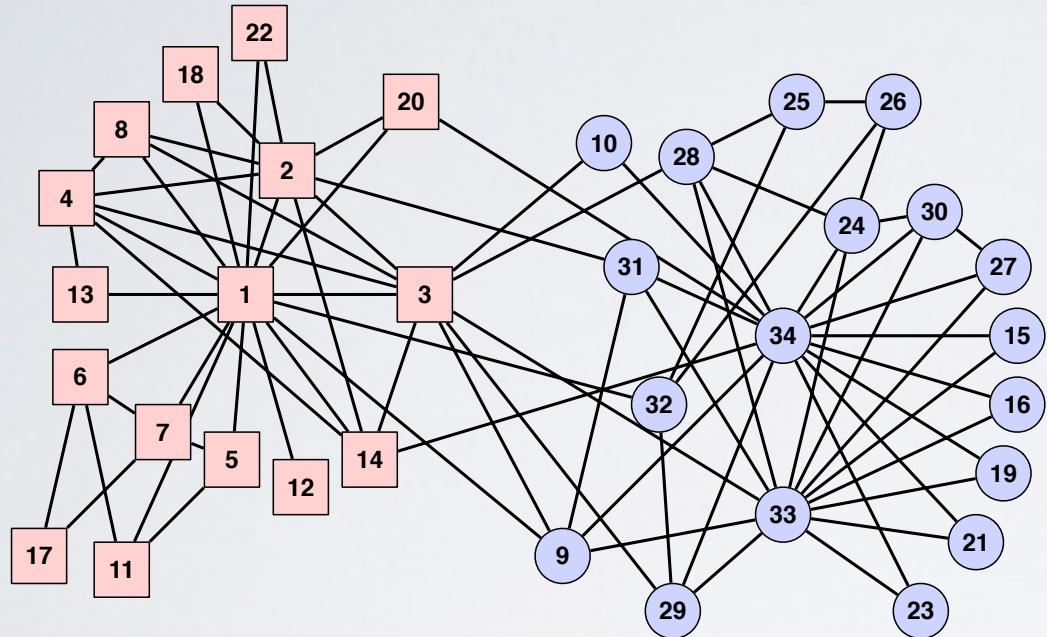
number of connections  $k$

$$k_i = \sum_j A_{ij}$$

degree sequence  $\{1, 2, 2, 2, 3, 4\}$

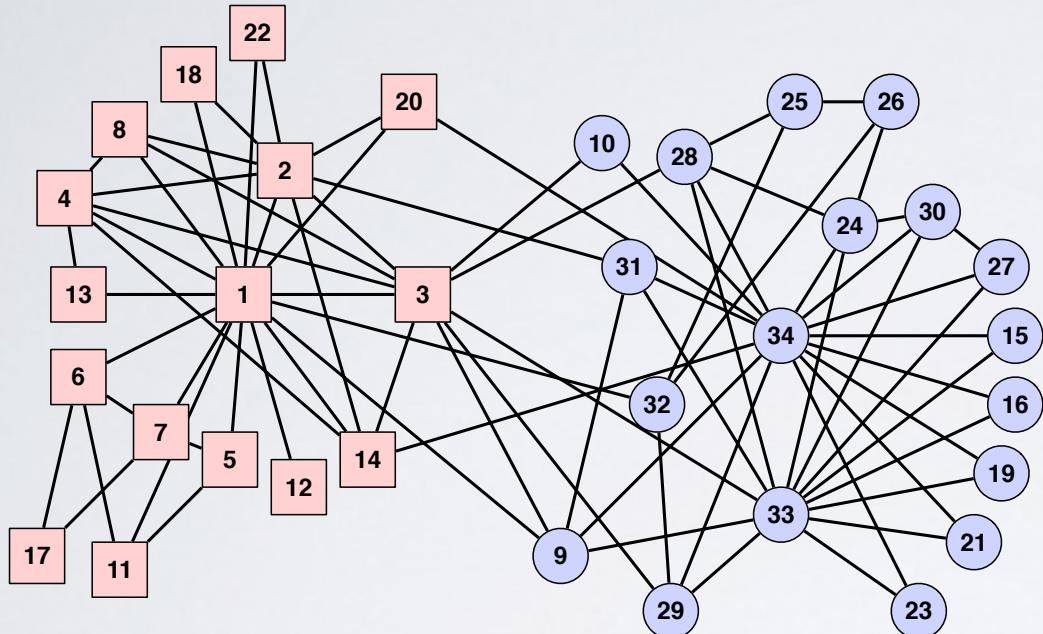
degree distribution  $\Pr(k) = \left[ \left(1, \frac{1}{6}\right), \left(2, \frac{3}{6}\right), \left(3, \frac{1}{6}\right), \left(4, \frac{1}{6}\right) \right]$

# degree distributions

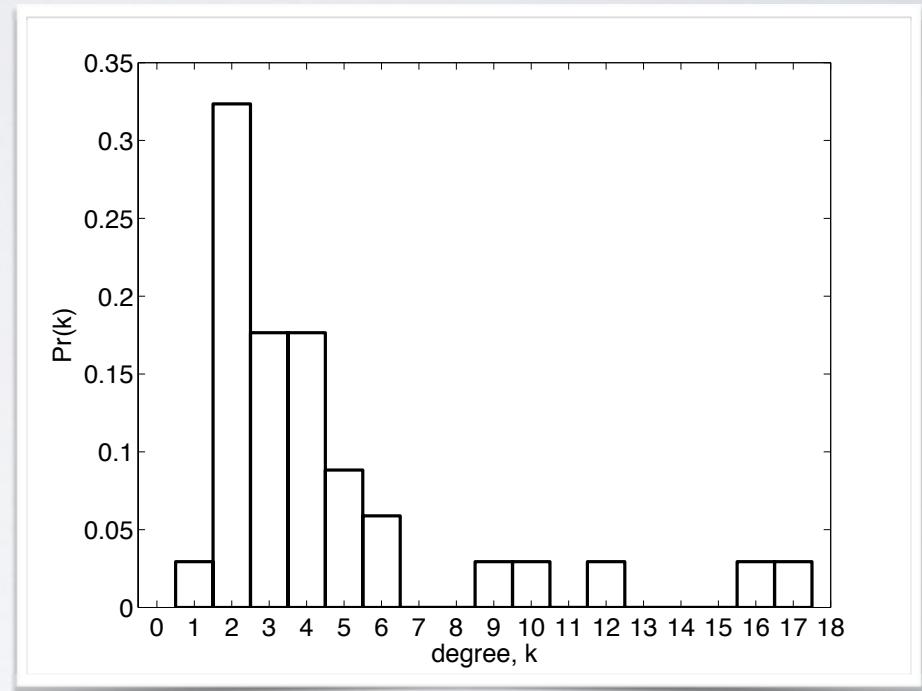


Zachary karate club\*

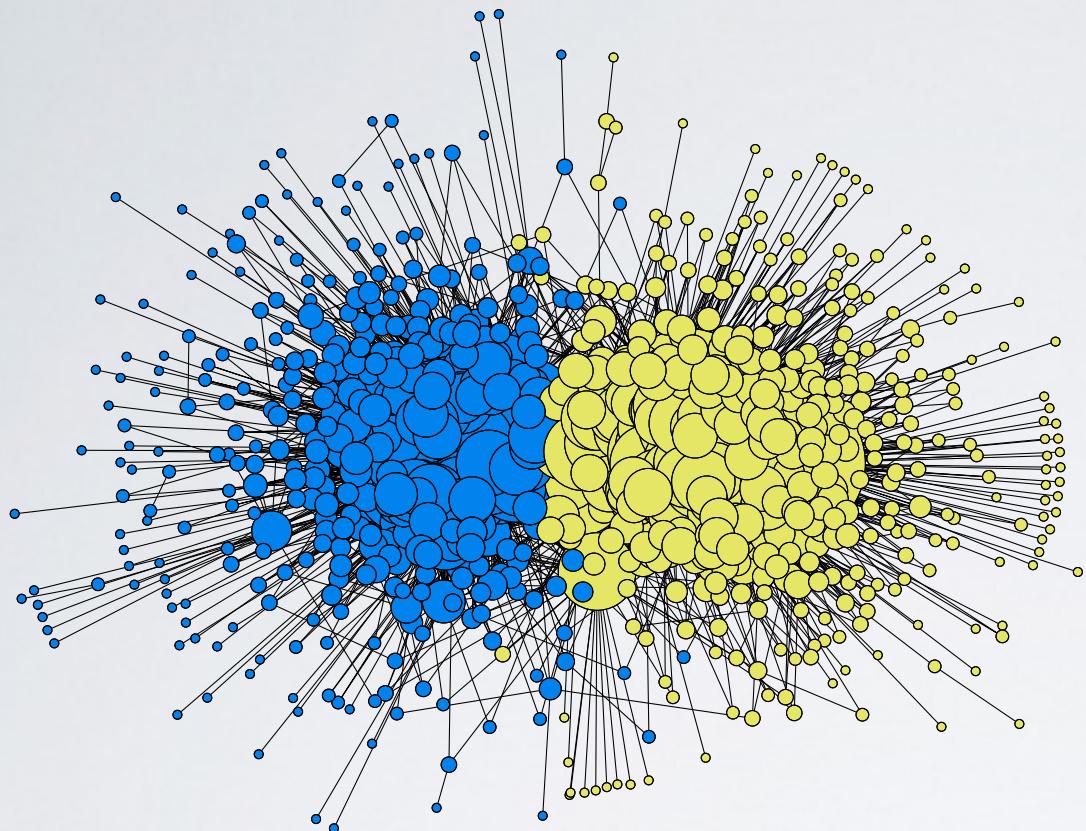
# degree distributions



Zachary karate club\*

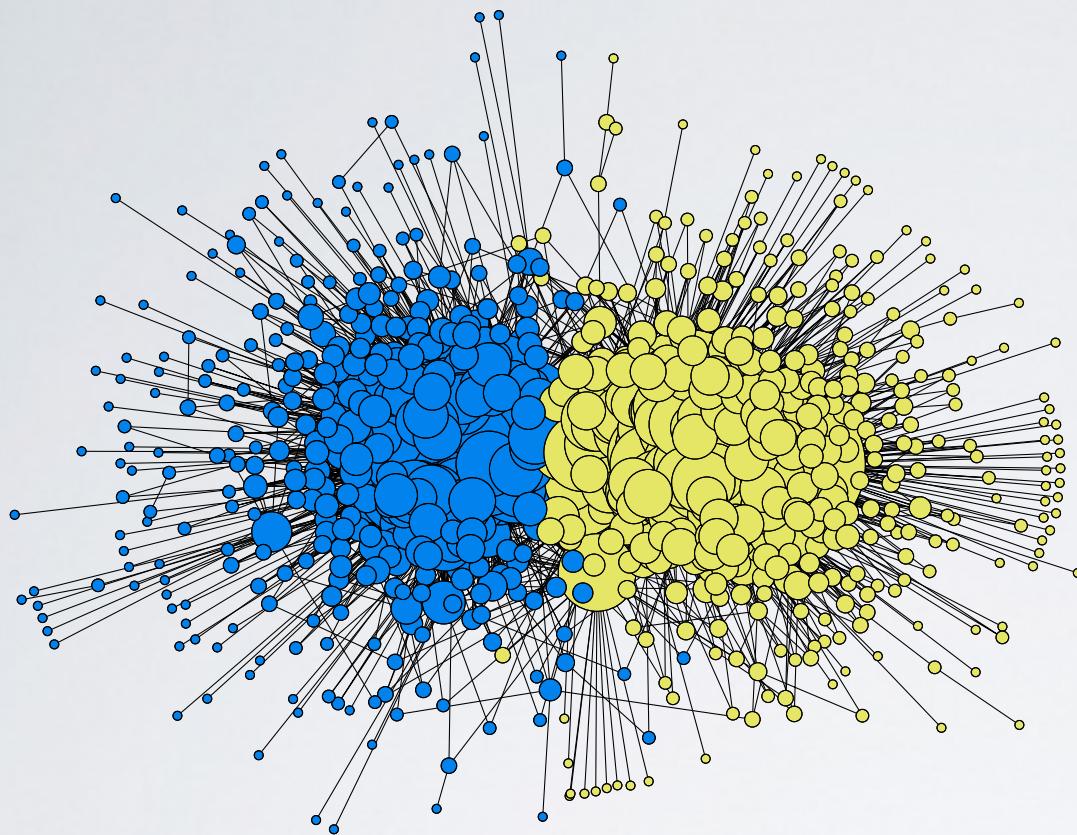


# degree distributions

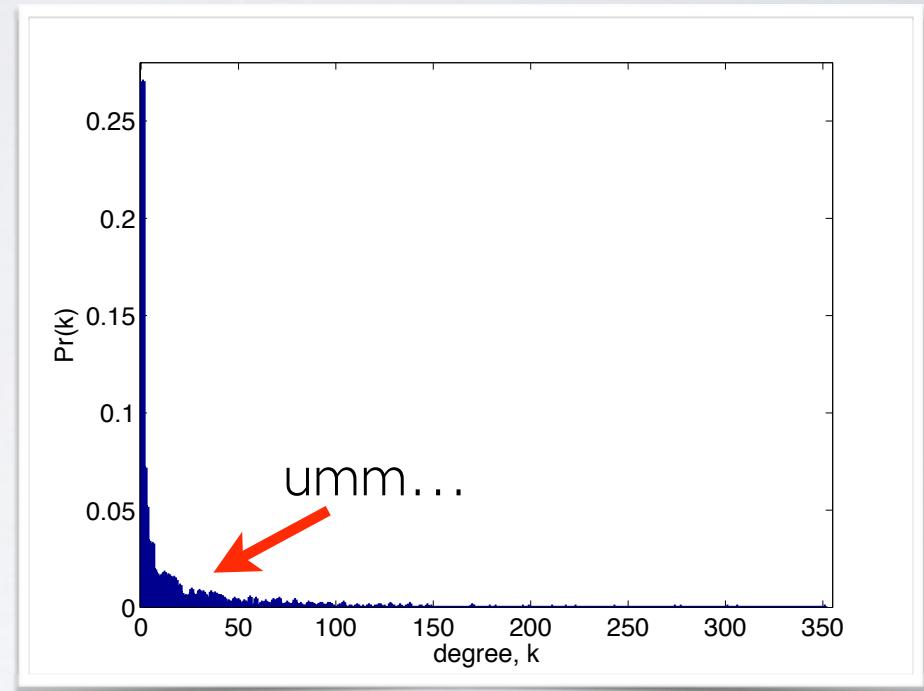


political blogs\*

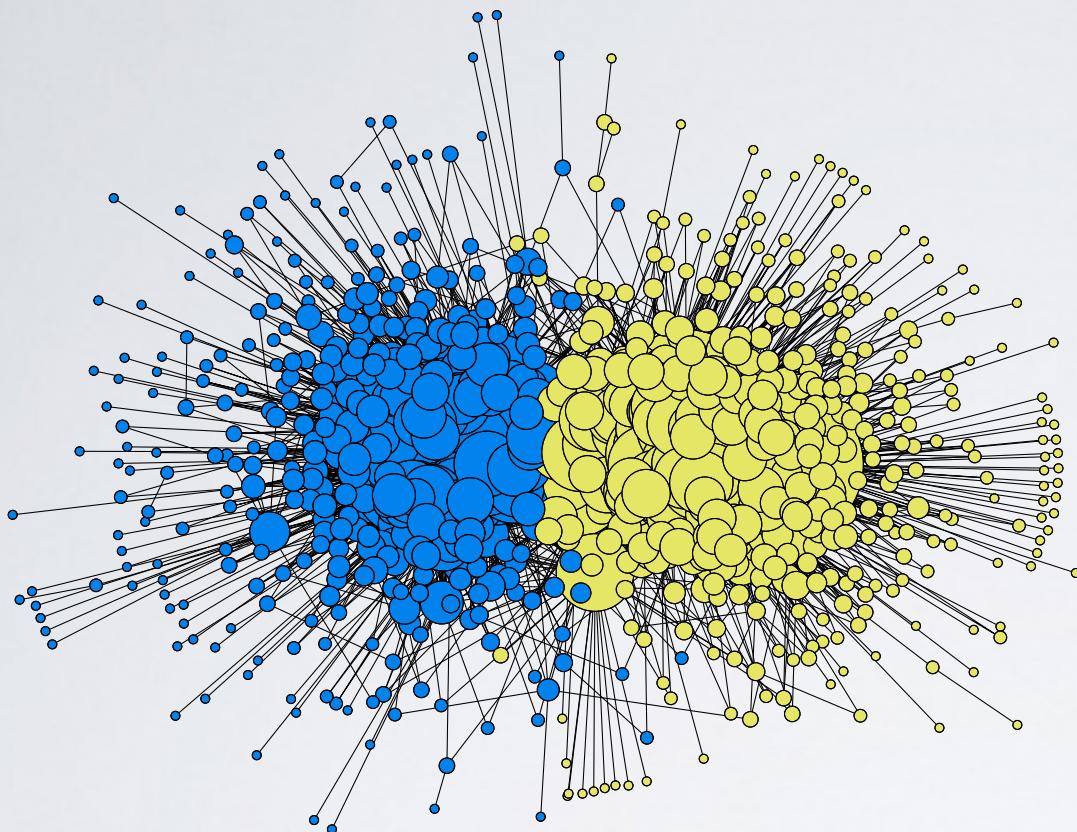
# degree distributions



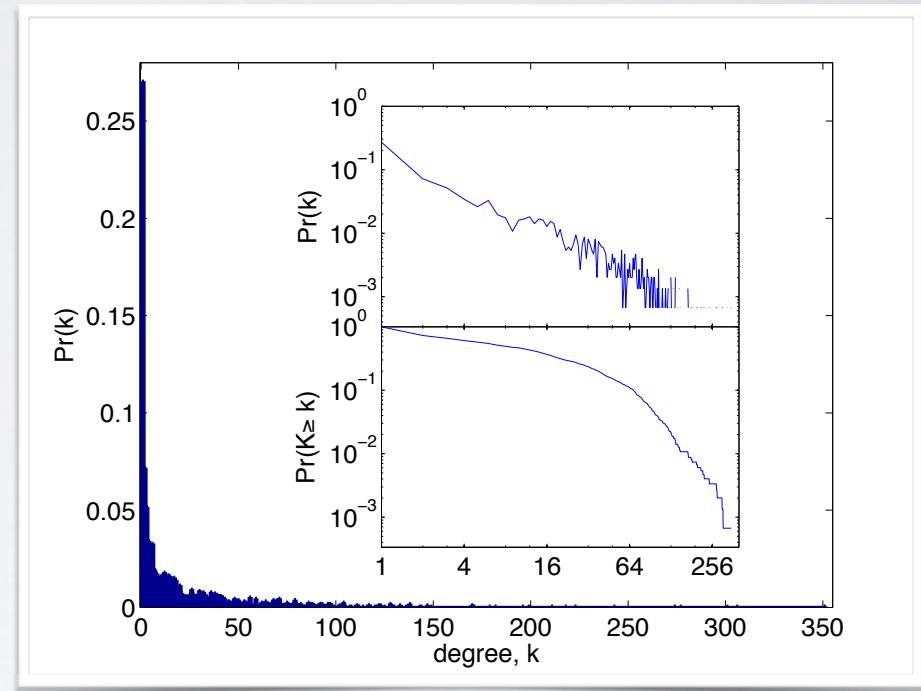
political blogs\*



# degree distributions



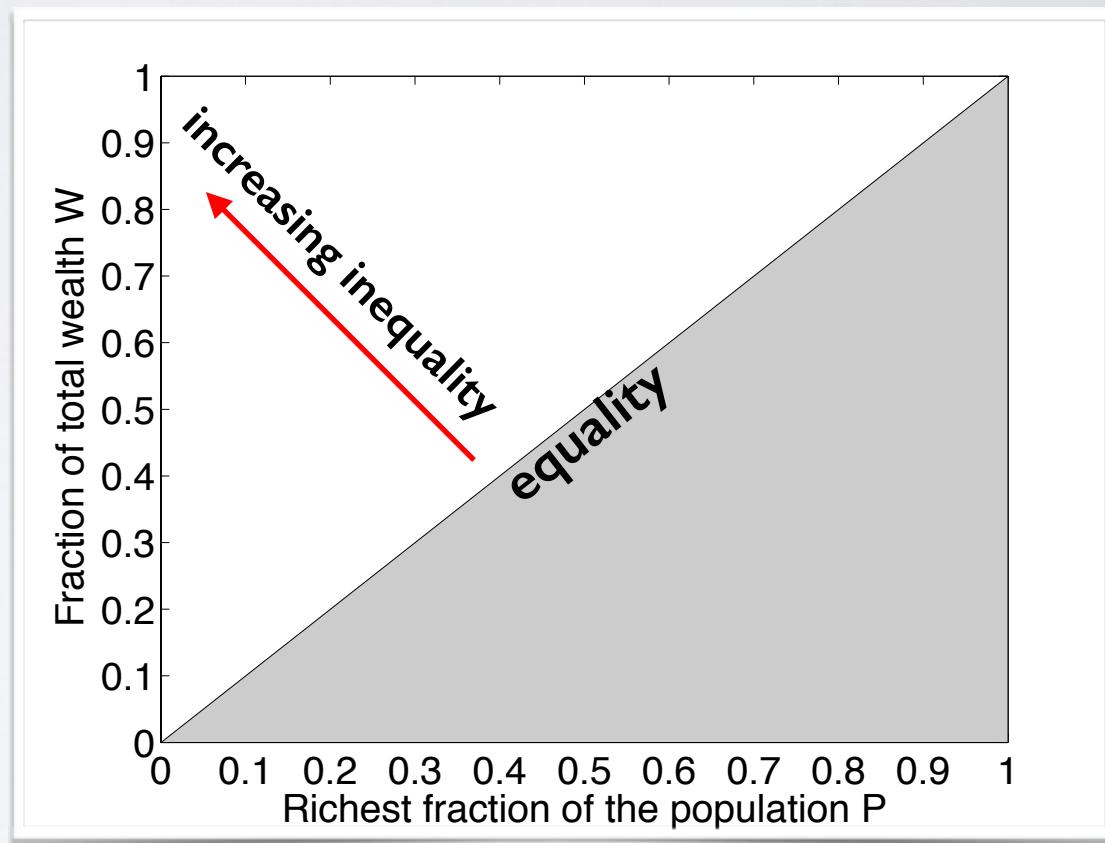
political blogs\*



# degree distributions

## degree "wealth"

what fraction of total wealth  $W$   
is owned by richest fraction  $P$



**Lorenz curve**

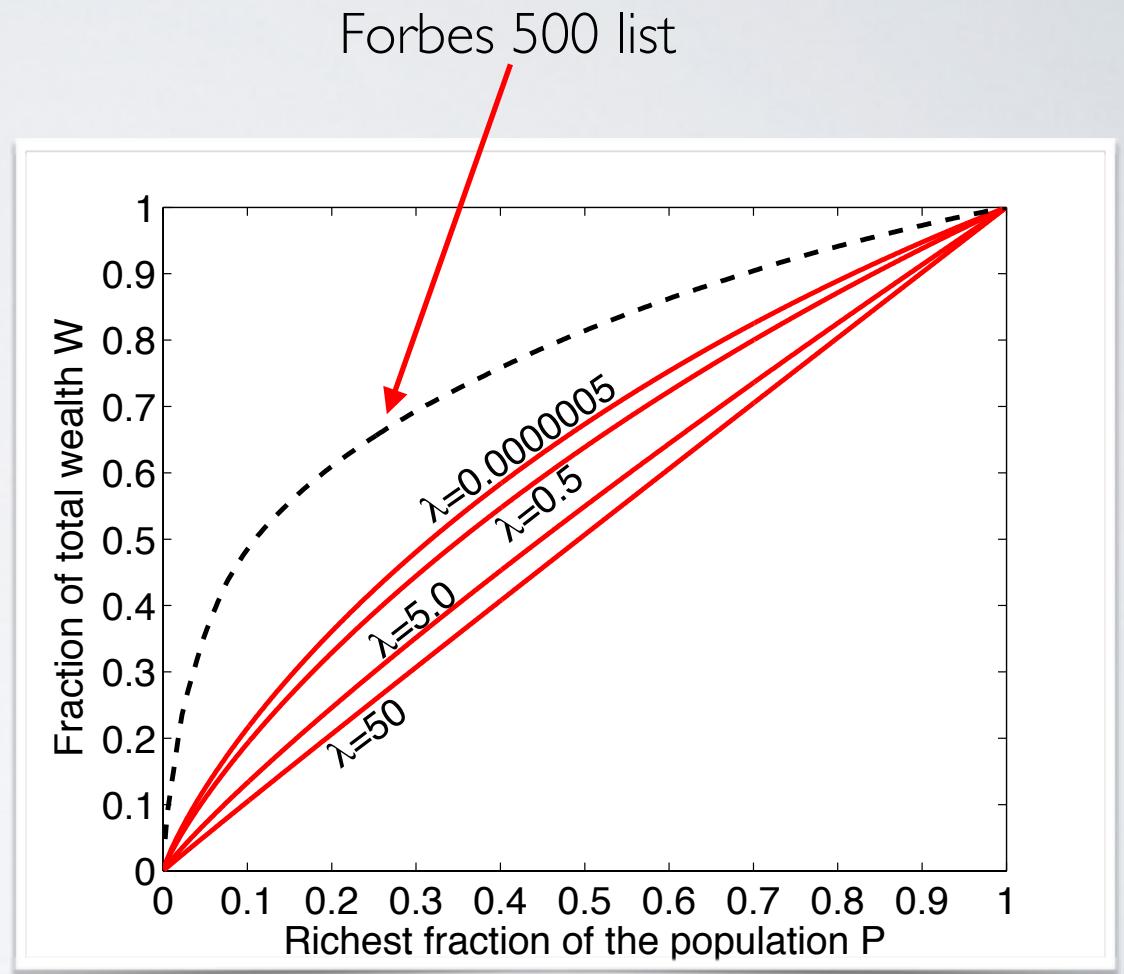
# degree distributions

## degree "wealth"

what fraction of total wealth  $W$   
is owned by richest fraction  $P$

$$\Pr(k) \propto e^{-\lambda k}$$

exponential distribution



Lorenz curve

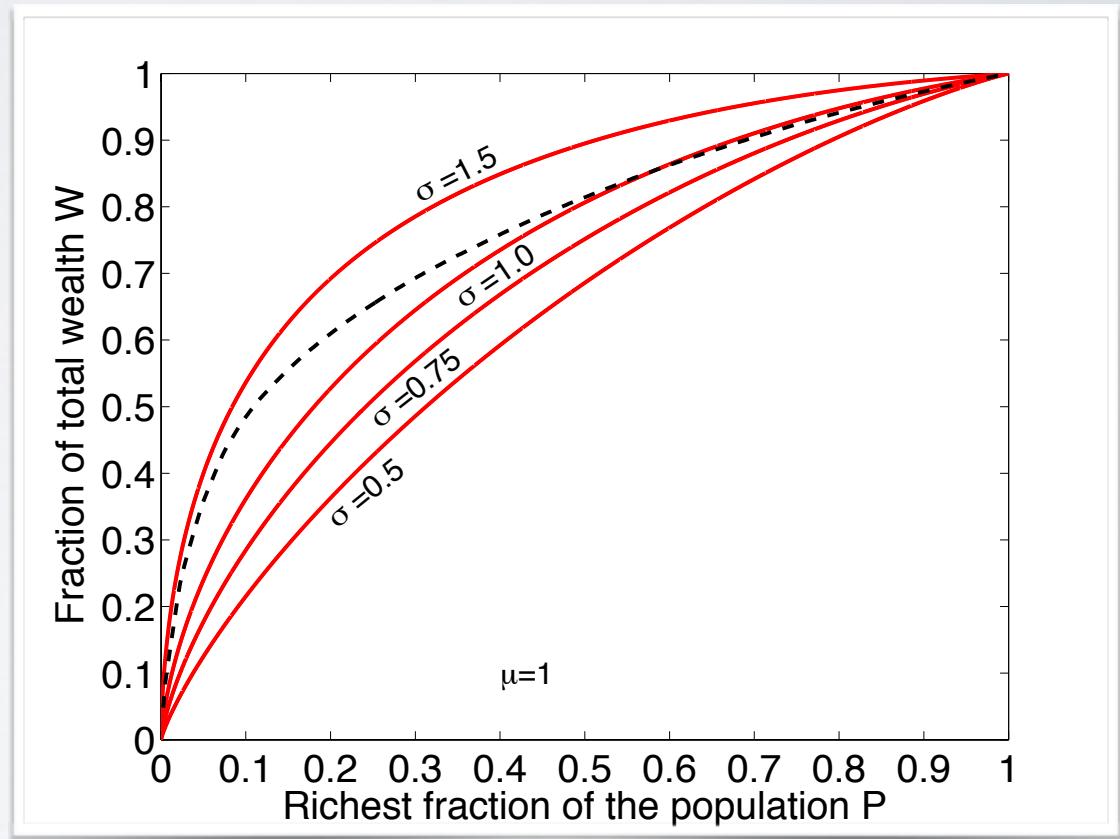
# degree distributions

## degree "wealth"

what fraction of total wealth  $W$   
is owned by richest fraction  $P$

$$\Pr(k) \propto \frac{1}{k} e^{-\left(\frac{\ln k - \mu}{\sigma \sqrt{2}}\right)^2}$$

log-normal distribution



Lorenz curve

# degree distributions

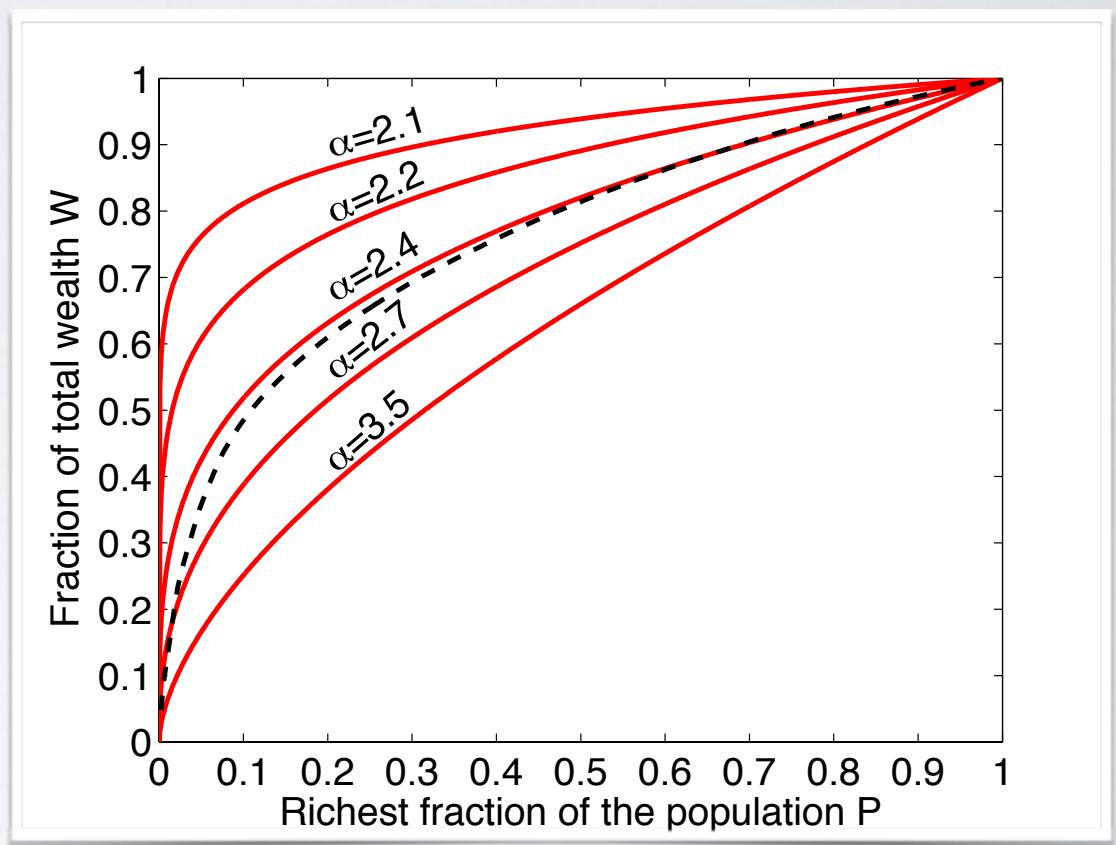
## degree "wealth"

what fraction of total wealth  $W$   
is owned by richest fraction  $P$

$$\Pr(k) \propto k^{-\alpha}$$

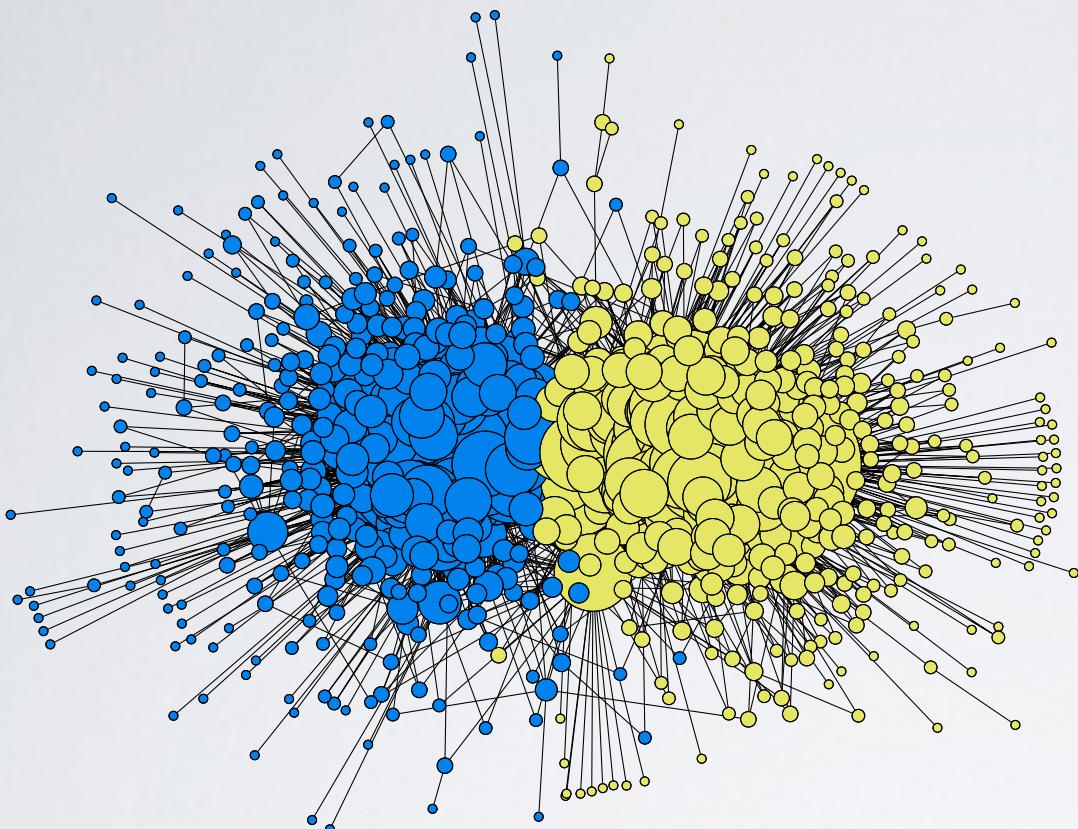
power-law distribution

**80/20 rule**

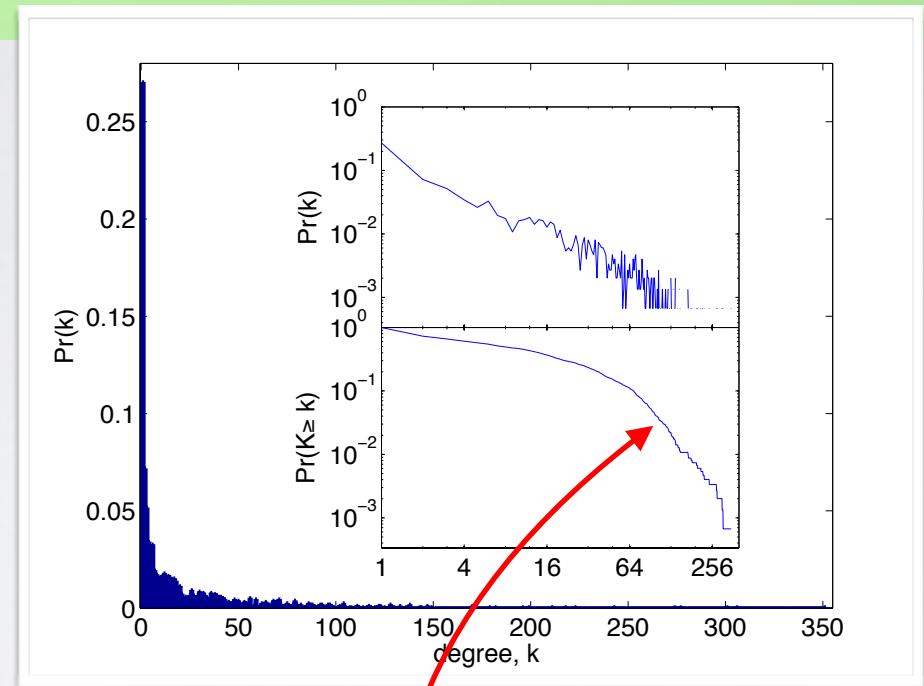


**Lorenz curve**

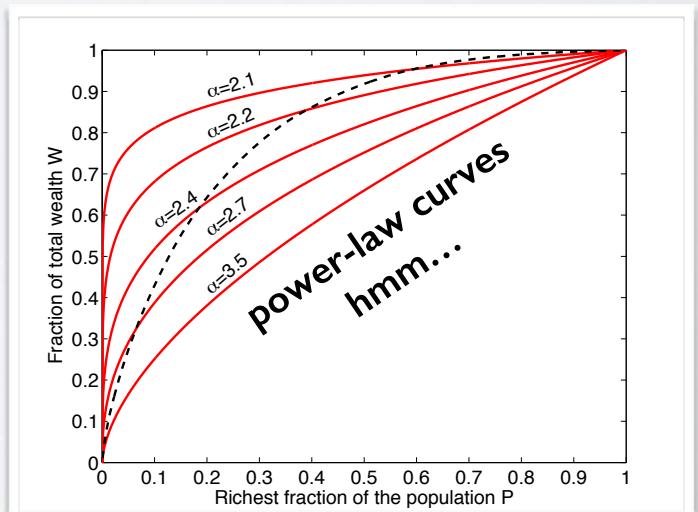
# degree distributions



political blogs\*



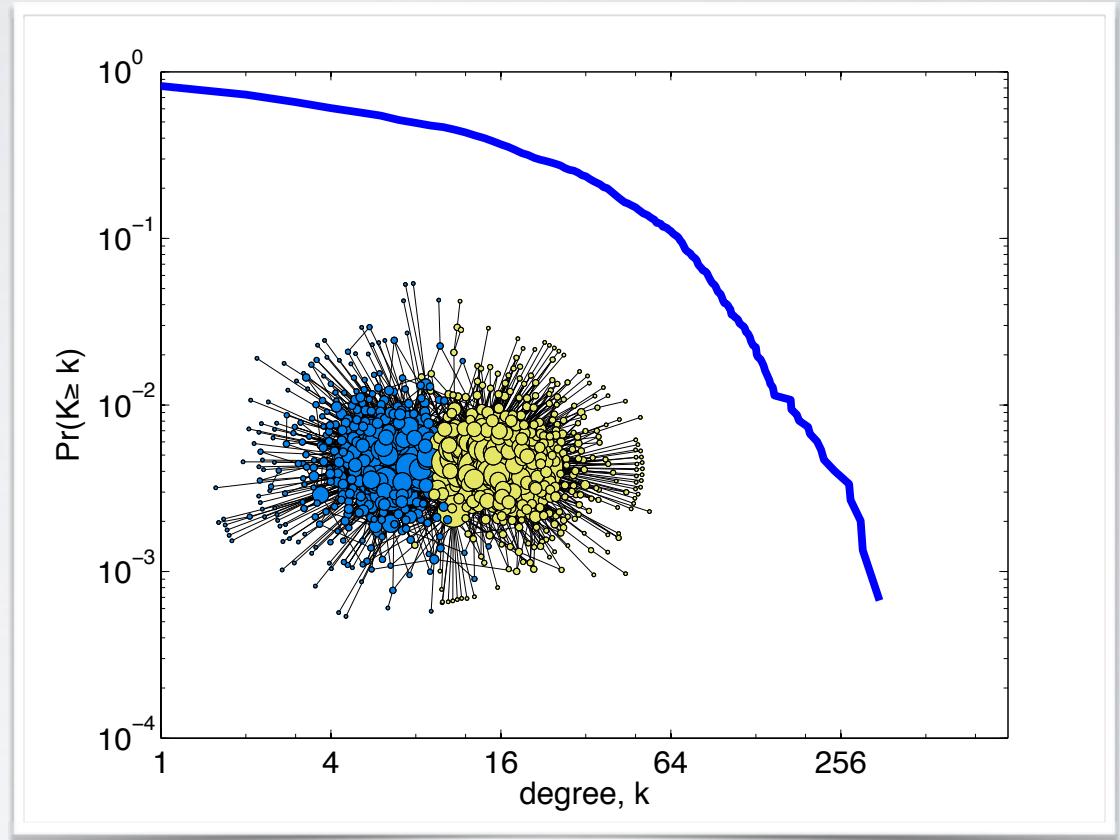
is this a power law?



# degree distributions

## fun facts:

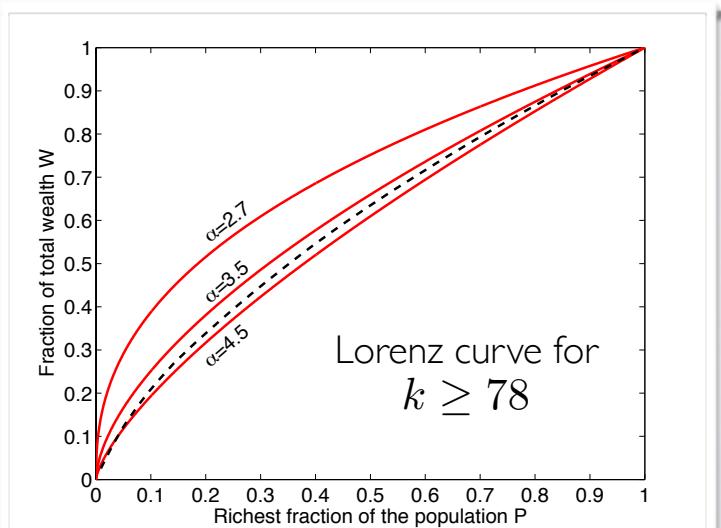
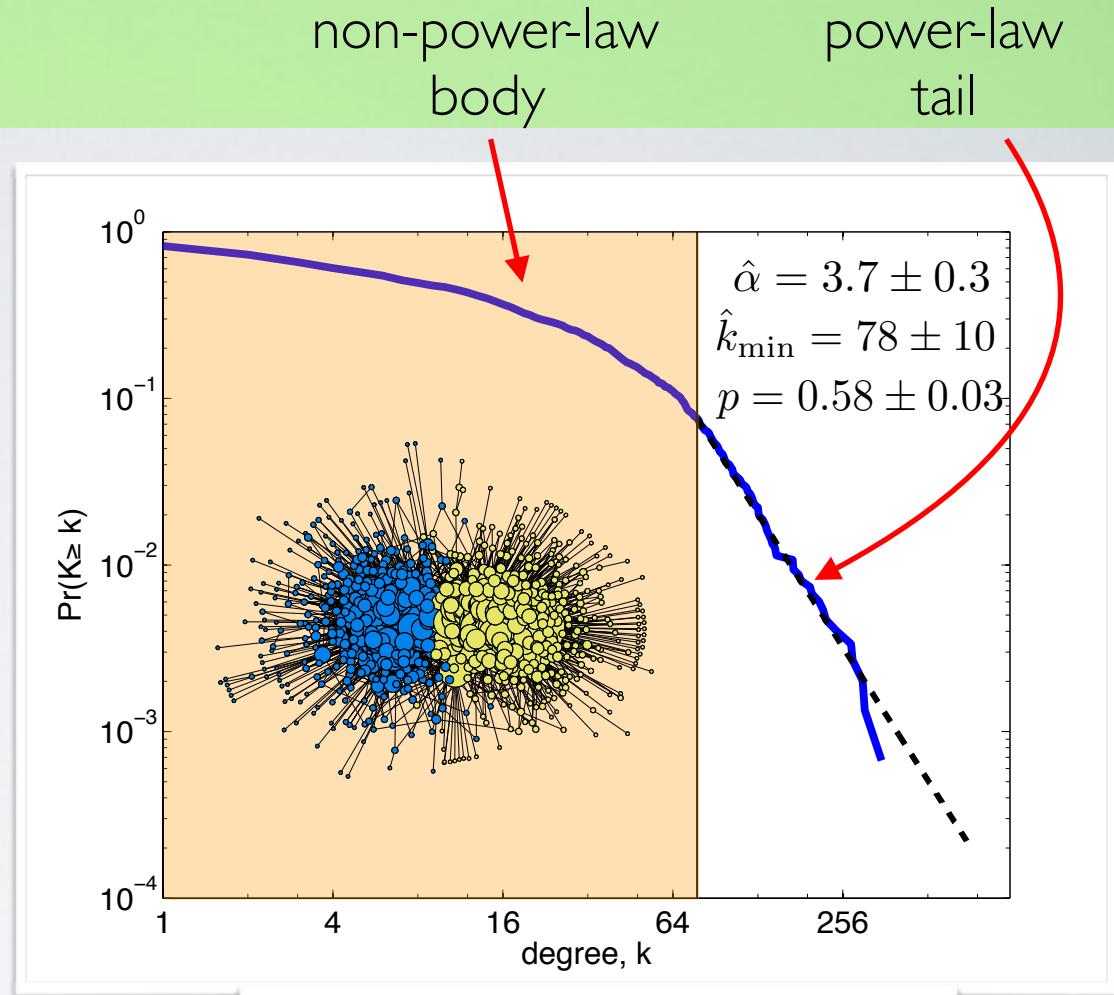
- nearly all real networks exhibit a ***heavy-tailed degree distribution***
- **very few** networks exhibit perfect power-law degree distributions
- **some** distributions exhibit power-law tails
- power laws are cool!  
but knowing one from garbage  
***requires statistics\****



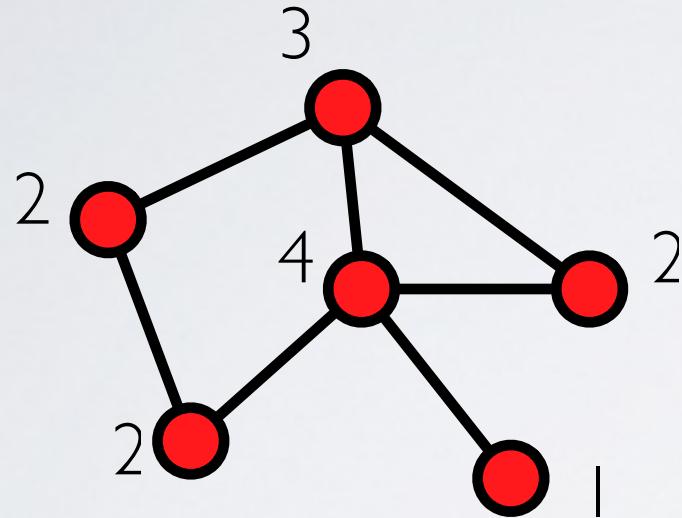
# degree distributions

## fun facts:

- nearly all real networks exhibit a ***heavy-tailed degree distribution***
- ***very few*** networks exhibit perfect power-law degree distributions
- ***some*** distributions exhibit power-law tails
- power laws are cool!  
but knowing one from garbage  
***requires statistics\****



# describing networks



**degree:**

number of connections  $k$

$$k_i = \sum_j A_{ij}$$

**when does node  
degree matter?**

# network degrees

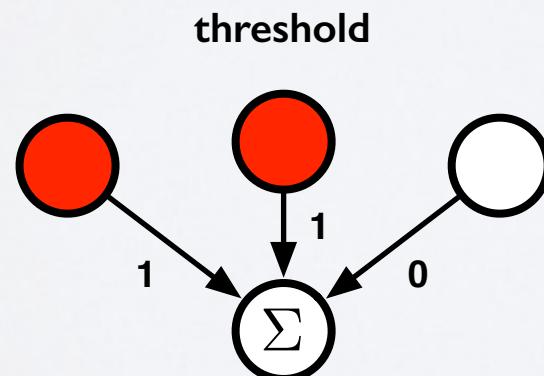
## spreading processes on networks

biological (diseases)

- SIS and SIR models

social (information)

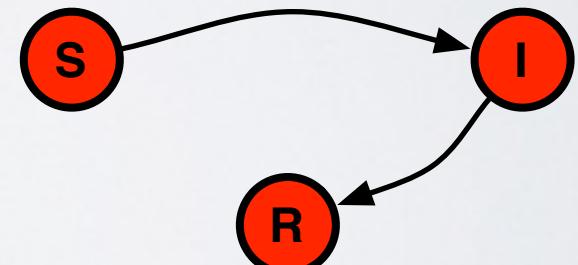
- SIS, SIR models
- threshold models



susceptible-infected-susceptible



susceptible-infected-recovered



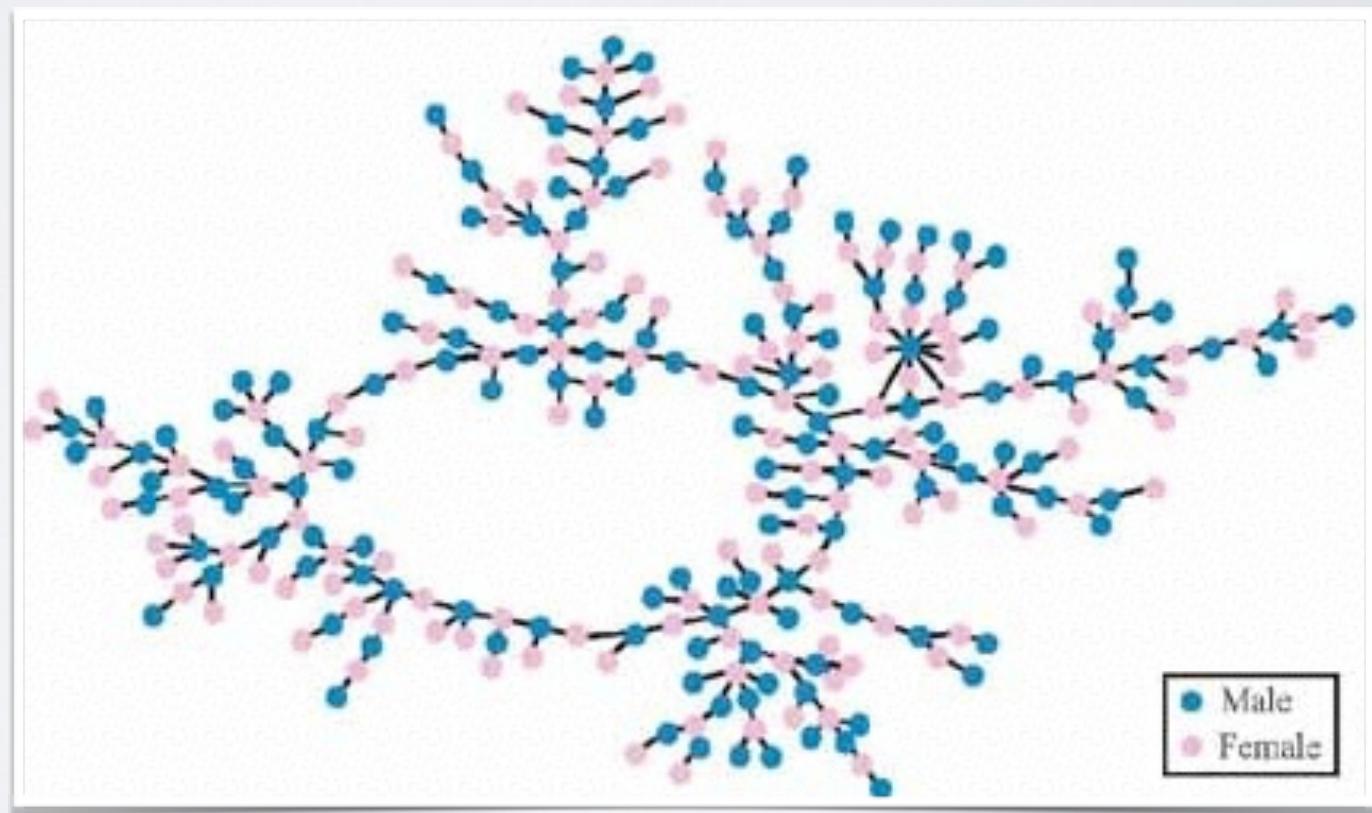
# network degrees

## Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks

2004

Peter S. Bearman      James Moody      Katherine Stovel  
*Columbia University*   *Ohio State University*   *University of Washington*

- relationship network in “Jefferson High”
- this subgraph is 52% of school
- who are most important disease spreaders?



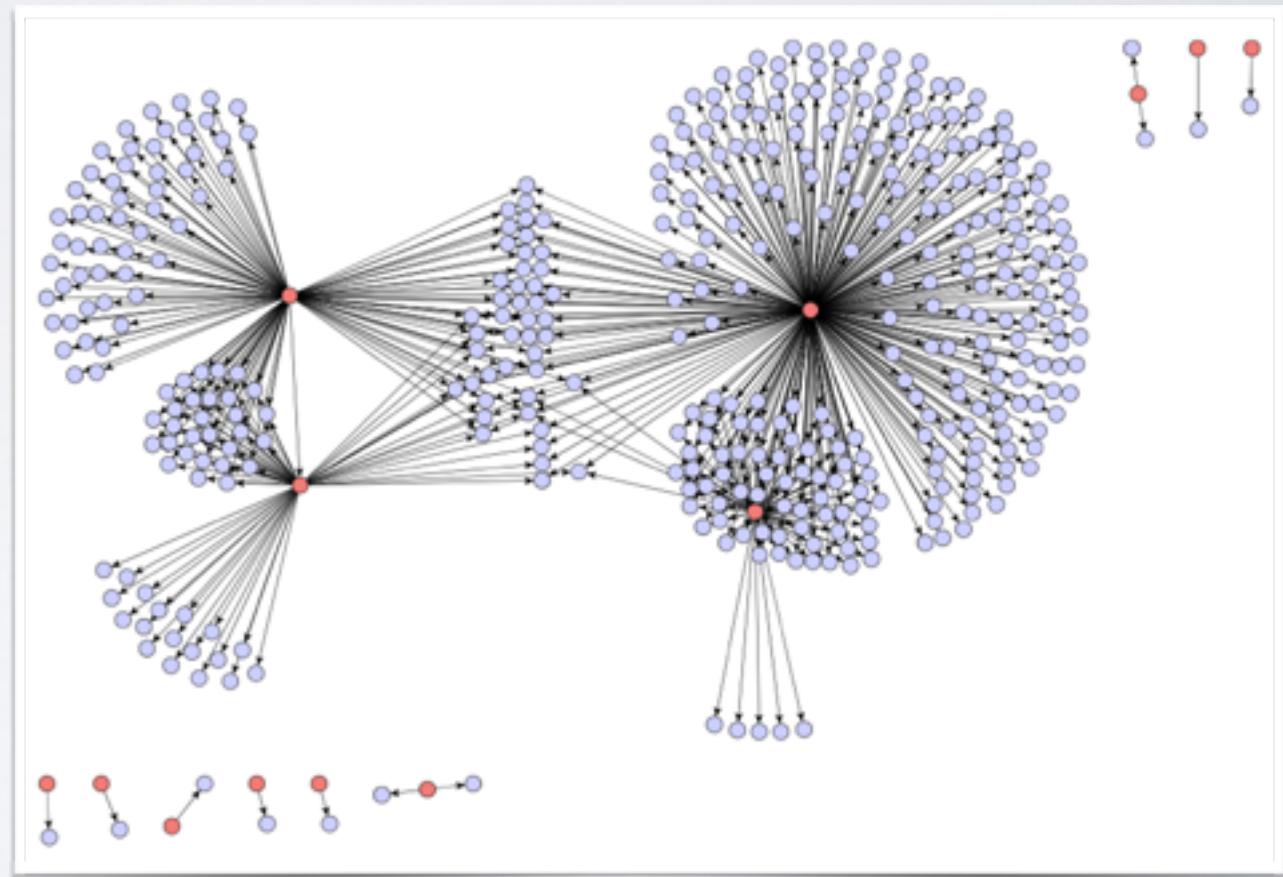
# network degrees

## The Dynamics of Viral Marketing

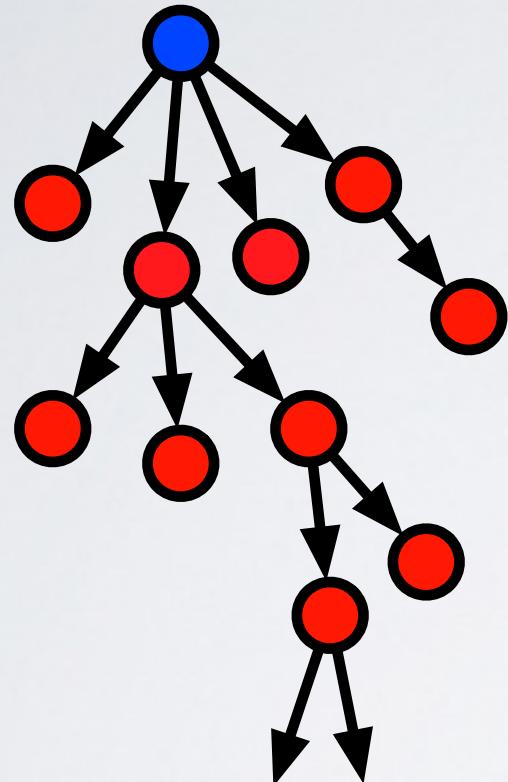
2007

JURE LESKOVEC LADA A. ADAMIC BERNARDO A. HUBERMAN

- amazon.com viral marketing
- viral trace for “Oh my Goddess!” community
- very high degrees!
- most attempts to “influence” fail



# network degrees



$$R_0 = 0.923 \dots$$

cascade  
epidemic  
branching process  
spreading process

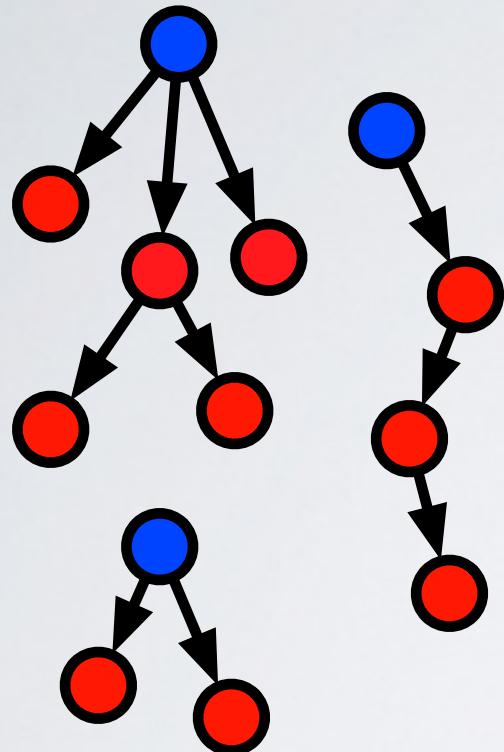
$$R_0 = \frac{\text{net reproductive rate}}{\text{average degree } \langle k \rangle}$$

If above 1 it will explode, below 1 it will implode

**caveat:**

ignores network structure,  
dynamics, etc.

# network degrees

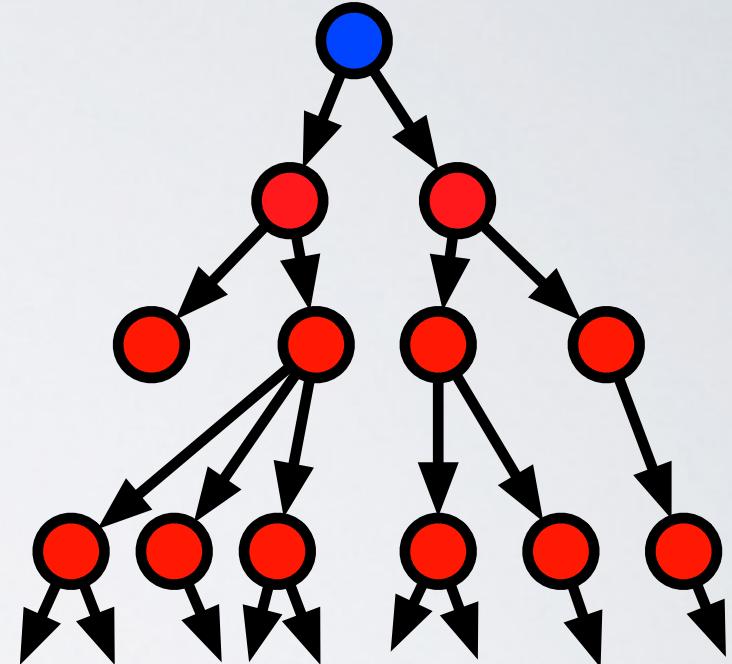


$$R_0 < 1$$

“sub-critical”  
small outbreaks

$$R_0 = 1$$

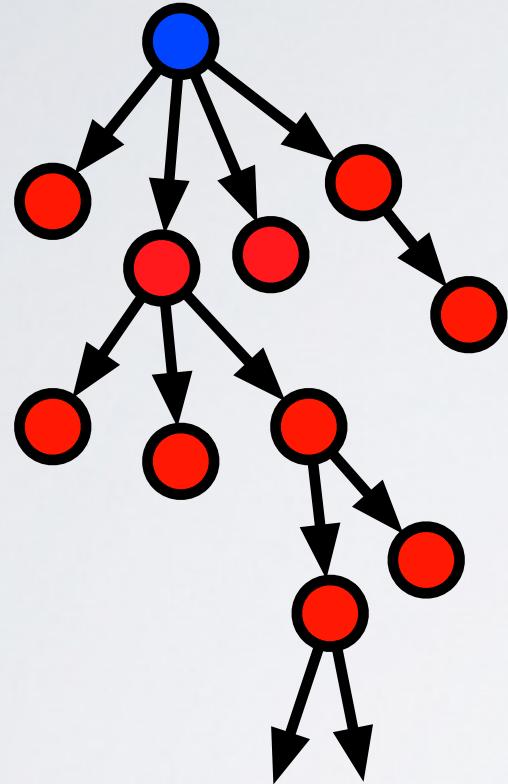
“critical”  
outbreaks of all sizes



$$R_0 > 1$$

“super-critical”  
global epidemics

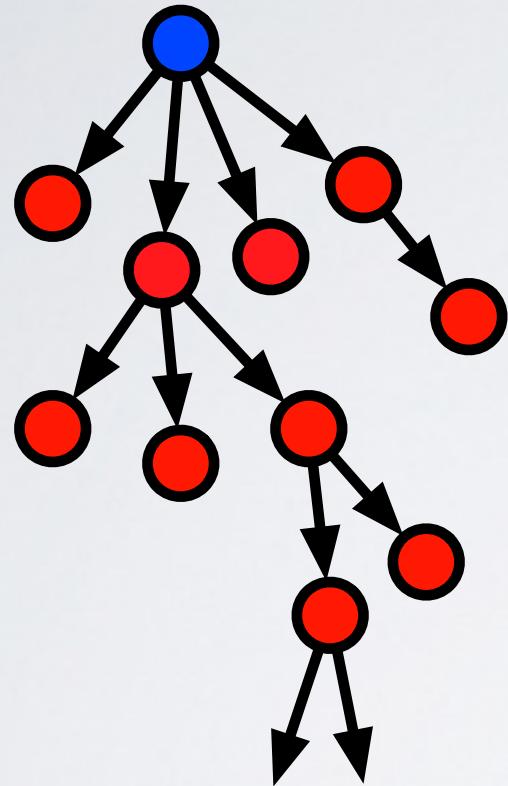
# network degrees



disease	R0	Theoretical value to decrease R_0 below 1 vaccination minimum
Measles	5-18	90-95%
Chicken pox	7-12	85-90%
Polio	5-7	82-87%
Smallpox	1.5-20+	70-80%
H1N1 influenza	1.0-3.0	

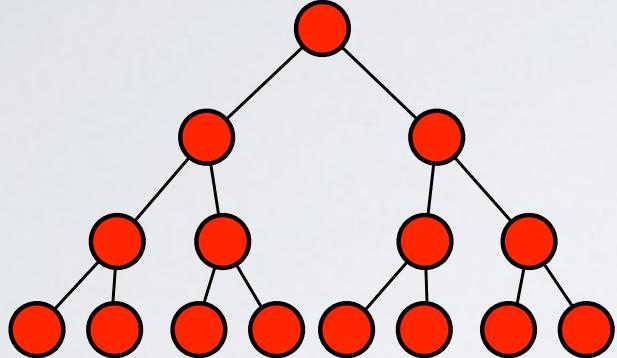
A large, solid black arrow pointing directly upwards, positioned above the text "all super-critical".

# network degrees



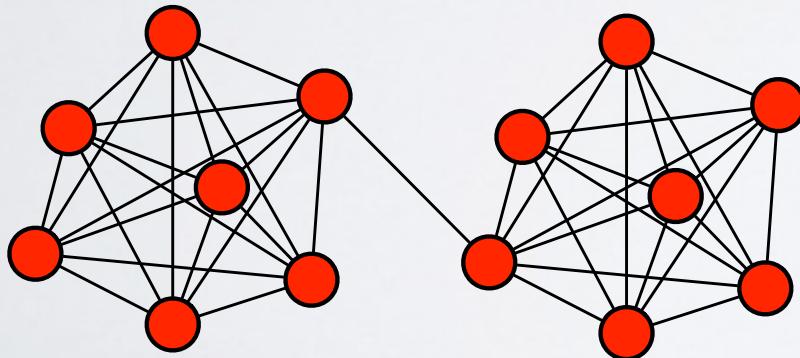
disease	R <sub>0</sub>	vaccination minimum
Measles	5-18	90-95%
Chicken pox	7-12	85-90%
Polio	5-7	82-87%
Smallpox	1.5-20+	70-80%
H1N1 influenza	1.0-3.0	

# network degrees



## bigger cascades

- smaller overlap among neighbors
- more expander-like  
[more like a random graph]
- higher transmission probability
- lower activation threshold



## smaller cascades

- larger overlap among neighbors
- more triangles
- smaller "communities"
- more spatial-like organization
- lower transmission probability
- higher activation threshold

Volz, J. Math. Bio. **56**, 293–310 (2008)

Bansal et al., J. Royal Soc. Interface **4**, 879–891 (2007)

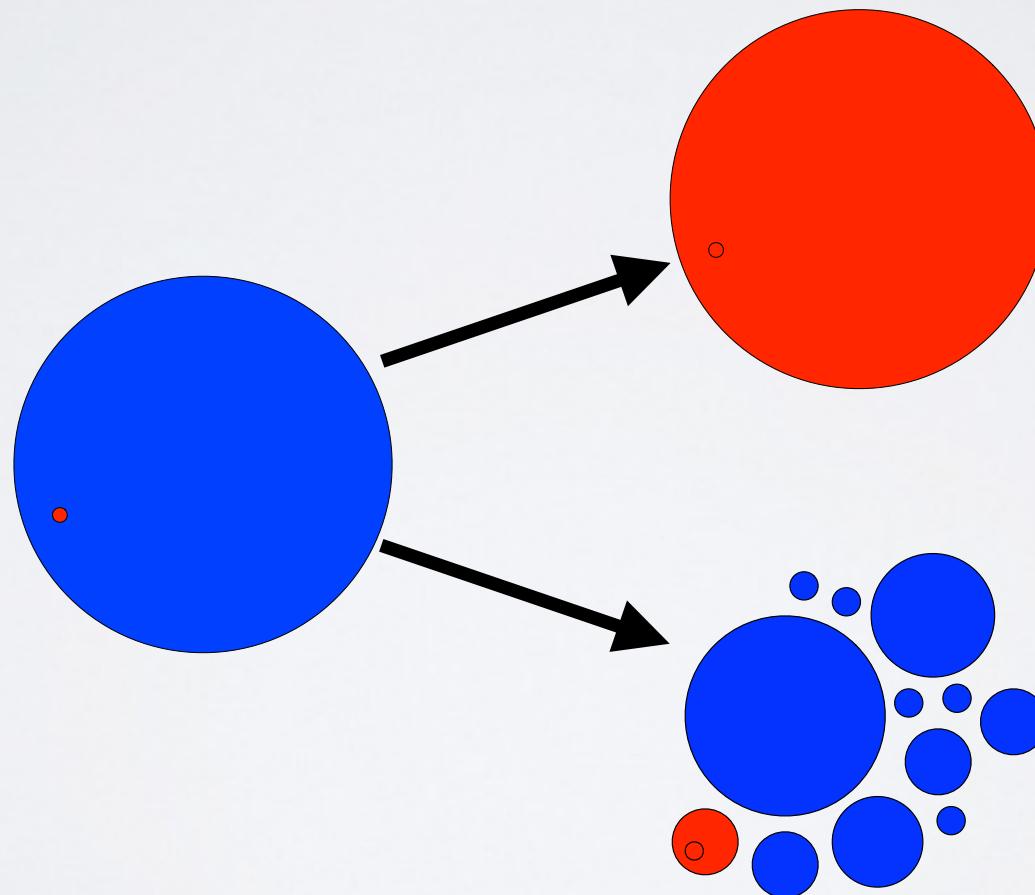
Karrer and Newman, Phys. Rev. E **82**, 016101 (2010)

Salathe and Jones, PLoS Comp. Bio. **6**, e1000736 (2010)

# network degrees

**how could we halt the spread?**

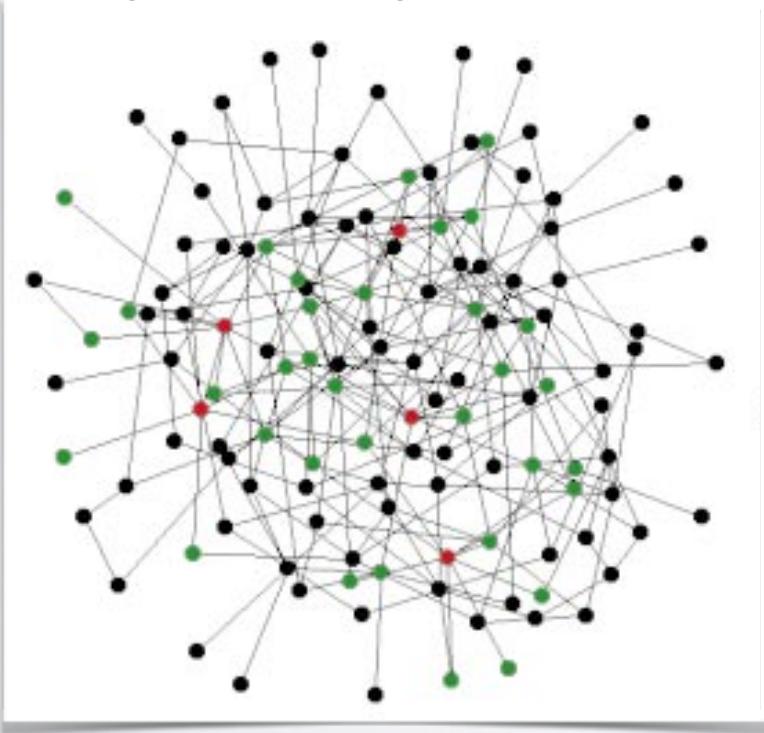
- break network into disconnected pieces



# network degrees

## two networks

homogeneous in degree

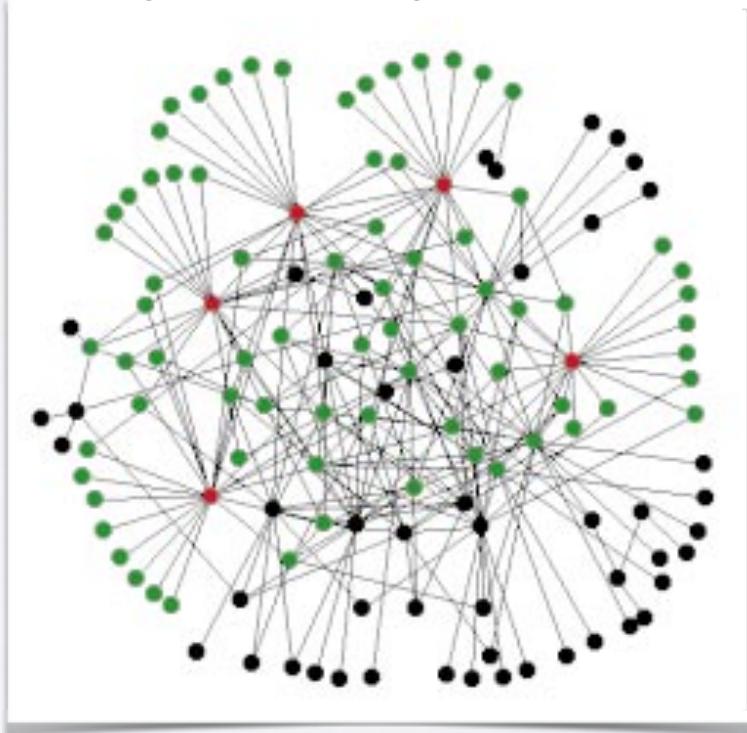


## Error and attack tolerance of complex networks

2000

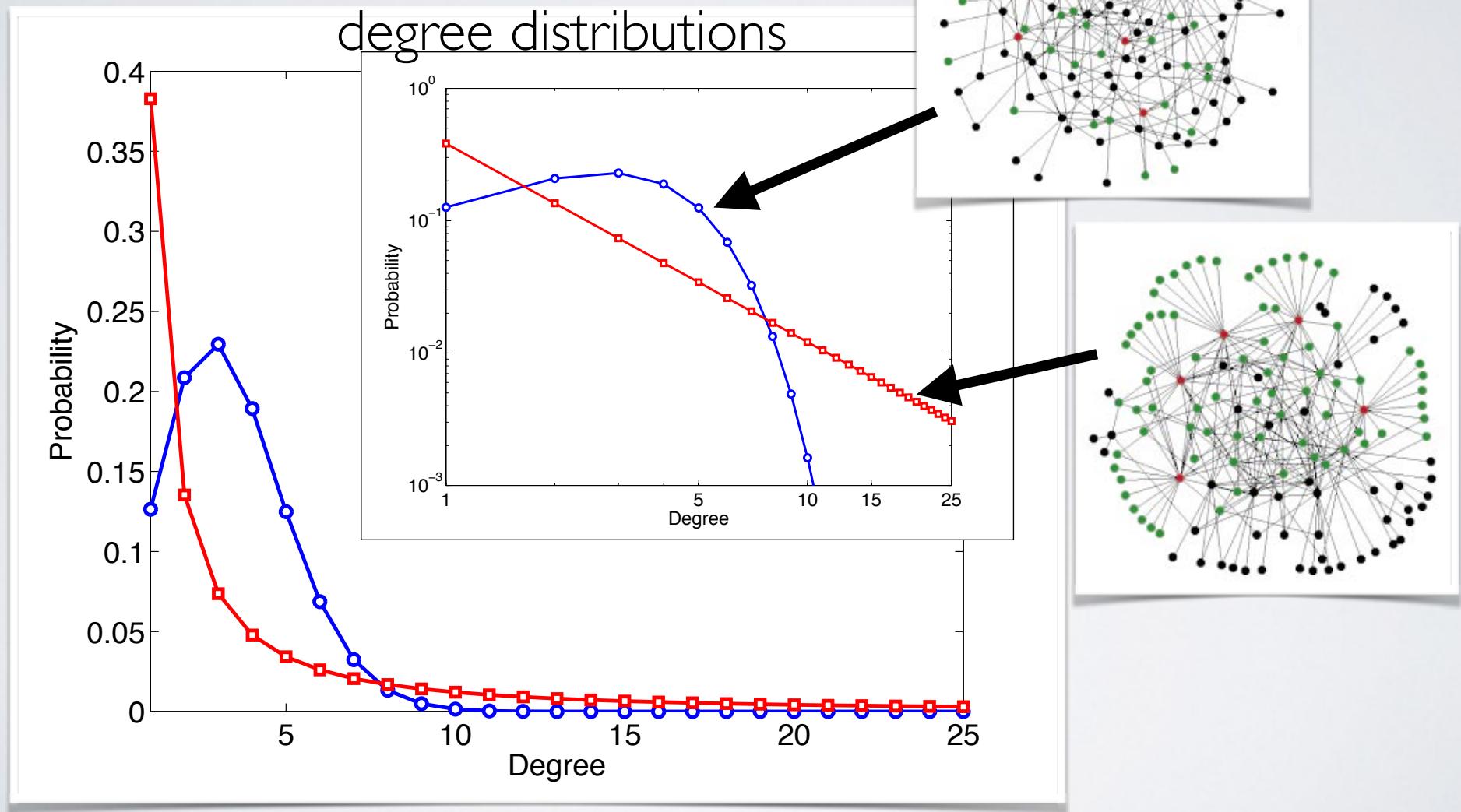
Réka Albert, Hawoong Jeong & Albert-László Barabási

heterogeneous in degree



# network degrees

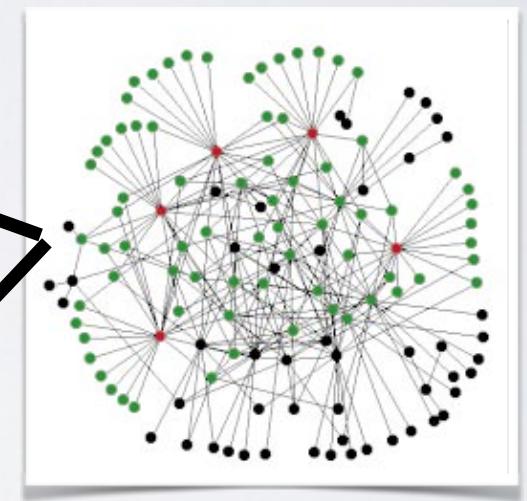
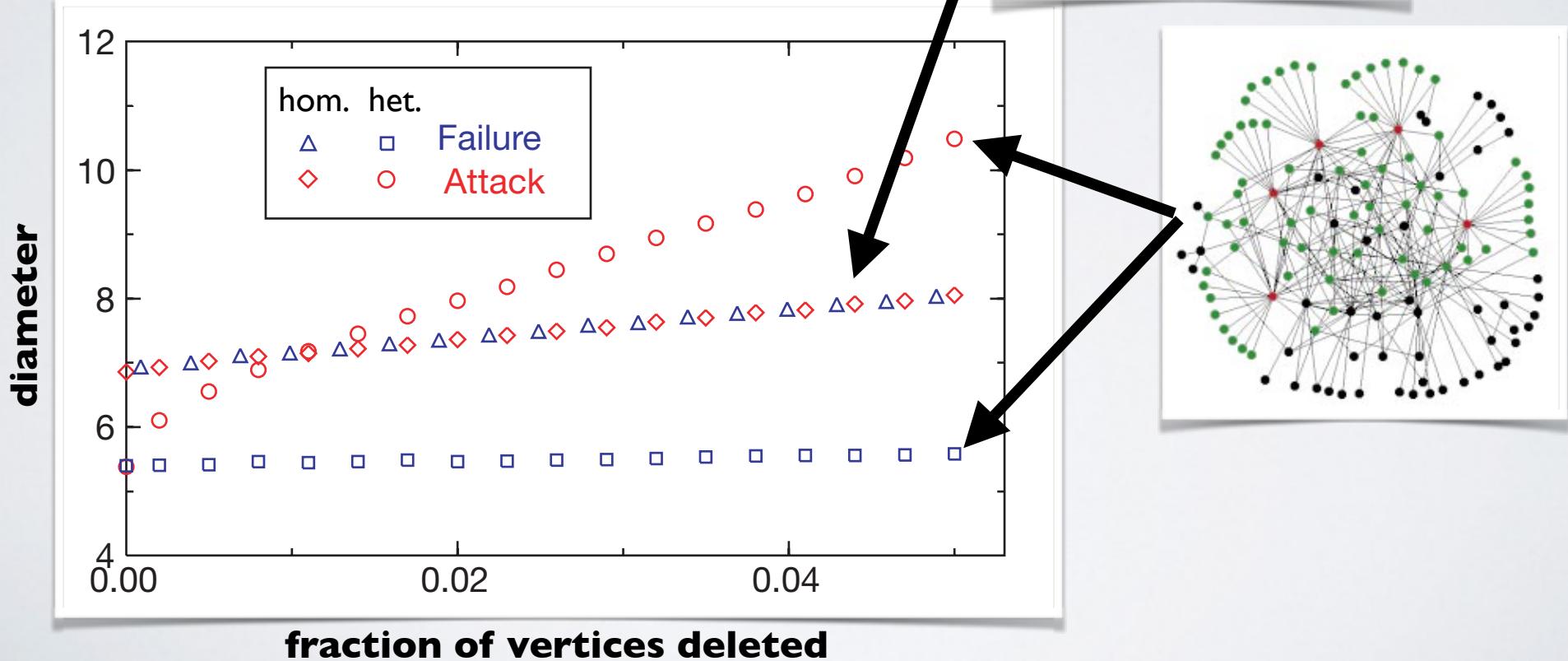
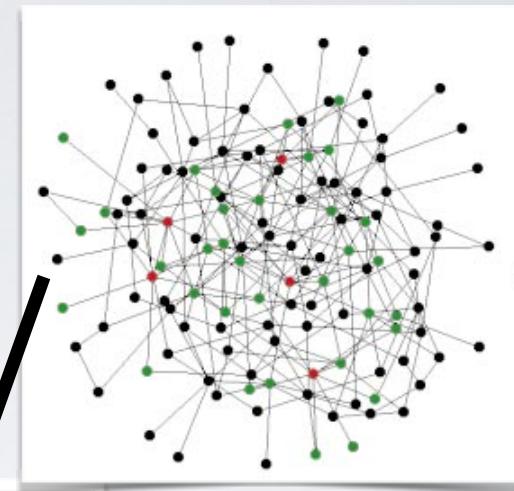
## two networks



# network degrees

## strategy: delete vertices

1. uniformly at random ("failure")
2. in order of degree ("attack")

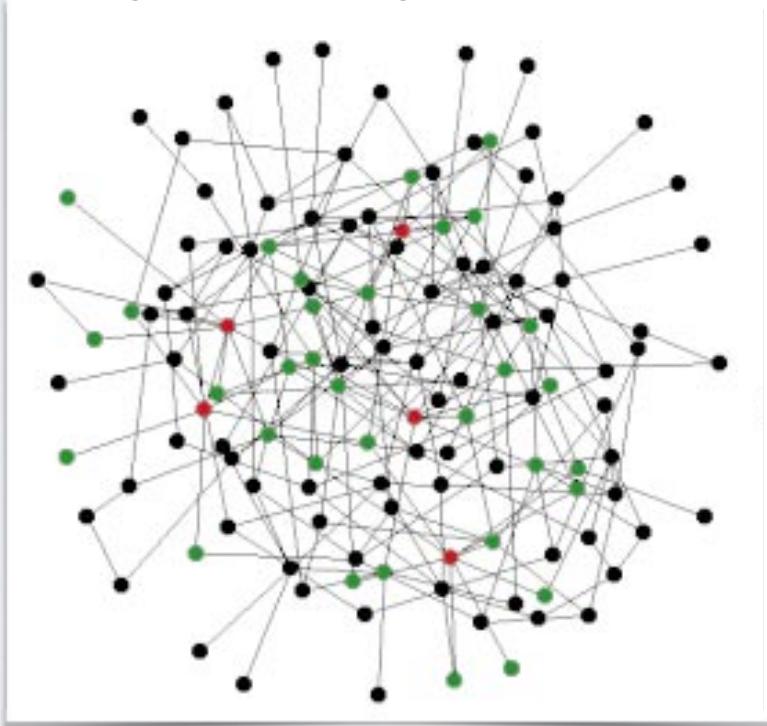


# network degrees

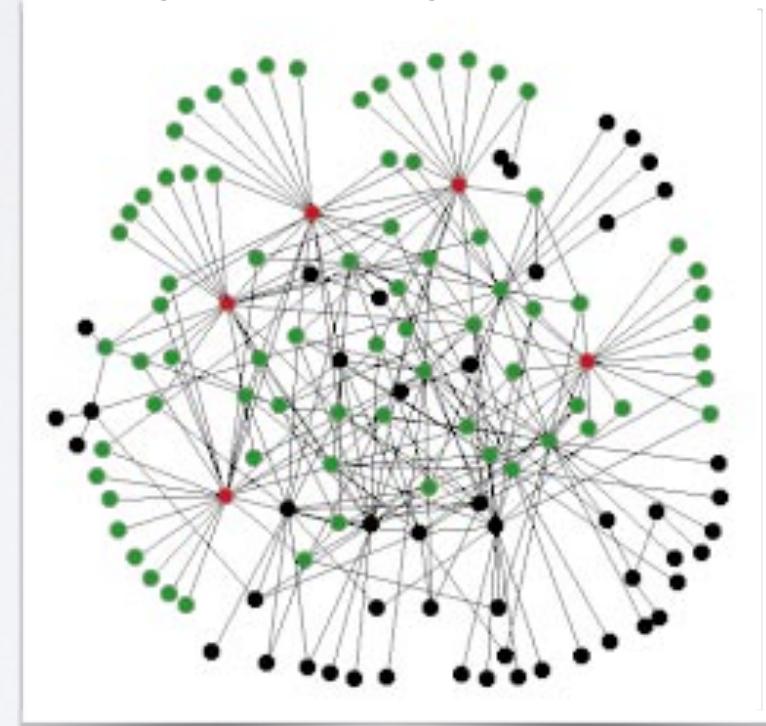
## what promotes spreading?

- high-degree vertices\*
- centrally-located vertices

homogeneous in degree



heterogeneous in degree

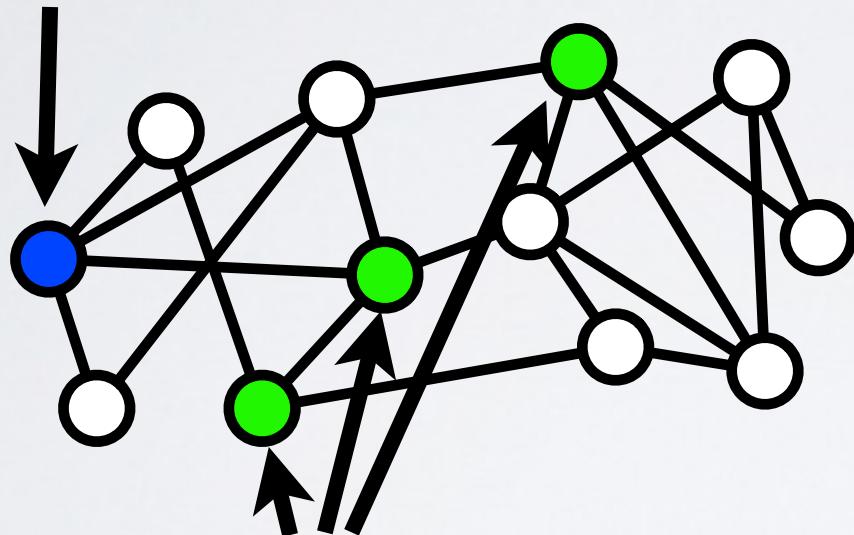


# network degrees

# strategy: delete vertices

3. build “fire breaks”

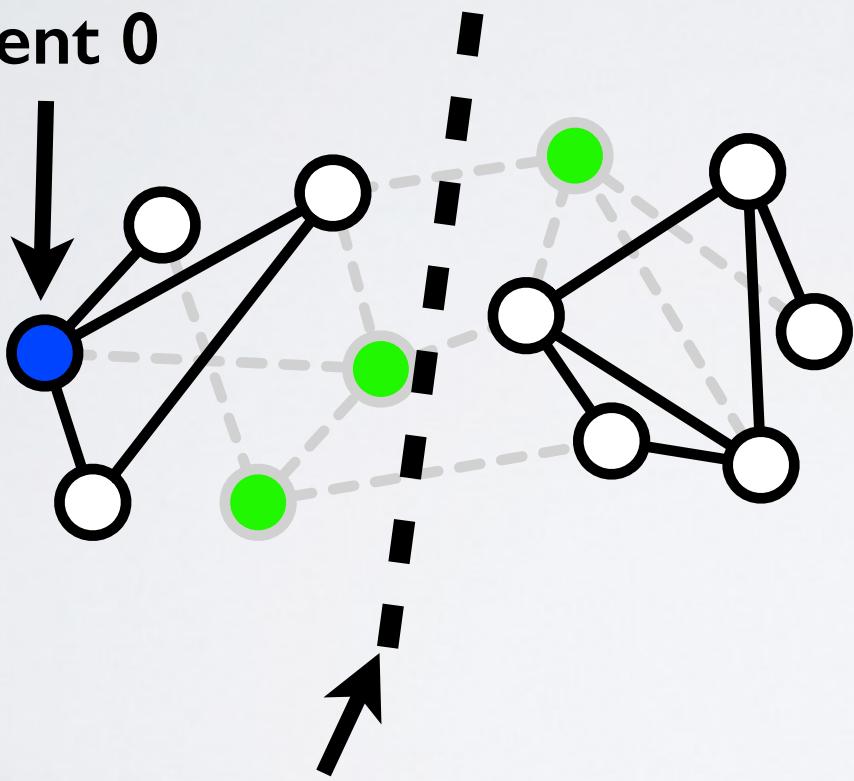
# patient 0



**vaccinated = deleted**  
("fire break")

# network degrees

**patient 0**



**effective buffer**

- **vaccination strategies**
  - the “front line” (hospitals)
  - high degree nodes
  - the vulnerable (old/young)

**network degrees**

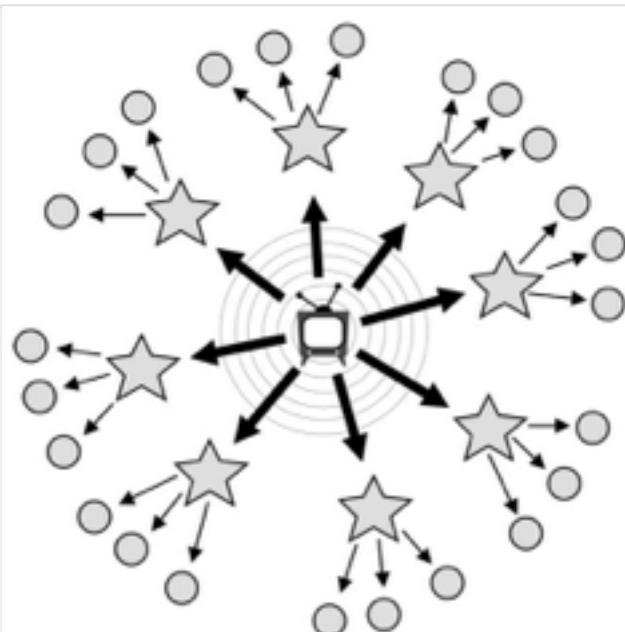
**but, in social networks...**

# network degrees

## Influentials, Networks, and Public Opinion Formation

DUNCAN J. WATTS  
PETER SHERIDAN DODDS\*

2007



- classic information marketing
- message saturation
- **degree** is most important

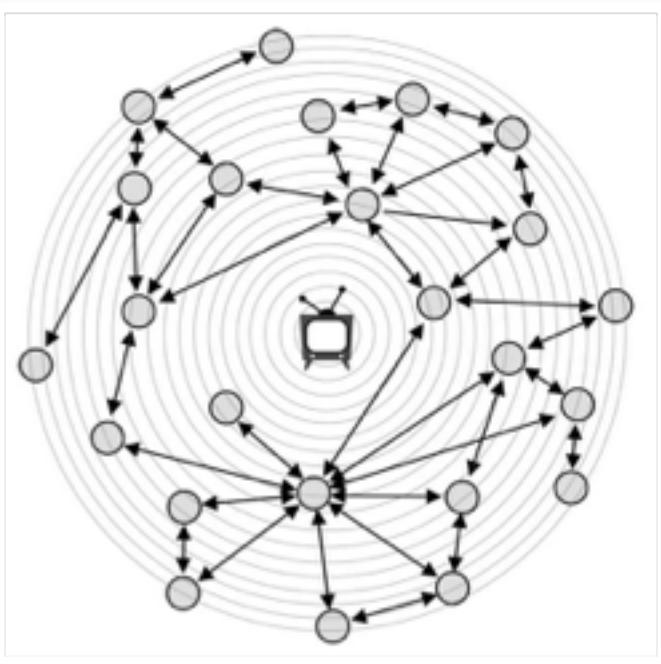
broadcast influence

# network degrees

## Influentials, Networks, and Public Opinion Formation

DUNCAN J. WATTS  
PETER SHERIDAN DODDS\*

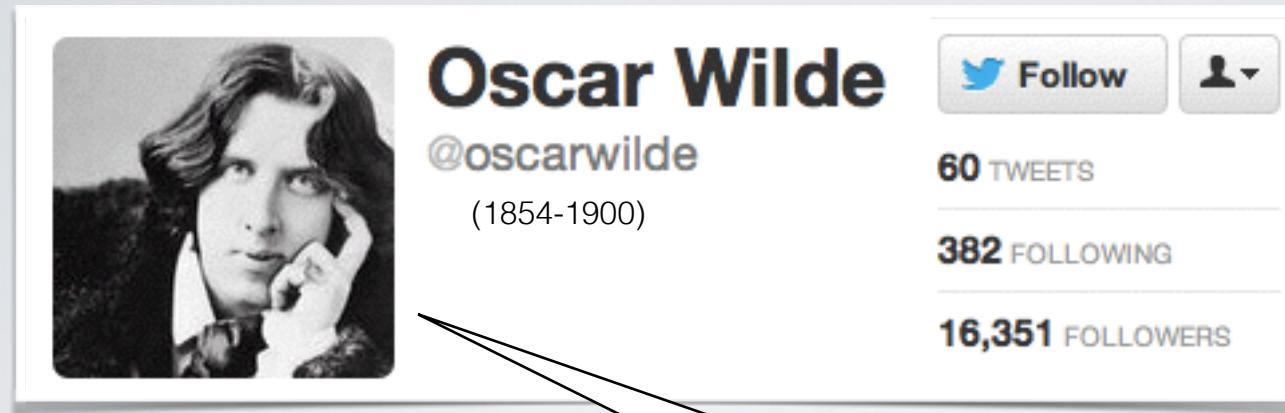
2007



network influence

- “network” (decentralized) marketing
- high-degree = “opinion leader”
- high-degree alone = **irrelevant**
- a cascade requires a legion of *susceptibles* (a system-level property)

# network degrees



“The only thing worse than being talked about is not being talked about.”

- "influence" not really about the influencer
- as much about the susceptibles

# network degrees

how to start a **social movement?**

# network degrees

how to start a **social movement?**

