Social Network Analysis for Computer Scientists

Vincent Traag

CWTS, Leiden University
https://liacs.leidenuniv.nl/~takesfw/SNACS

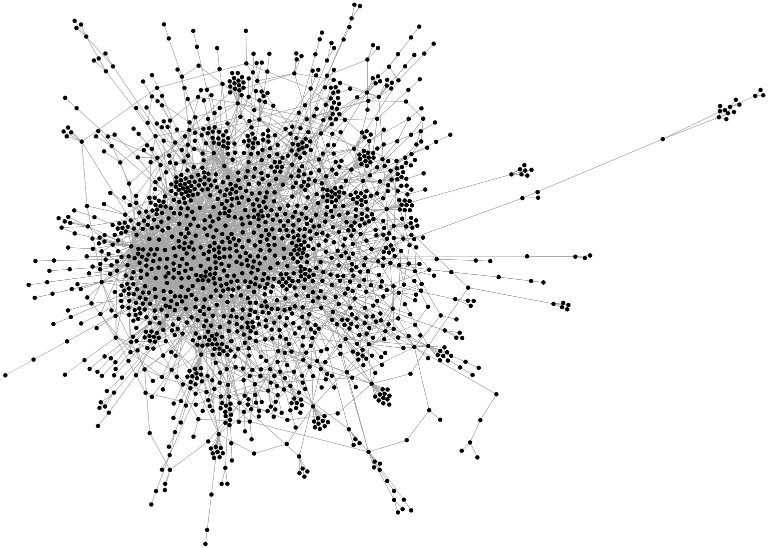
Lecture 6 — Dynamics on Networks

Overview today

- Recap
- Epidemic spreading
- Complex contagion
- Social influence & homophily

Recap

Networks



Notation

Concept	Symbol
Network (graph)	$G = (V, E)$
Nodes (objects, vertices, ...)	V
Links (ties, relationships, ...)	E
- Directed — $E \subseteq V \times V$ — "links"	
- Undirected — "edges"	
Number of nodes — $\|V\|$	n
Number of edges — $\|E\|$	m
Degree of node u	$deg(u)$
Distance from node u to v	$d(u, v)$

Real-world networks

1. Sparse networks — density
2. Fat-tailed power-law degree distribution — degree
3. Giant component — components
4. Low pairwise node-to-node distances — distance
5. Many triangles — clustering coefficient

Real-world networks

1. Sparse networks — density
2. Fat-tailed power-law degree distribution — degree
3. Giant component — components
4. Low pairwise node-to-node distances — distance
5. Many triangles — clustering coefficient

- Many examples: communication networks, citation networks, collaboration networks (Erdös, Kevin Bacon), protein interaction networks, information networks (Wikipedia), webgraphs, financial networks (Bitcoin) . . .

Advanced concepts

- Assortativity, homophily
- Reciprocity
- Power law exponent
- Planar graphs
- Complete graphs
- Subgraphs
- Trees
- Spanning trees
- Diameter, eccentricity
- Bridges
- Graph traversal: DFS, BFS

Centrality measures

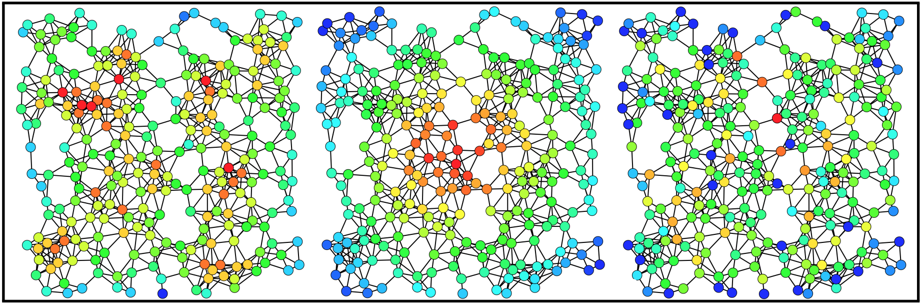
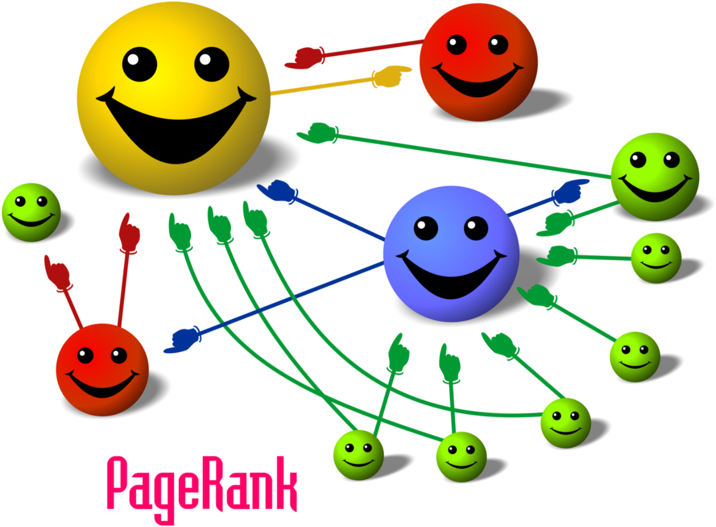


Figure: Degree, closeness and betweenness centrality

Source: "Centrality"' by Claudio Rocchini, Wikipedia `File:Centrality.svg`

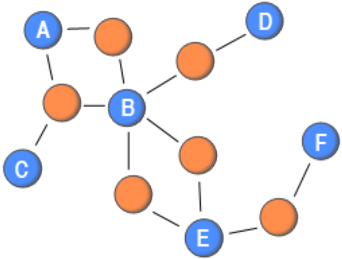
Centrality measures: PageRank



Centrality measures

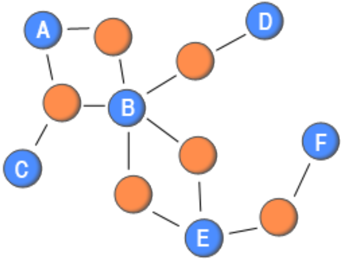
- Distance/path-based measures:
  - Degree centrality $O(n)$
  - Closeness centrality $O(mn)$
  - Betweenness centrality $O(mn)$
  - Eccentricity centrality $O(mn)$
- **Propagation-based** measures:
  - Hyperlink Induced Topic Search (HITS) $O(m)$
  - PageRank $O(m)$

Network projection



→

Network projection



→

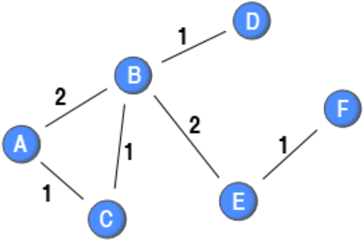
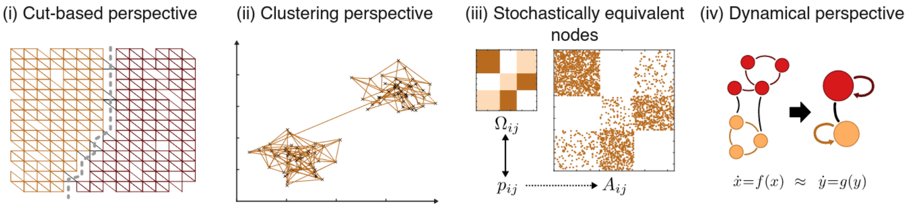


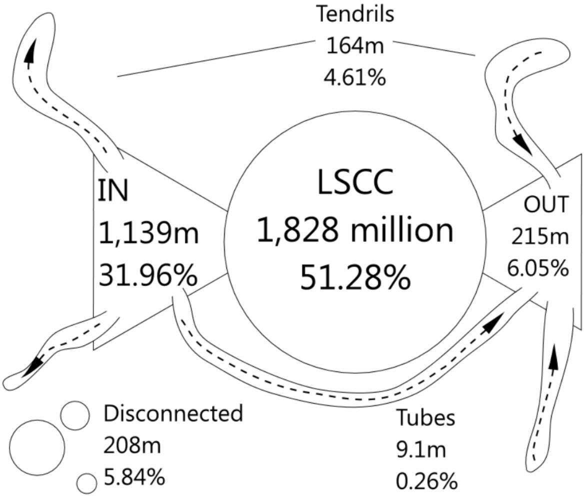
Image: `http://toreopsahl.com`

Community detection



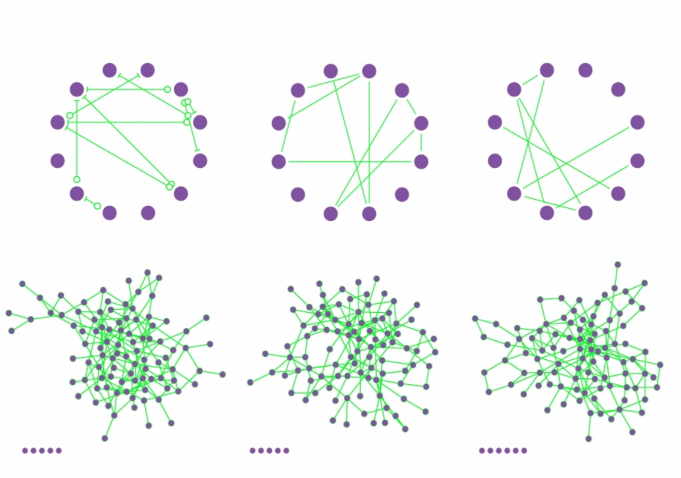
M Rosvall *et al.*, in *Advances in Network Clustering and Blockmodeling*, pp. 105–119, DOI 10.1002/9781119483298.ch4.

Bow-tie structure of the web



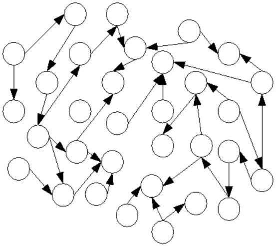
Meusel et al., Graph Structure in the Web — Revisited, WWW 2014: 427–431, 2014.

Erdös-Rényi random graphs

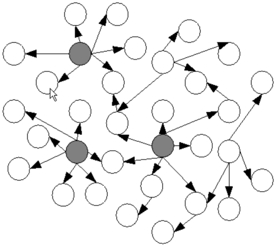


https://networksciencebook.com/chapter/3

Scale-free / Barábasi-Albert graphs



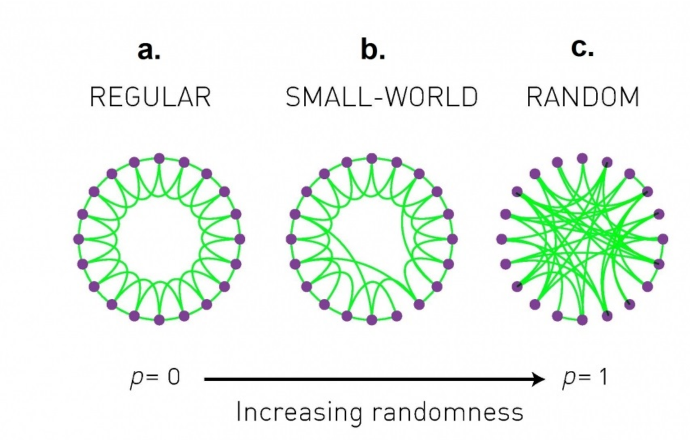
(a) Random network



(b) Scale-free network

B. Svenson, Complex networks and social network analysis in information fusion

Small-world / Watts & Strogatz graphs



http://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/bgc-sci.jpg

Dynamics on networks

Possible processes

- Random walks
- Flows / diffusion
- Contagion
  - Simple contagion
  - Complex contagion
- Social influence
- Opinion models

Possible processes

- Random walks
- Flows / diffusion
- **Contagion**
  - **Simple contagion**
  - **Complex contagion**
- **Social influence**
- Opinion models

Types of contagion

Simple contagion

- Single contact enough to "infect".
- e.g. epidemiology, computer viruses, information flow

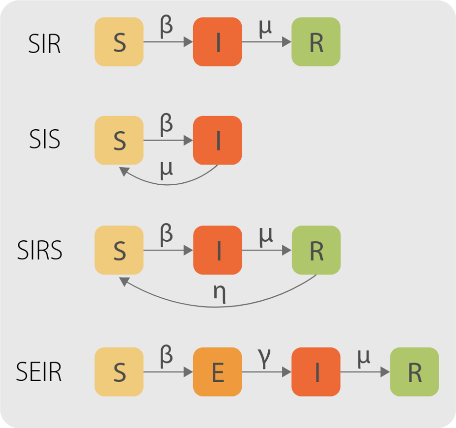
Complex contagion

- Multiple contacts reinforce "infection"
- Need multiple contacts before "infection" (i.e. threshold)
- e.g. social behaviour, especially collective action

Simple contagion

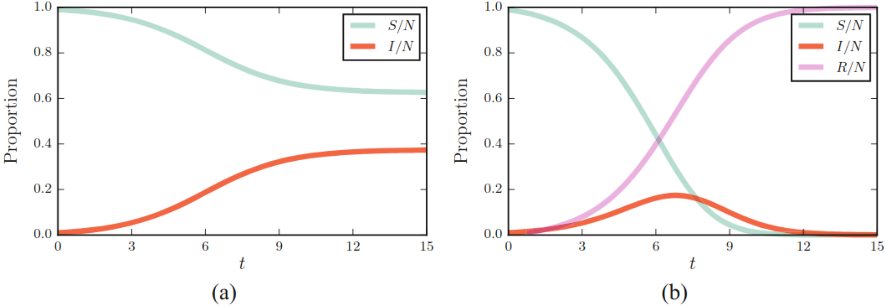
Epidemiological models

- **S**usceptible
- **E**xposed
- **I**nfected
- **R**ecovered



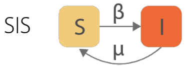
R Pastor-Satorras *et al.*, *Rev. Mod. Phys.* **87**, 925–979, DOI 10.1103/RevModPhys.87.925 (2014).

Epidemiological model example



IZ Kiss, JC Miller, PL Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models*, vol. 46, DOI 10.1007/978-3-319-50806-1.

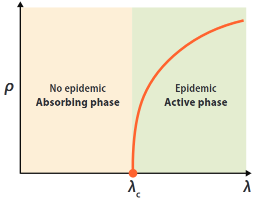
Epidemiological threshold



$$\dot{S} = \mu I - \beta I \frac{S}{N}$$

$$\dot{I} = \beta I \frac{S}{N} - \mu I$$

- Two steady states
  - $I = 0$
  - $I = N\left(1 - \frac{\mu}{\beta}\right)$
- Basic reproduction number $R_0 = \frac{\beta}{\mu}$
- Epidemic when $R_0 > 1$.



R Pastor-Satorras *et al.*, *Rev. Mod. Phys.* **87**, 925–979, DOI 10.1103/RevModPhys.87.925 (2014).

Epidemics on networks

- Fixed contacts, no switching.
- Epidemic threshold related to appearance of giant component.
- Scale-free graphs: no epidemic threshold.

Erdős-Rényi graphs

- Giant component appears when $np = \langle k \rangle > 1$.
- Same condition as basic reproduction number $R_0 = \frac{\beta}{\mu} > 1$.

Scale-free graphs

- Giant component appears when $\langle k^2 \rangle - 2\langle k \rangle > 0$.
- For scale free degree distributions $\Pr(k) \sim k^{-\alpha}$, always giant component if $\alpha \leq 3$.

Simulate epidemics on networks

Three approaches

- Naive approach
  - Loop over nodes, infect/recover with some probability
- Gillespie algorithm
  - Calculate waiting time for infect/recover
- Event-based algorithm
  - Pre-calculate time of events, process events in order

Naive SIR simulation

Naive approach (discrete time)

- Loop over nodes
  - Infect neighbour with probability $\sim \beta$.
  - Recover infected node with probability $\sim \mu$.

Gillespie SIR algorithm

- Loop while dynamics
  - Sample time to next event
  - Sample event (Infection or Recovery)
  - Sample node to apply event to
  - Update event rates

Event-based SIR algorithm

- Add initial infections to queue
- Loop while event queue
    - If infection event
        - Infect node, and create infect event for all neigbhours
    - else if (recovery event)
        - Recover node

Immunisation

Immunise nodes so that effective $R_0 < 1$

- Select optimal set $S \subseteq V(G)$ of size $k = |S|$.
- NP-hard problem
- Random immunisation
- Global heuristics
  - Degree centrality
  - Eigenvector centrality
  - Betweenness centrality
  - Community separation
- Local heuristics
  - Acquaintance immunisation
  - Community bridge finding

Immunisation

Immunise nodes so that effective $R_0 < 1$

- Select optimal set $S \subseteq V(G)$ of size $k = |S|$.
- NP-hard problem
- **Random immunisation**
- Global heuristics
    - **Degree centrality**
    - Eigenvector centrality
    - Betweenness centrality
    - Community separation
- Local heuristics
    - **Acquaintance immunisation**
    - Community bridge finding

Random immunisation

Immunise fraction g of the nodes randomly.

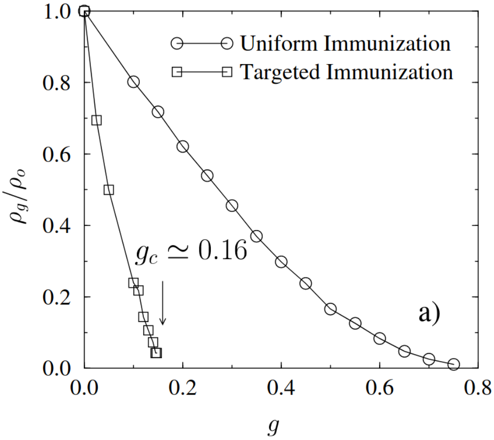
Erdős-Rényi graphs

- Every node infects $\langle k \rangle$ nodes on average.
- If g nodes are immunised, every node infects $(1-g)\langle k \rangle$ nodes.
- No epidemic spreading if $(1-g)\langle k \rangle < 1$, so $g > 1 - \frac{1}{\langle k \rangle}$.
- Same general reasoning, basic reproduction number R_0, critical vaccination rate $g_c = 1 - \frac{1}{R_0}$.

Scale-free

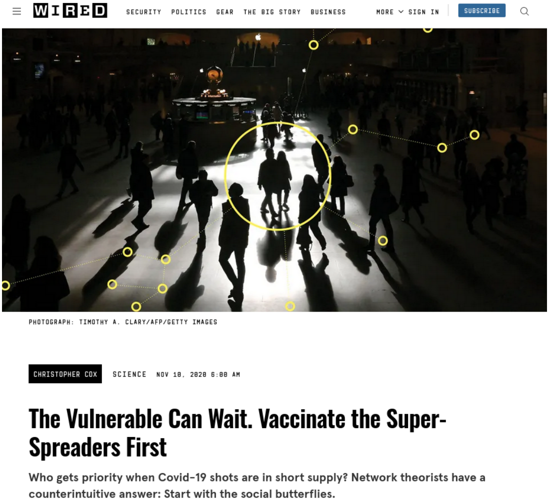
- Random removal, critical immunisation $g_c = 1 - \frac{\langle k \rangle}{\langle k^2 \rangle}$.
- If $\langle k^2 \rangle \to \infty$, only effective immunisation, $g_c = 1$.
- No epidemic threshold, no effective immunisation.

Targeted immunisation



R Pastor-Satorras, A Vespignani, *Phys. Rev. Lett.* **86**, 3200–3203, DOI 10.1103/PhysRevLett.86.3200 (2001).

Targeted immunisation



WIRED SECURITY POLITICS GEAR THE BIG STORY BUSINESS MORE SIGN IN SUBSCRIBE

PHOTOGRAPH: TIMOTHY A. CLARY/AFP/GETTY IMAGES

CHRISTOPHER COX SCIENCE NOV 10, 2020 6:00 AM

The Vulnerable Can Wait. Vaccinate the Super-Spreaders First

Who gets priority when Covid-19 shots are in short supply? Network theorists have a counterintuitive answer: Start with the social butterflies.

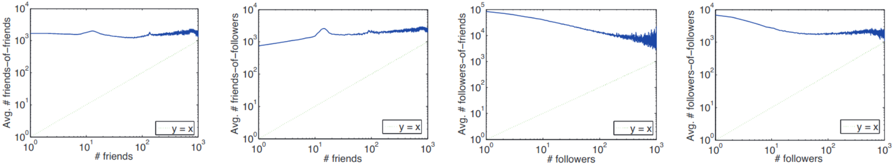
Friendship paradox

- Choose random neighbour of random node, probability $\frac{k_i}{2m}$ to end up at node i.
- Average degree of random neighbour

$$\sum_i k_i \frac{k_i}{2m} = \frac{1}{2m} \sum_i k_i^2 = \frac{n}{2m} \langle k^2 \rangle = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq \frac{\langle k \rangle^2}{\langle k \rangle} = \langle k \rangle$$

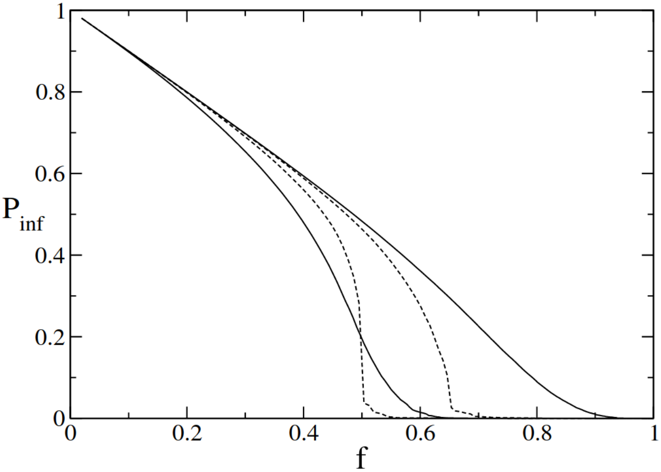
- Average degree of friends greater than average degree.

Friendship paradox



N Hodas, F Kooti, K Lerman, *Proceedings Int. AAAI Conf. on Web & Soc. Media* **7**, 225–233, DOI 10.1609/icwsm.v7i1.14440 (2013).

Acquaintance immunisation



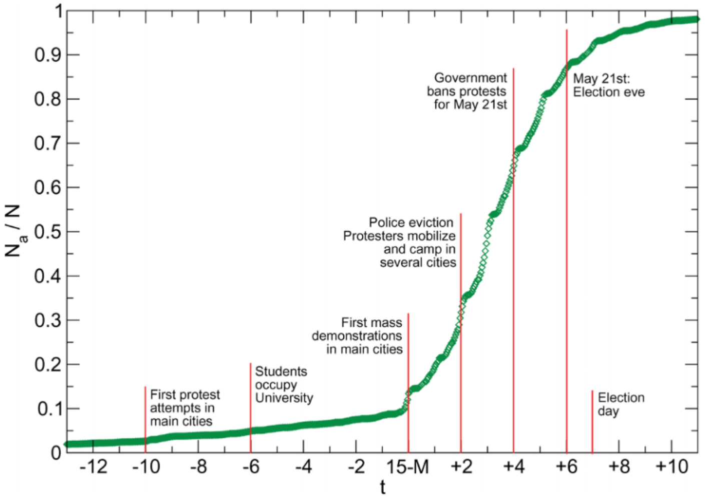
R Cohen, S Havlin, D ben-Avraham, *Phys. Rev. Lett.* **91**, 247901, DOI 10.1103/PhysRevLett.91.247901 (2003).

Complex contagion

Collective action

https://commons.wikimedia.org/wiki/File:
20110629_Moutza_demonstrations_Greek_parliament_Athens_Greece.jpg

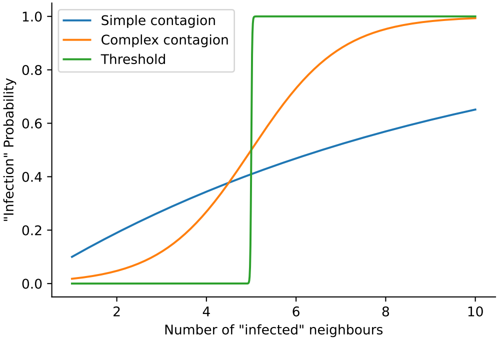
Spain 15-M movement



S González-Bailón *et al.*, *Scientific reports* **1**, 197, DOI 10.1038/srep00197 (2011).

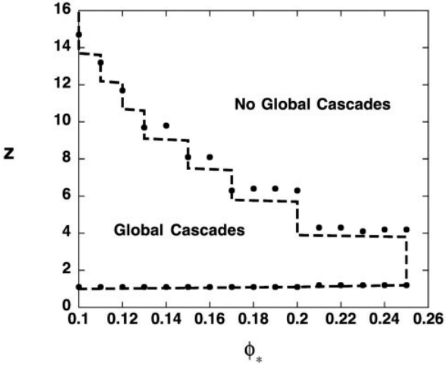
Complex contagion

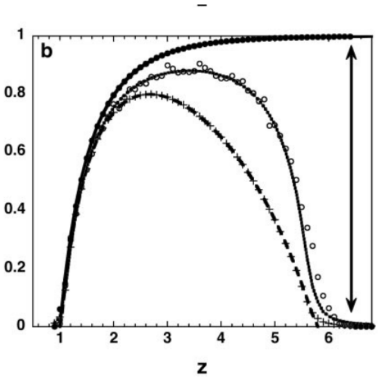
- "Infection" probability $f(N(v))$ not simple contagion $1-(1-p)^{k_i}$.
- Reinforcement, additional neighbour increases "infection" probability.
- Threshold models, i.e. step-wise function.



M Granovetter, *American journal of sociology* **83**, 1420–1443 (1978).

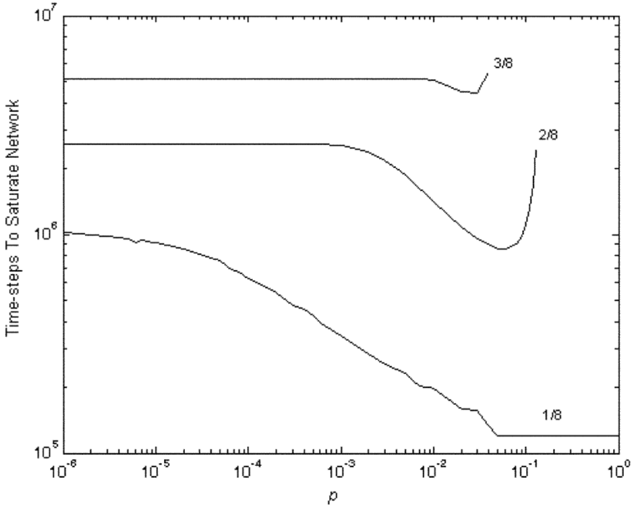
Complex contagion





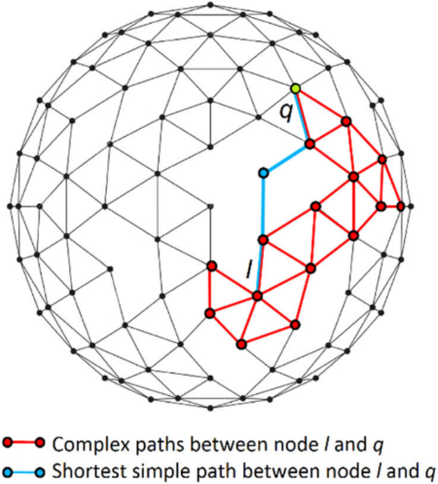
DJ Watts, *PNAS* **99**, 5766–5771, DOI 10.1073/pnas.082090499 (2002).

Complex and simple contagions in small-world networks



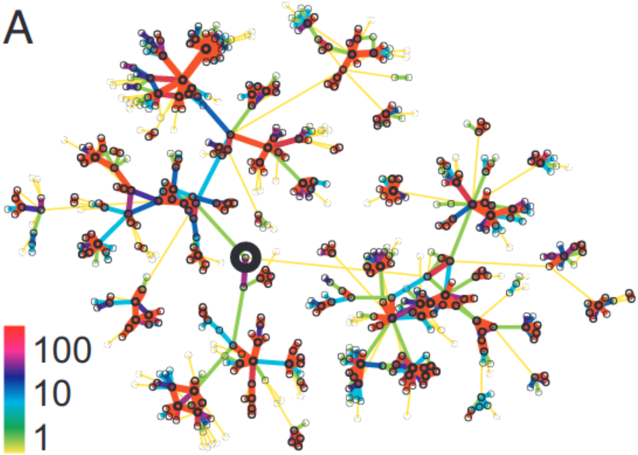
D Centola, M Macy, *Am. J. Sociol.* **113**, 702–734, DOI 10.1086/521848 (2007).

Complex vs simple contagions



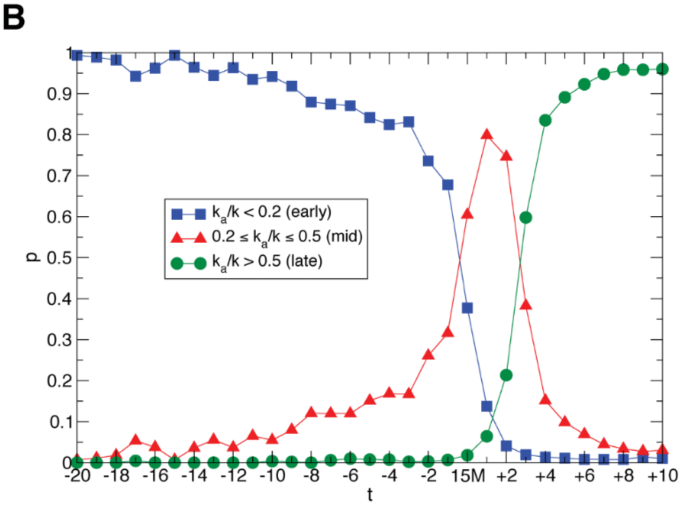
D Guilbeault, D Centola, *Nat Commun* **12**, 4430, DOI 10.1038/s41467-021-24704-6 (2021).

Weak ties



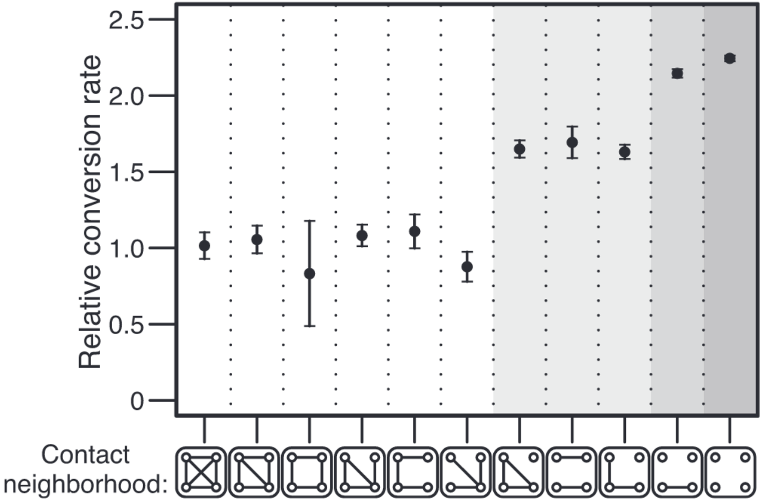
JP Onnela *et al.*, *Proc. Nat. Acad. Sci. USA* **104**, 7332–6, DOI 10.1073/pnas.0610245104 (2007).

Protest spread



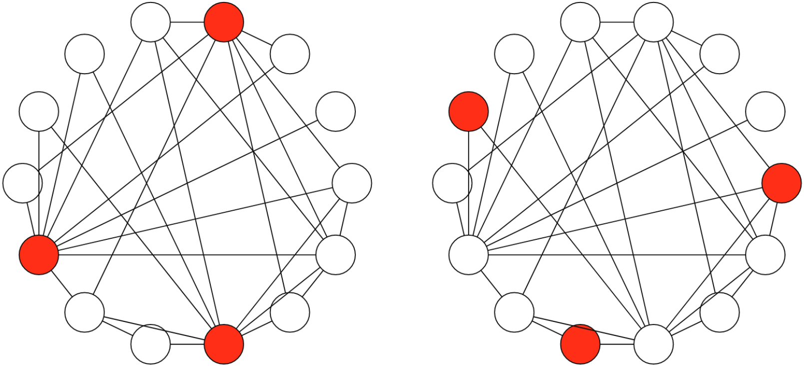
S González-Bailón *et al.*, *Scientific reports* **1**, 197, DOI 10.1038/srep00197 (2011).

Structural diversity



J Ugander *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5962–5966, DOI 10.1073/pnas.1116502109 (2012).

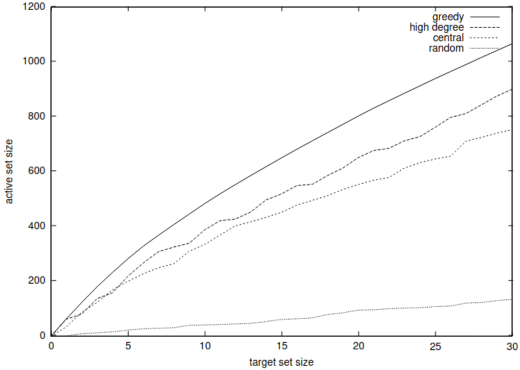
Majority illusion



K Lerman, X Yan, XZ Wu, *PLoS ONE* **11**, e0147617, DOI 10.1371/journal.pone.0147617 (2016).

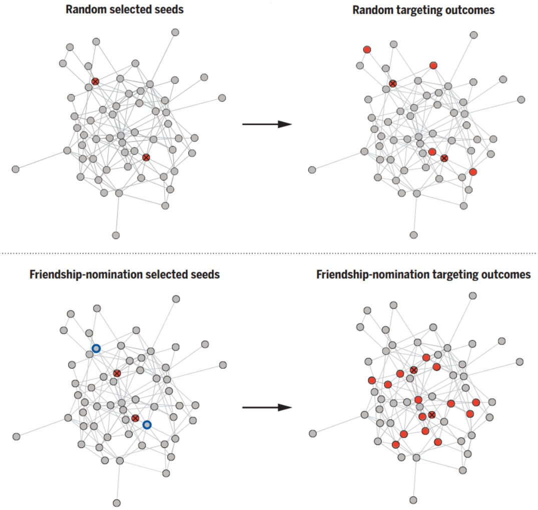
Social influence maximisation

Maximise social influence ∼ minimise spreading



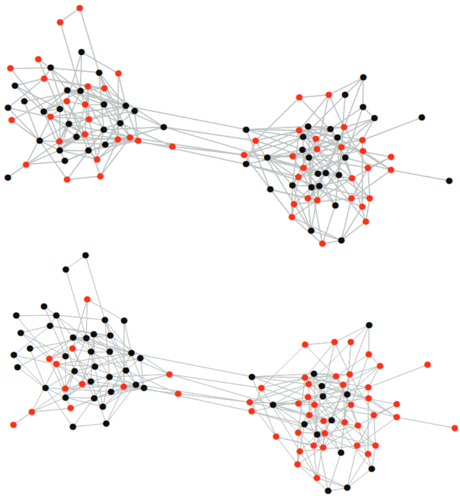
D Kempe, J Kleinberg, É Tardos, presented at the Proceed. 9th ACM SIGKDD Int. Conf. on Knowledge Disc. and Data Min. Pp. 137–146.

Friendship nomination



EM Airoldi, NA Christakis, *Science* **384**, eadi5147, DOI 10.1126/science.adi5147 (2024).

Social influence & homophily



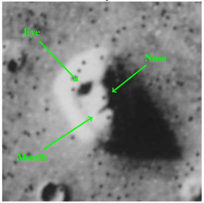
CR Shalizi, AC Thomas, *Sociological Methods & Research* **40**, 211–239, DOI 10.1177/0049124111404820 (2011).

Descriptive vs inferential community detection

TP Peixoto, *Elements in the Structure and Dynamics of Complex Networks*, DOI 10.1017/9781009118897 (2023).

Descriptive vs inferential community detection

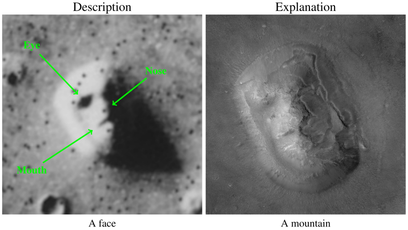
Description



A face

TP Peixoto, *Elements in the Structure and Dynamics of Complex Networks*, DOI 10.1017/9781009118897 (2023).

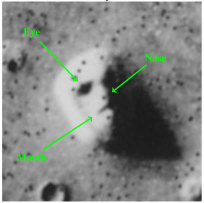
Descriptive vs inferential community detection



TP Peixoto, *Elements in the Structure and Dynamics of Complex Networks*, DOI 10.1017/9781009118897 (2023).

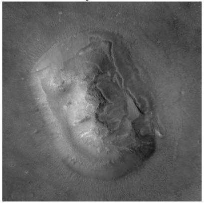
Descriptive vs inferential community detection

Description

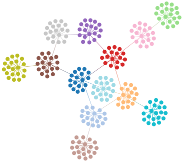


A face

Explanation



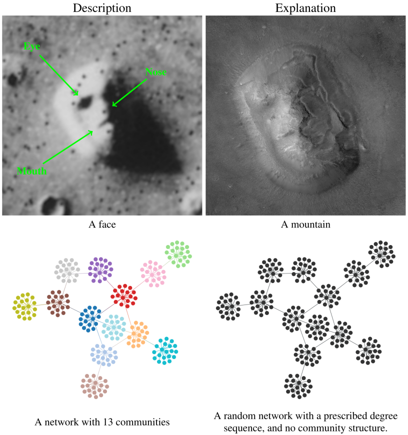
A mountain



A network with 13 communities

TP Peixoto, *Elements in the Structure and Dynamics of Complex Networks*, DOI 10.1017/9781009118897 (2023).

Descriptive vs inferential community detection



TP Peixoto, *Elements in the Structure and Dynamics of Complex Networks*, DOI 10.1017/9781009118897 (2023).

Practicalities

Upcoming lab session and next week

- Lab session: make serious progress with Assignment 2.
- Next week: student presentations; know your track letter.
  - Track A: room BW 0.08
  - Track B: room BW 0.19
  - Track C: room BW 0.20
- Detailed schedule: see course website; rooms are fixed; lecturer differs.
- Still no presentation slot? Contact course staff ASAP
- Presenting? On the Tuesday before your Friday presentation, send your slides to `snacs@liacs.leidenuniv.nl` for some feedback, or if you want, make an appointment the week before with the lecturer assigned to your room.