Network Science

#17 Network robustness

Robustness

A.L. Barabási, Network science, http://barabasi.com/networksciencebook

Ch.8 "Network robustness"

Network robustness

- We are now interested in network robustness to failures
- Want to understand how real networks work under imperfect conditions/malfunctioning

  e.g., why some mutations lead to diseases (biology & medicine)

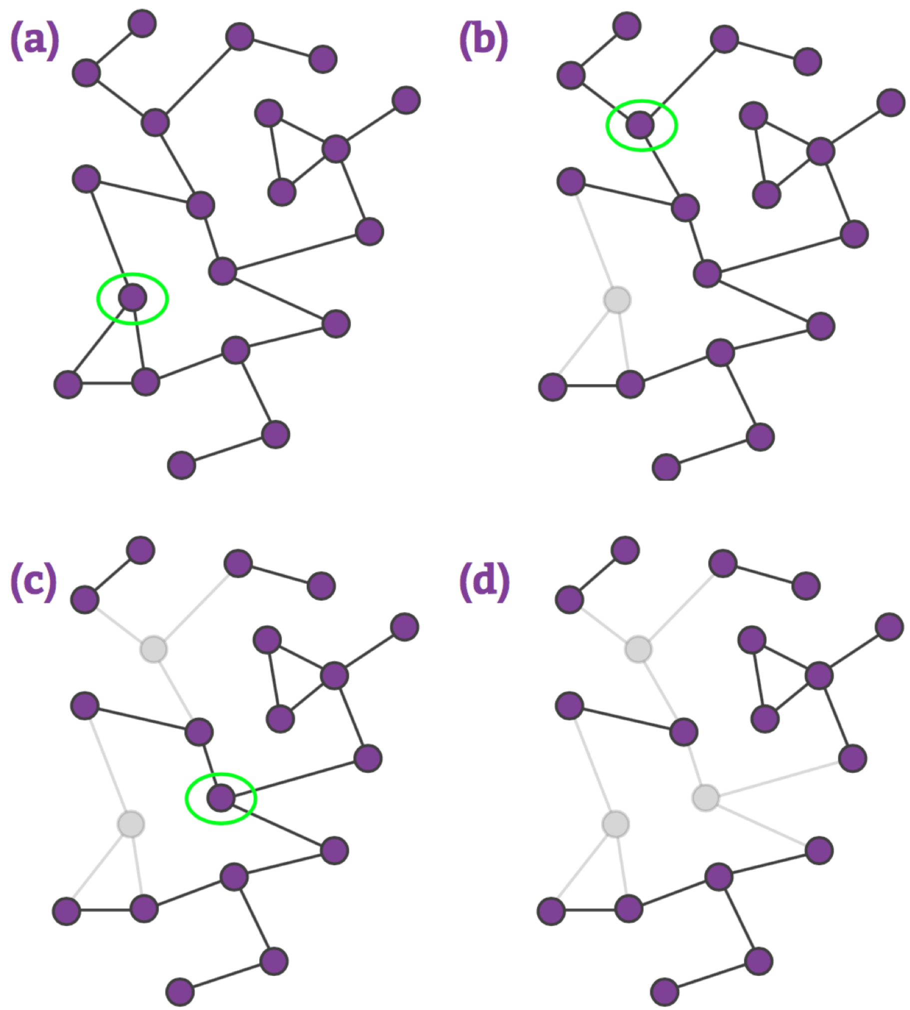
  stability of social networks to disruptive events (war, famine, etc)

  robustness to occasional failures in telecom networks/the www

Oak, Quercus Robur → robust

Network robustness

- Would the network still "work" in the presence of missing nodes?
- Failures can lead to either just isolating nodes or breaking the whole network apart
- What is the limit/phase transition?



Applications

This can serve to identify:

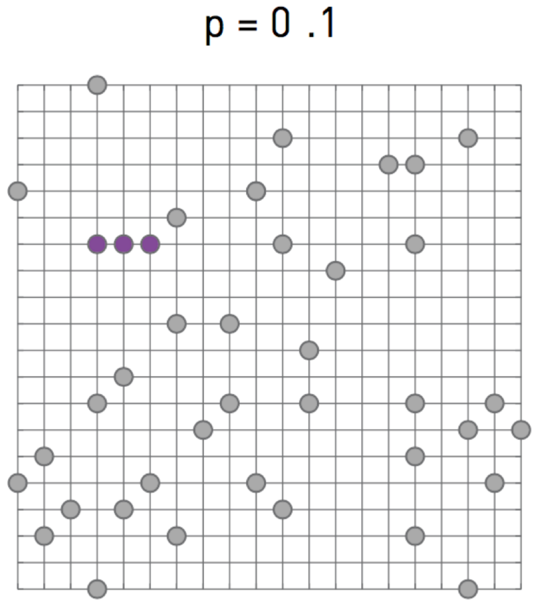
- robustness of air transportation under random strikes
- robustness of social contacts even when someone is off
- possibility of destroying of criminal/terror networks
- eradication of an epidemics

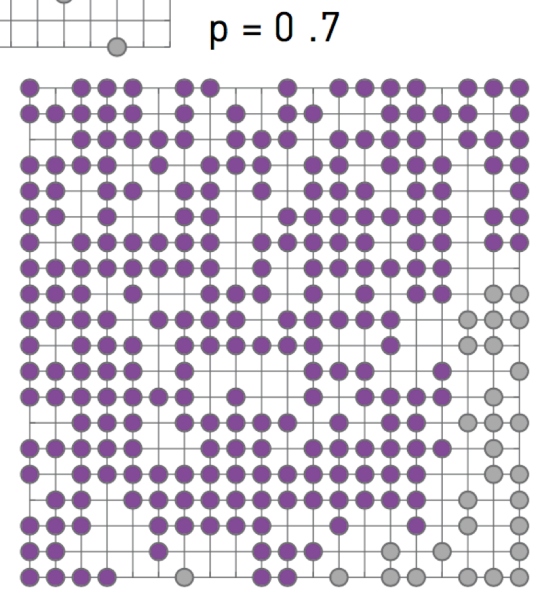
Percolation Theory

Percolation theory

- Pebbles are randomly paced, with probability *p*, over a square lattice
- What is the average cluster size?
- What is the expected size of the largest cluster? (percolating cluster)

→ Percolation theory predicts a sudden phase transition

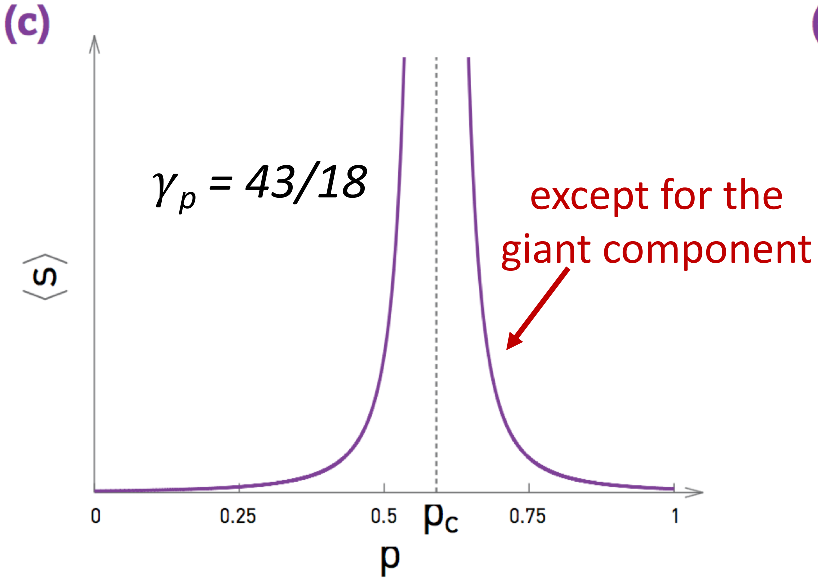




Critical transition

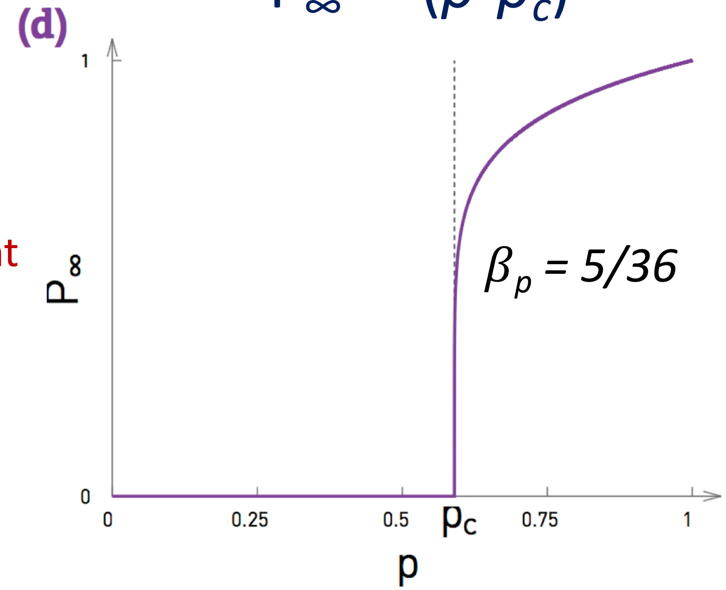
average cluster size

$\langle s \rangle \sim |p-p_c|^{-\gamma_p}$



probability of belonging to the largest cluster

$P_\infty \sim (p-p_c)^{\beta_p}$



- Critical transition at $p_c \sim 0.6$
- Around p_c small clusters grow and coalesce, leading to the emergence of a large cluster

Universality?

- Value p_c depends on the lattice type, and # of dimension

  p_c = 0.593 for a 2D square lattice

  p_c = 0.5 for a 2D triangular lattice

- Critical exponents γ_p and β_p only depend on # of dimensions

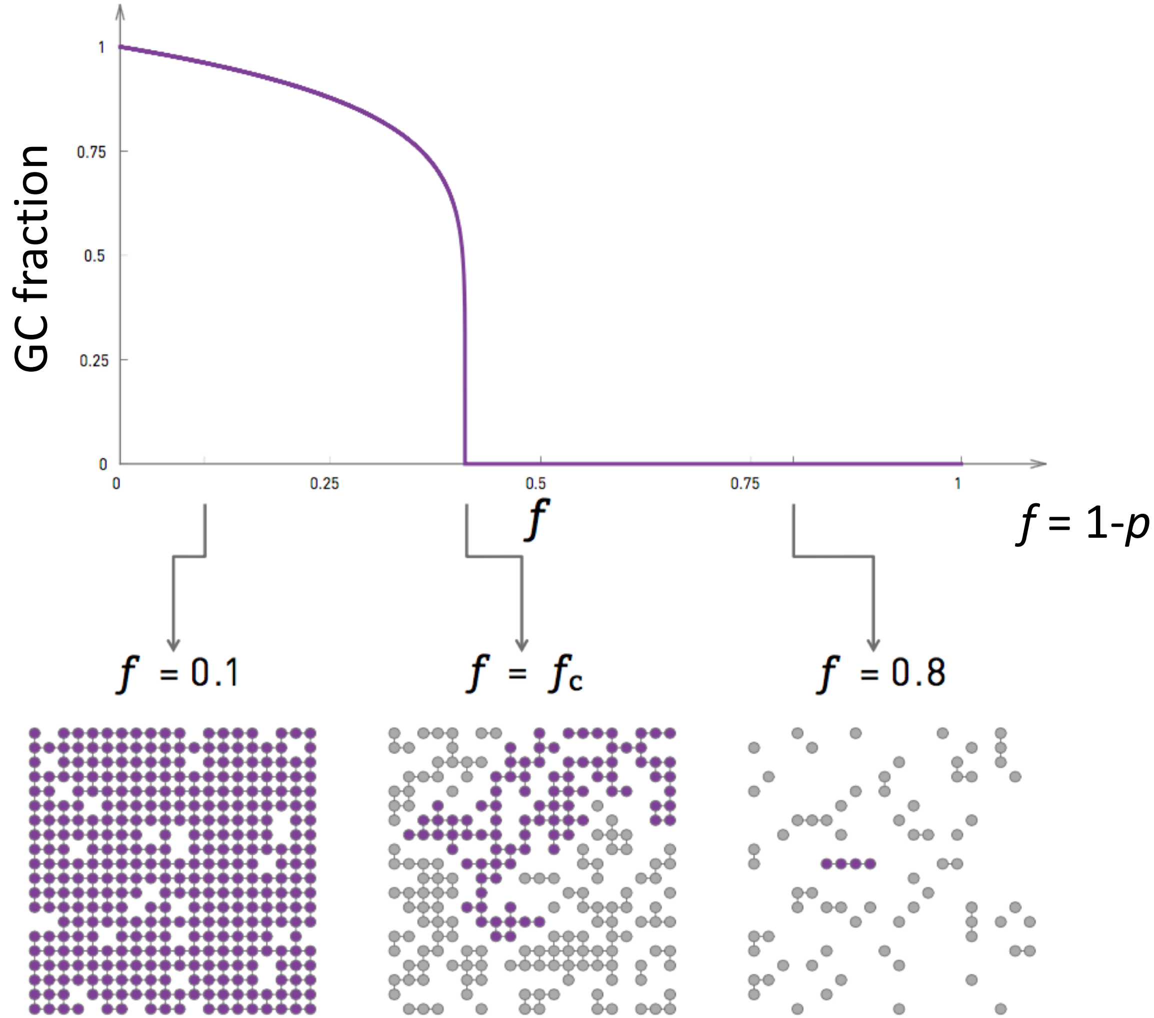
	γ_p	β_p	ν_p
2D lattice	43/18	5/36	4/3
3D lattice	1.80	0.41	0.88
7+D lattice	1	1	1/2

ν_p ← average distance inside clusters $\sim |p-p_c|^{-\nu_p}$

MIME.

Inverse percolation

Can be also interpreted as inverse percolation (node removal)



$0 < f < f_c$:

There is a giant component.

$P_\infty \sim |f - f_c|^\beta$

$f = f_c$:

The giant component vanishes.

$f > f_c$:

The lattice breaks into many tiny components.

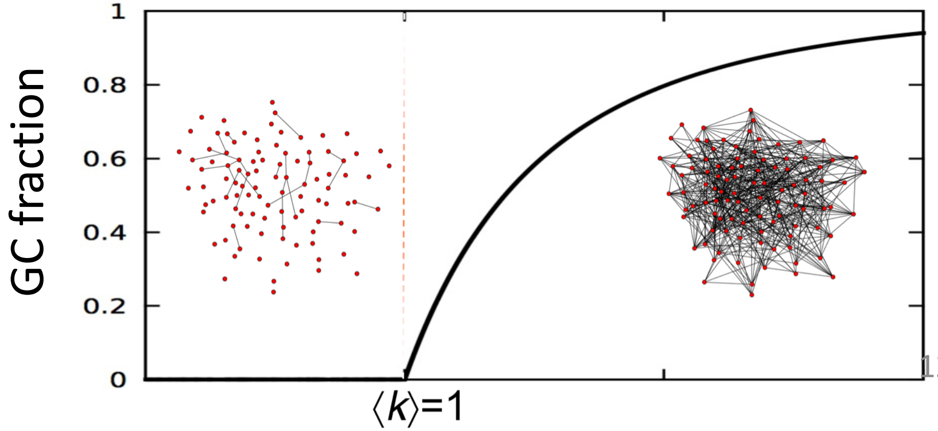
MIME.

Molloy-Reed Criterion

Molloy-Reed criterion

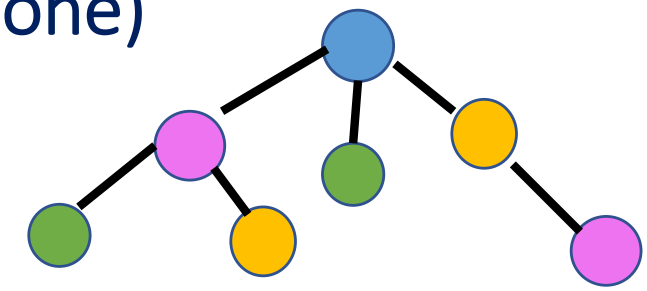
- The inhomogeneity ratio is $\kappa = \langle k^2 \rangle / \langle k \rangle$
- A randomly wired network has a giant component if $\kappa > 2$ (this identifies a breaking point)
- Networks with $\kappa < 2$ lack a giant component

E.g.: in random networks $\langle k^2 \rangle = \sigma^2 + \langle k \rangle^2 = \langle k \rangle (1+\langle k \rangle)$ so $\kappa = 1+\langle k \rangle > 2$ for $\langle k \rangle > 1$



Proof

- To hold a GC together at least 2 links are needed per node (it formally is 2 -2/|GC| for tree topology, the minimum one)



N = 7
L = 6

- That node *i* belongs to the GC can be derived recursively by asking $i \rightarrow j$ with *j* in the GC, ad we write $i \rightarrow j_{GC}$
- The average degree of the GC is $\langle k_i | \; i \rightarrow j_{GC} \rangle > 2$

… we then work on $\langle k_i | \; i \rightarrow j_{GC} \rangle$

Proof (cont'd)

- $\langle k_i |\ i \rightarrow j_{GC} \rangle = \sum_{k_i} k_i \, \mathrm{P}(k_i |\ i \rightarrow j_{GC})$
- $\mathrm{P}(k_i |\ i \rightarrow j_{GC}) = \mathrm{P}(i \rightarrow j_{GC} | k_i)\, \mathrm{P}(k_i) \,/\, \mathrm{P}(i \rightarrow j_{GC})$ by Bayes' rule
- $\mathrm{P}(i \rightarrow j_{GC} | k_i) = k_i\ (G\text{-}1)/(N\text{-}1)$ since there are $(G\text{-}1)/(N\text{-}1)$ random chances to connect to the GC, an k_i trials

reliable approximation for low G, i.e., close to the breaking point

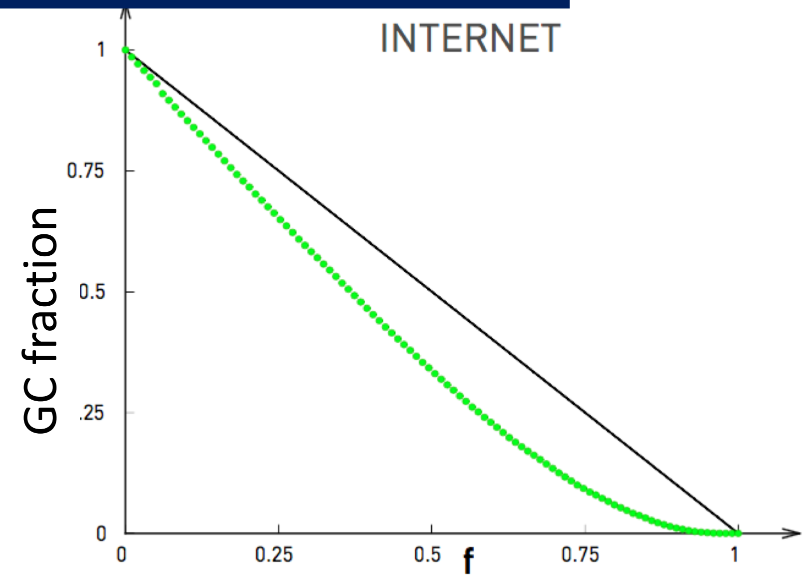
- Then $\mathrm{P}(i \rightarrow j_{GC}) = \langle k \rangle\ (G\text{-}1)/(N\text{-}1)$
- Hence $\langle k_i |\ i \rightarrow j_{GC} \rangle = \sum_{k_i} k_i^2 \, \mathrm{P}(k_i) \,/\, \langle k \rangle = \langle k^2 \rangle \,/\, \langle k \rangle$

MIME.

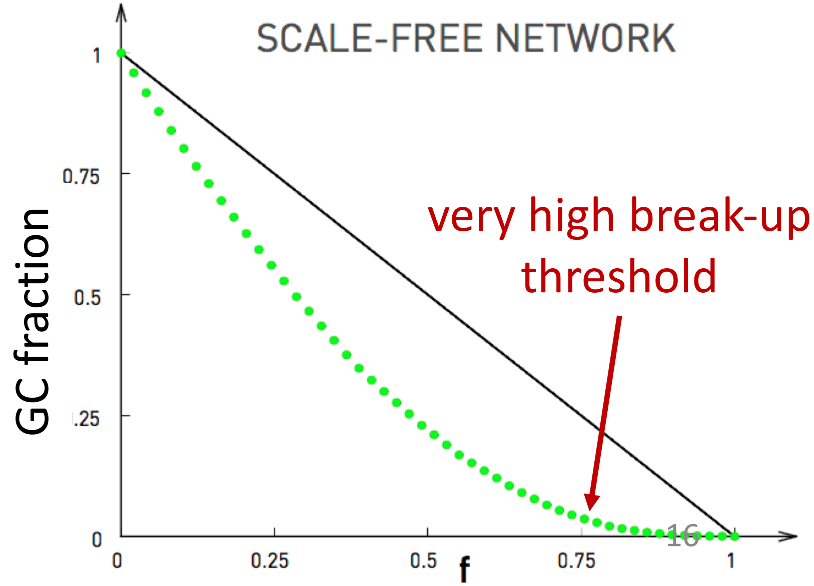
Robustness of scale-free networks

Robustness of scale-free nets

- Robustness of the Internet due to scale-free properties
- Nodes linked to the GC after random removal with rate *f* → still large if *f*<1



- Experiments aligned with a scale-free model
- Reason: random removal of (many) hubs is very unlikely



MIME

Inhomogeneity ratio under removal

- Assume a network with arbitrary degree distribution p_k and node removal at rate f

- It is $\langle k\rangle_f = (1\text{-}f)\,\langle k\rangle$

  and $\langle k^2\rangle_f = (1\text{-}f)^2\,\langle k^2\rangle + f(1\text{-}f)\langle k\rangle$

- Hence the inhomogeneity ratio $\boldsymbol{\kappa}_f = f + (1\text{-}f)\,\boldsymbol{\kappa}$

Sketch of the proof

- Probability that a node of degree k turns into a node of degree m

f can be fraction of deleted nodes/links

$$P(k \rightarrow m) = \text{binom}(k,m)\, f^{\,k-m}\, (1-f)^m$$

counting the # of cases

k-m links are deleted

m links are kept

- Then $P(m) = \sum_{k \geq m} P(k \rightarrow m)\, p_k$
- $\langle k \rangle_f = \sum_m m\, P(m)$
- $\langle k^2 \rangle_f = \sum_m m^2\, P(m)$ where $m^2 = m(m-1) + m$

the trick is to swap the order of sums

… then just replace and do boring substitutions

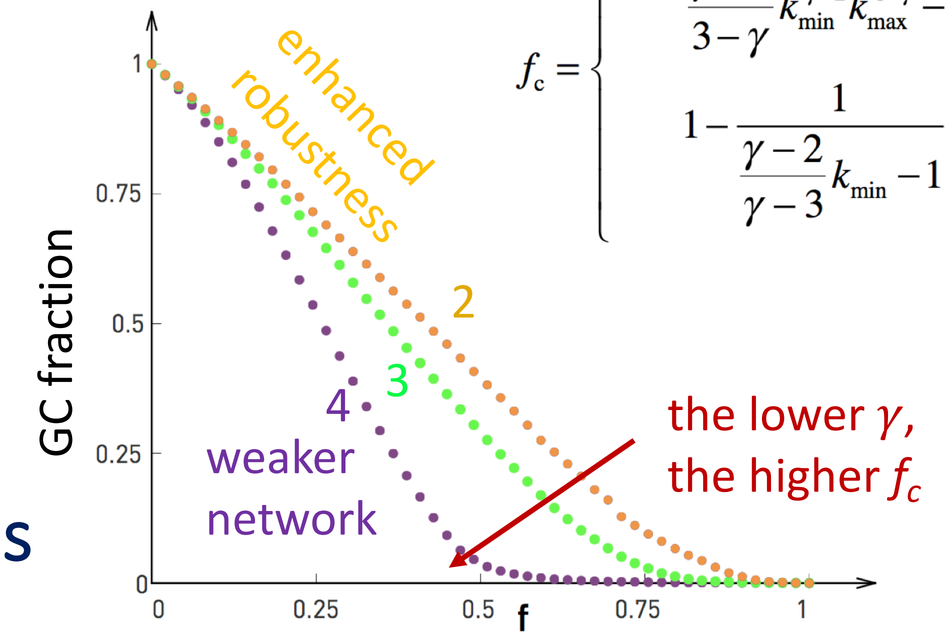
Breaking point

- Assume a network with arbitrary degree distribution p_k and node removal at rate f
- The breaking point

is found

at $\kappa_f = f + (1-f)\,\kappa = 2$

$$f_c = \begin{cases} 1 - \dfrac{1}{\dfrac{\gamma-2}{3-\gamma} k_{\min}^{\gamma-2} k_{\max}^{3-\gamma} - 1} & 2 < \gamma < 3 \\[2ex] 1 - \dfrac{1}{\dfrac{\gamma-2}{\gamma-3} k_{\min} - 1} & \gamma > 3 \end{cases}$$



- The breaking point is

$f_c = 1 - 1/(\kappa-1)$ which

solely depends on the degree distribution

MIME.

Some implications

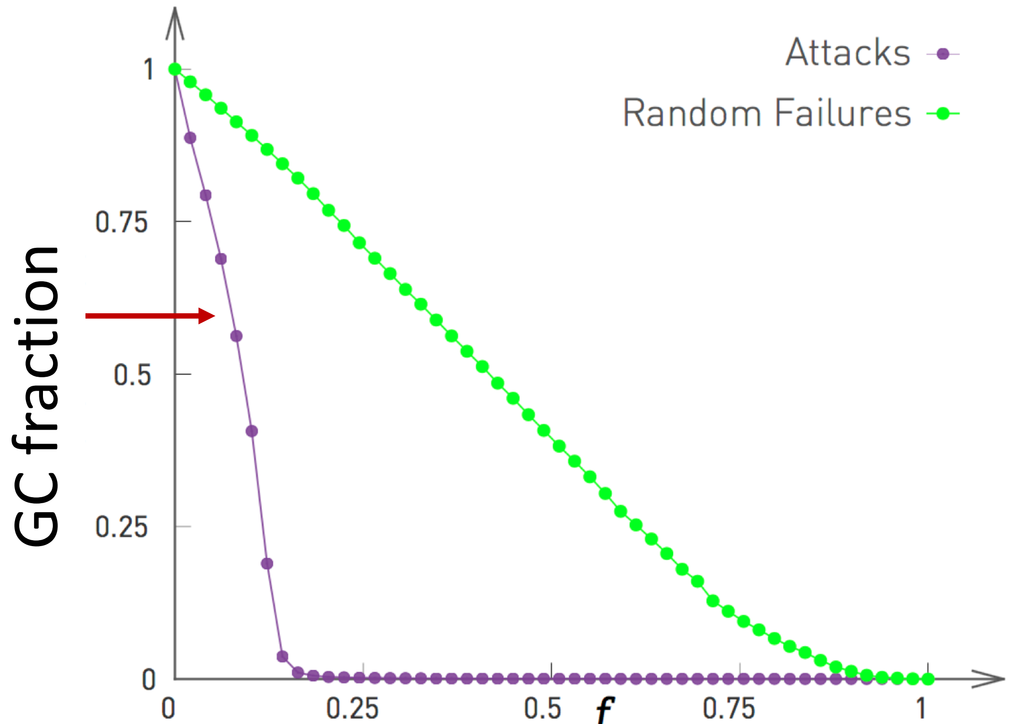
- networks with big hubs (causing wide deviations from $\langle k \rangle$) are hard to die
- in random networks $f_c = 1 - 1/\langle k \rangle$, i.e., large average degrees strengthen the network
- in scale-free networks the exponent γ sets the network robustness

Attacks

Attack tolerance

- What if removals are not by chance, but caused by an adversary with sufficient insights on our network?

an adversary would remove all hubs first, i.e., it removes nodes in decreasing order of their degree



probability/percentage of removed nodes

Fragility of scale-free nets

- Scale-free networks are not very robust to targeted attacks exactly because they have vulnerable hubs
- Recall that $f_c = 1 - 1/(\kappa - 1)$ meaning that robustness depends on κ, and removing hubs reduces κ

  - good news in medicine (vulnerability of bacteria) ☺
  - bad news for the Internet ☹

MIME.

Breaking point in scale-free nets

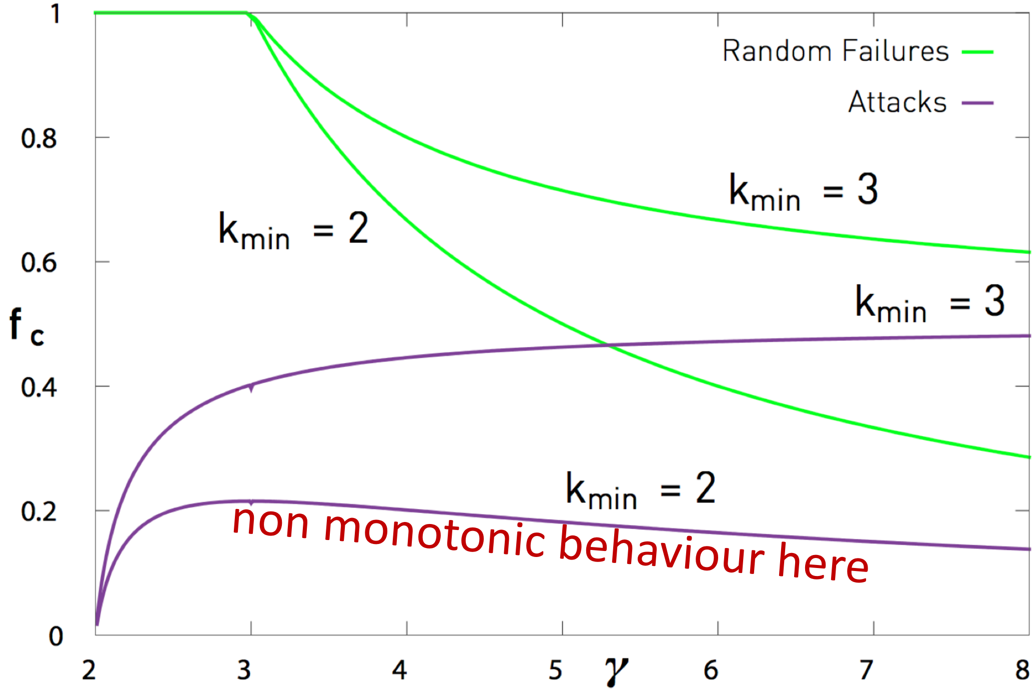
estimated value

NETWORK	RANDOM FAILURES (REAL NETWORK)	RANDOM FAILURES (RANDOMIZED NETWORK)	ATTACK (REAL NETWORK)
Internet	0.92	0.84	0.16
WWW	0.88	0.85	0.12
Power Grid	0.61	0.63	0.20
Mobile-Phone Call	0.78	0.68	0.20
Email	0.92	0.69	0.04
Science Collaboration	0.92	0.88	0.27
Actor Network	0.98	0.99	0.55
Citation Network	0.96	0.95	0.76
E. Coli Metabolism	0.96	0.90	0.49
Yeast Protein Interactions	0.88	0.66	0.06

Not robust to random failures (exponential degree distribution)

Fragility of scale-free nets

$$f_c^{\frac{2-\gamma}{1-\gamma}} = 2 + \frac{2-\gamma}{3-\gamma} k_{min} \left(f_c^{\frac{3-\gamma}{1-\gamma}} - 1 \right)$$



random failures & attacks have similar behaviour here

MIME

Analysis of an attack

An attack reduces $k_{max} \rightarrow k'_{max}$

- Degree distribution

$$p_k = C\,k^{-\gamma}, \qquad C = (\gamma - 1) / (k_{min}^{1-\gamma} - k_{max}^{1-\gamma})$$

- Percentage of removed nodes is

$$f = \int_{k'_{max}}^{k_{max}} p_k \, dk = C/(\gamma - 1)\,(k'^{\,1-\gamma}_{max} - k_{max}^{1-\gamma})$$

- Hence $k'_{max} = k_{min}\, f^{-1/(\gamma-1)}$

MIME.

Fraction of removed links

- The fraction a of removed links is $a = b / \langle k \rangle$ where

$$b = \int_{k'_{max}}^{k_{max}} k\, p_k\, dk = C/(\gamma - 2)\, (k'^{\,2-\gamma}_{max} - \cancel{k_{max}^{2-\gamma}})$$

$$\langle k \rangle = \int_{k_{min}}^{k_{max}} k\, p_k\, dk = C/(\gamma - 2)\, (k_{min}^{2-\gamma} - \cancel{k_{max}^{2-\gamma}})$$

- Hence $a = (k'_{max}/k_{min})^{2-\gamma} = \boxed{f^{(\gamma-2)/(\gamma-1)}}$

Final proof

An attack distorts the degree distribution $p_k \rightarrow p'_k$

- We assume that links were randomly assigned, so that if a is the fraction of removed links, then

$$p'_m = \sum_{k=m}^{k'_{max}} \underbrace{\text{binomial}(k,m)\, a^{k-m}\, (1-a)^m}_{\text{transition probability } P(k \rightarrow m)}\, p_k$$

- As a consequence

$$\kappa_f = a + (1-a)\, \kappa'$$

$$\kappa' = k_{min}\, (\gamma-2)/(\gamma-3)\, (f^{(\gamma-3)/(\gamma-1)} - 1)/(f^{(\gamma-2)/(\gamma-1)} - 1)$$

Same as before but $f \rightarrow a = f^{(\gamma-2)/(\gamma-1)}$
and $k_{max} = k_{min}\, N^{1/(\gamma-1)} \rightarrow k'_{max} = k_{min}\, f^{-1/(\gamma-1)}$

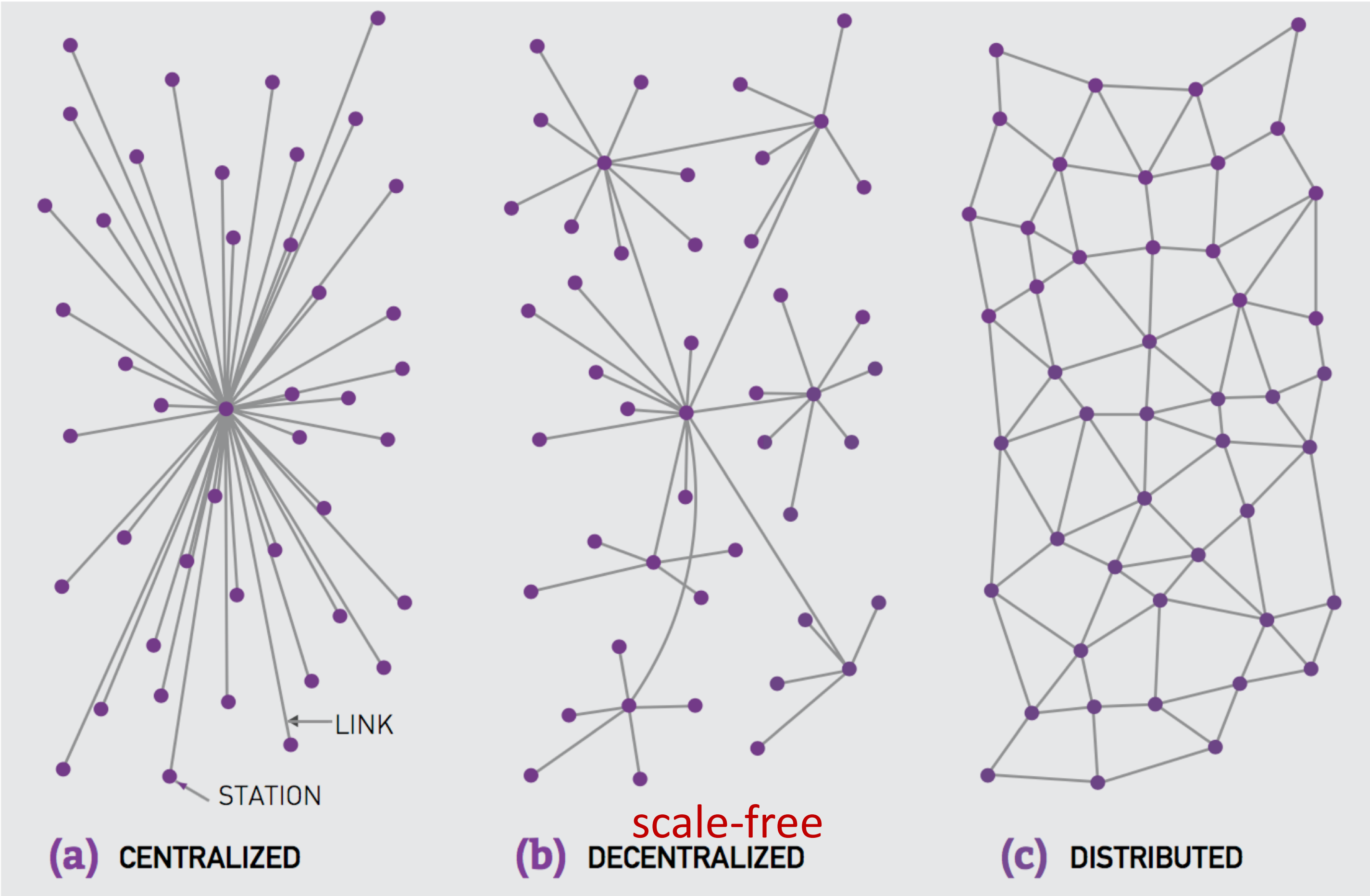
- Set $\kappa_f = 2$ to obtain the equation

MiME.

Optimizing robustness

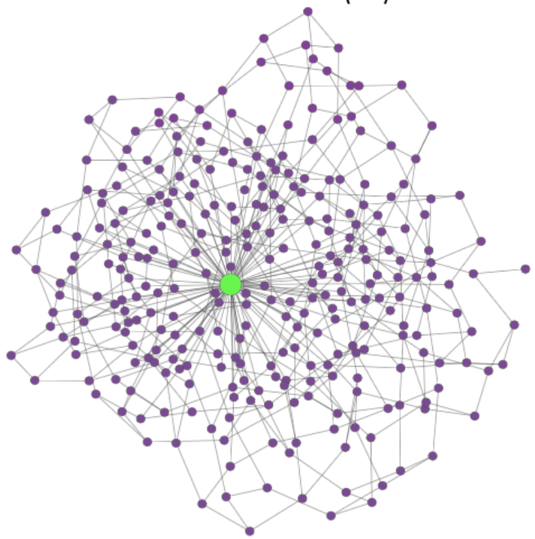
Optimizing robustness

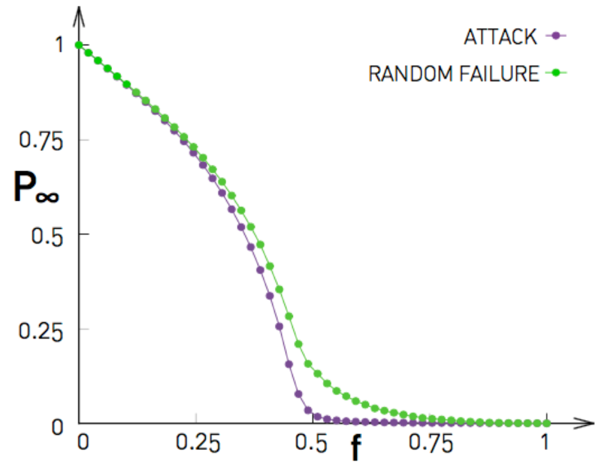
An early attempt by Paul Baran [1959]



Optimizing robustness

$\langle k \rangle = 3$

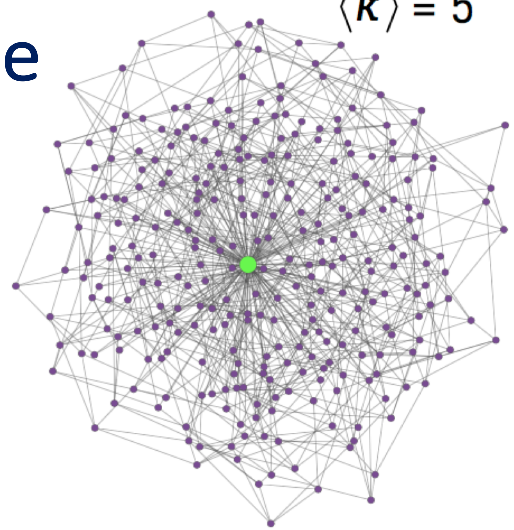




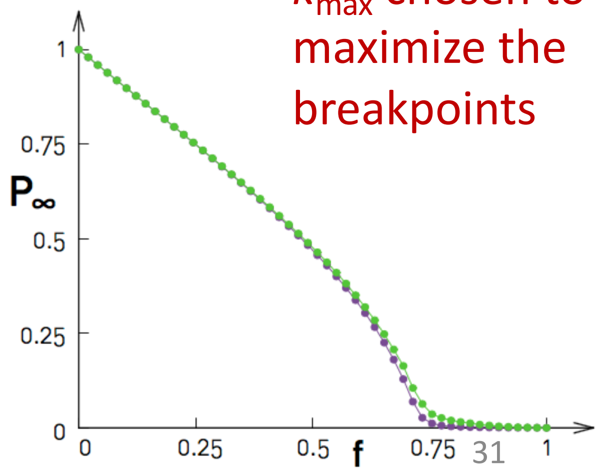
The best option is a bimodal distribution

$p_k = r\,\delta_{k_{max}} + (1\text{-}r)\,\delta_{k_{min}}$

$\langle k \rangle = 5$



r = 1/N
k_{max} chosen to maximize the breakpoints



But not always possible to implement it in practice (mainly for cost reasons)

Analysis

Random failure - Assume $p_k = r\, \delta_{k\max} + (1\text{-}r)\, \delta_{k\min}$

- Average degree $\langle k \rangle = r\, k_{max} + (1\text{-}r)\, k_{min}$
- Inhomogeneity ratio $\kappa = (r\, k_{max}{}^2 + (1\text{-}r)\, k_{min}{}^2)/\langle k \rangle$
- Breakpoint $f_c = 1\text{-}\ 1/(\kappa\text{-}1)$

Attack – Assume that all hubs are removed ($f > r$) and that only nodes of degree k_{min} are surviving

- Fraction of removed links $a = 1 - k_{min}(1\text{-}f)/\langle k \rangle$
- $\kappa_f = a + (1\text{-}a)\, \kappa'$ with $\kappa' = k_{min}$
- Breakpoint $f_c = 1\text{-}\ \langle k \rangle/(k_{min}(k_{min} -1))$

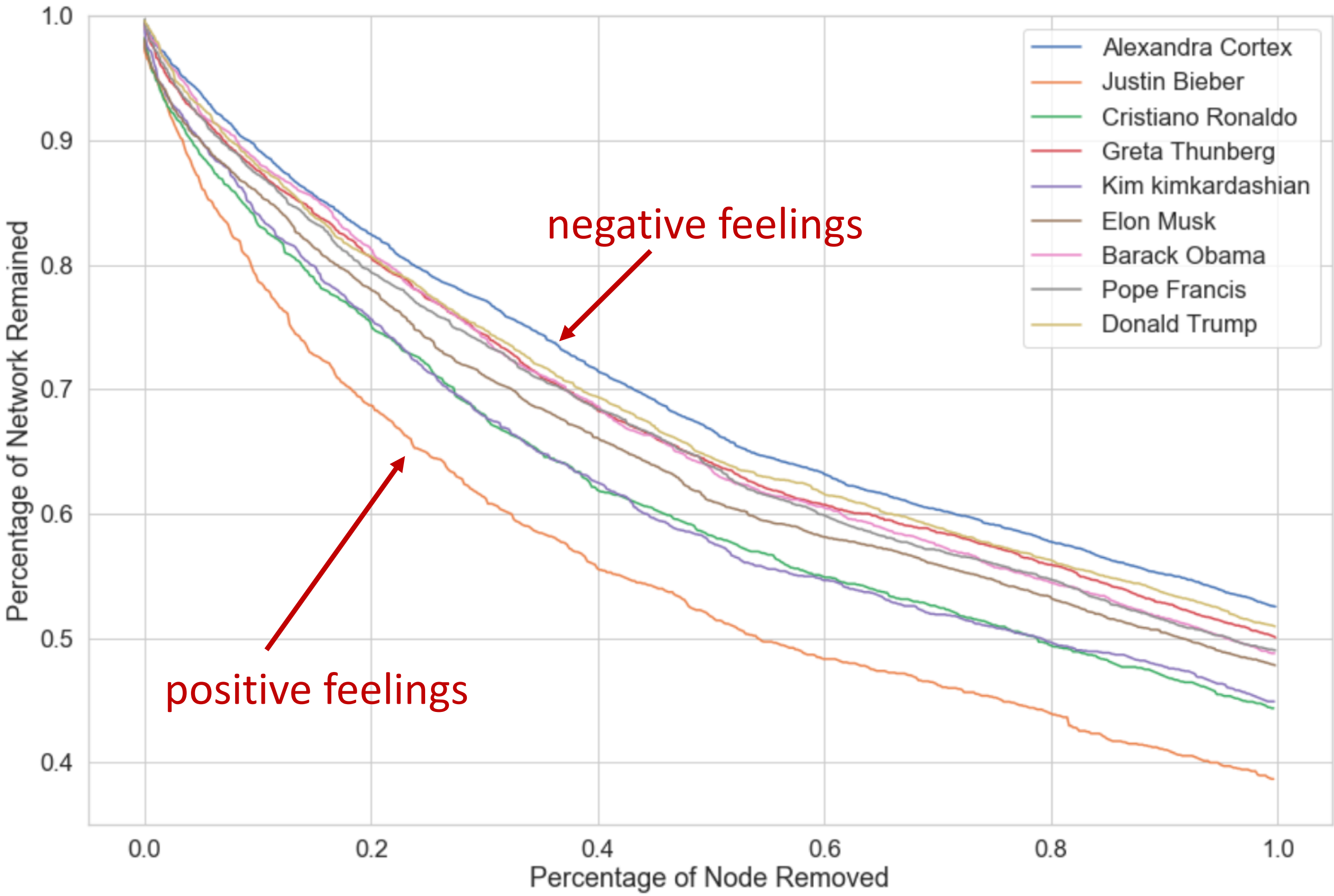
MiME.

Application example

Network analysis of Tweets' sentiment

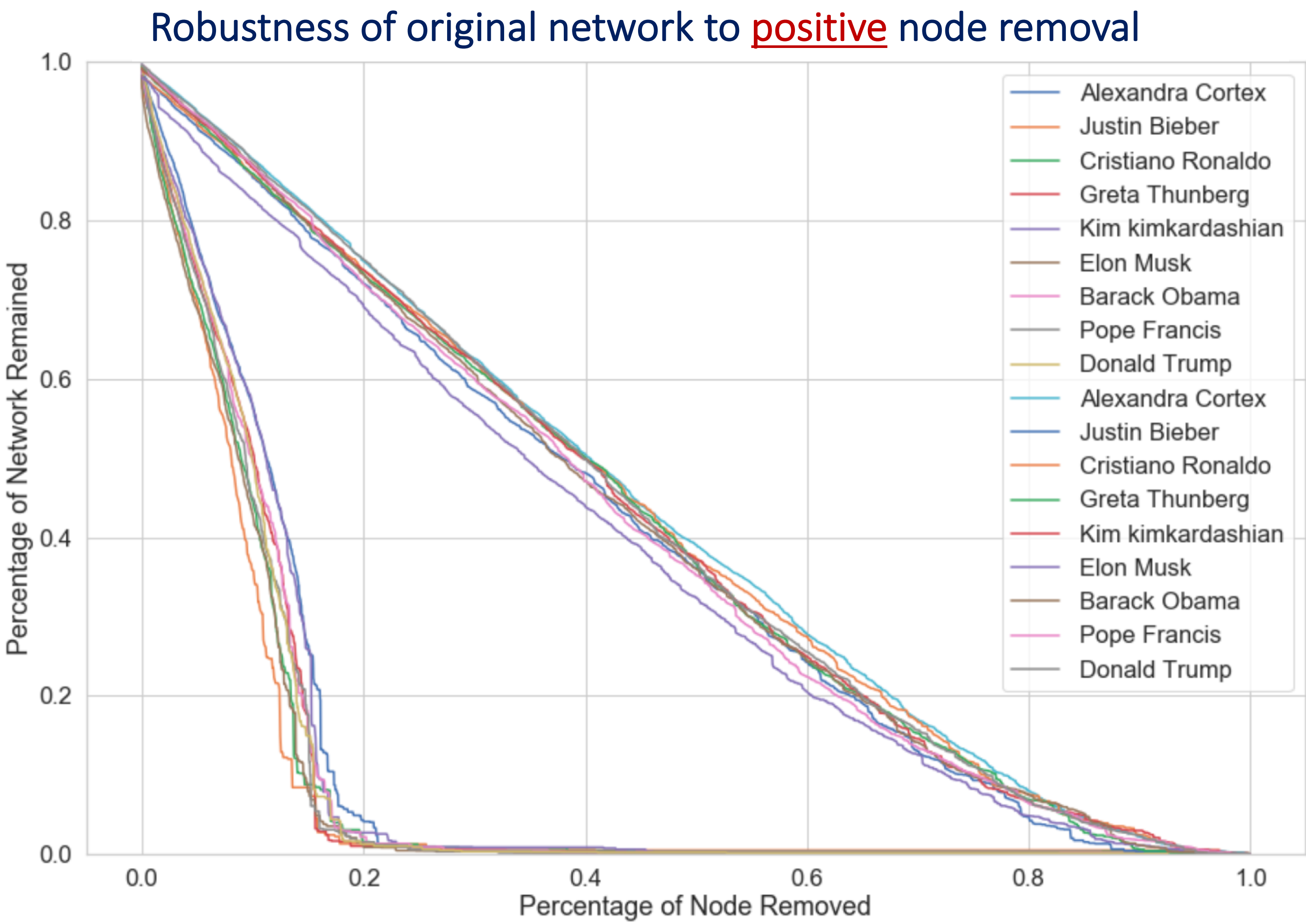
Salvatore Romano, Alberto Zancanaro, Enrico Lanza, Carlo Facchin

Robustness of original network to positive node removal



MiME.

Network analysis of Tweets’ sentiment



Questions ?