Social and Graph Data Management
Network Robustness

Benoît Groz (slides and content are mostly from Silviu Maniu)

November 17th, 2024

M2 Data Science

Table of contents

Robustness

Robustness is a central issue in network science.

What happens to a network if some parts of it are *removed*?

- mutations in medicine
- network attack in online social networks
- diseases, famines, wars, ...

Robustness

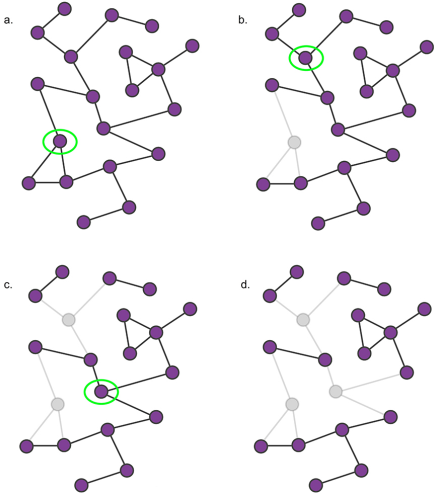
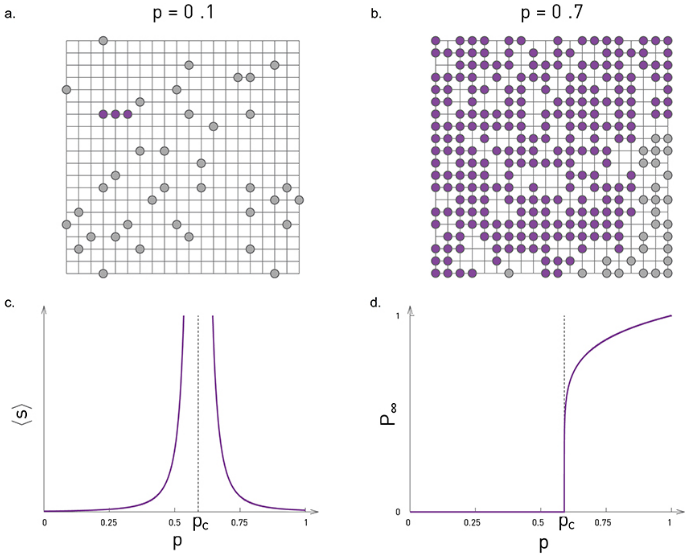


Table of contents

Percolation: term coming from statistical physics, applied in our case: what is the *expected size of the largest cluster* and *the average cluster size*

Example: a square lattice, where "pebbles" are placed with probability p at random intersections. If two or more pebbles are connected they form clusters. As p approaches a critical value p_c, a large cluster emerges.

Percolation in Lattices



Percolation in Lattices

We track:

- **average cluster size** $\langle s \rangle \sim |p - p_c|^{-\gamma_p}$ – diverges as we approach p_c
- **order parameter** $p_\infty \sim (p - p_c)^{\beta_p}$ – probability that a pebble belongs to the largest cluster
- **correlation length** $\xi \sim |p - p_c|^{-\nu}$ – mean distance between two pebbles belonging to the same cluster

γ_p, β_p, and ν are **critical exponents** – they characterize the behavior near the critical point

Percolation theory says that the exponents are **universal**: independent of p_c or the nature of the lattice.

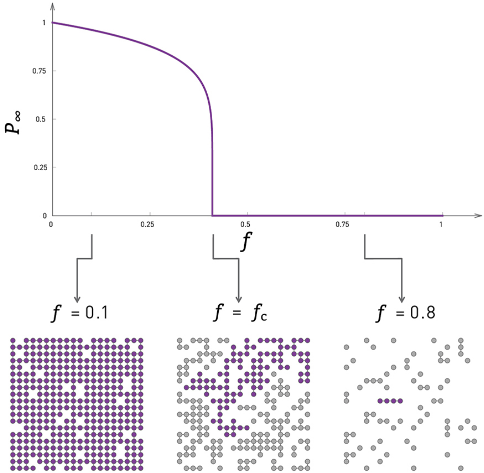
Percolation and Robustness

Inverse percolation: what happens when we remove a fraction f of nodes from the giant component of the lattice

As f increases, the lattice is more and more likely to break up in tiny components

However, the process is **not gradual**! It is characterized by a **critical threshold** f_c at which point the lattice is broken.

Inverse Percolation in Lattices



$0 < f < f_c$:

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.

Percolation and Networks

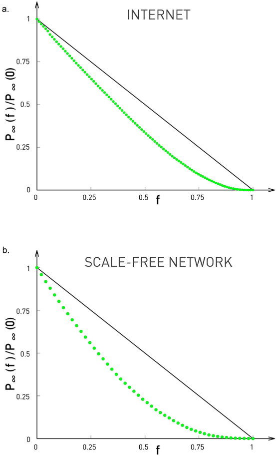
Random networks under random node failures have the same exponents as the infinite-dimensional percolation.

The critical exponents in random networks are $\gamma_p = 1$, $\beta_p = 1$ and $\nu = 1/2$.

Table of contents

What happens to scale-free networks under random removals?

Empirical results show that they are surprisingly resilient. Why?



Molloy-Reed Criterion

f_c in scale free networks is extremely high.

Molloy-Reed criterion: a randomly wired network has a giant component if:

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle} > 2; \tag{1}$$

this works for any degree distribution p_k.

For a random network:

$$\kappa = \frac{\langle k \rangle (1 + \langle k \rangle)}{\langle k \rangle} = 1 + \langle k \rangle > 2,$$

or

$$\langle k \rangle > 1.$$

Applying Molloy-Reed in Random Networks

We can apply the criterion to a network with arbitrary degree we have that:

$$f_c = 1 - \frac{1}{\kappa - 1}; \tag{2}$$

depending **only** on $\langle k \rangle$ and $\langle k^2 \rangle$.

In a **random network**:

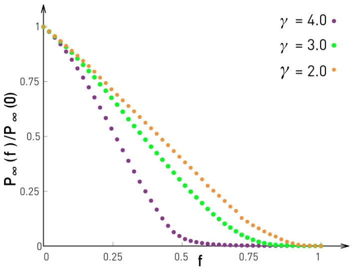
$$f_c = 1 - \frac{1}{\langle k \rangle}.$$

We only need to remove a **finite number of nodes**, and f_c is higher as the network is **denser**

Applying Molloy-Reed in Scale-Free Networks

In scale-free networks, f_c depends on the degree exponent γ:

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



Robustness in Scale-Free Networks

For $\gamma < 3$, $f_c \to 1$, meaning that we have to remove almost all nodes in order that the network breaks.

Main takeaway: scale-free networks are resilient under random removals, we can remove an arbitrary number of nodes.

Table of contents

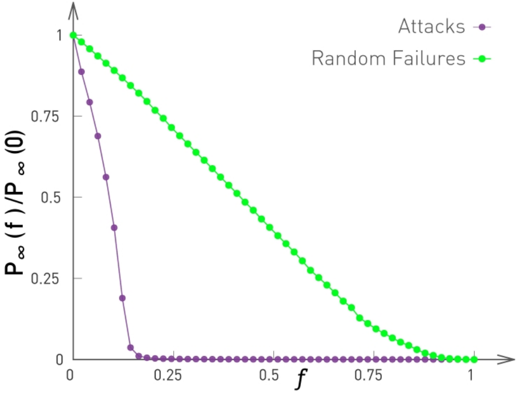
f_c under Attacks

What happens when we attack the network (we choose deliberately the nodes, prioritizing *high degree nodes*?

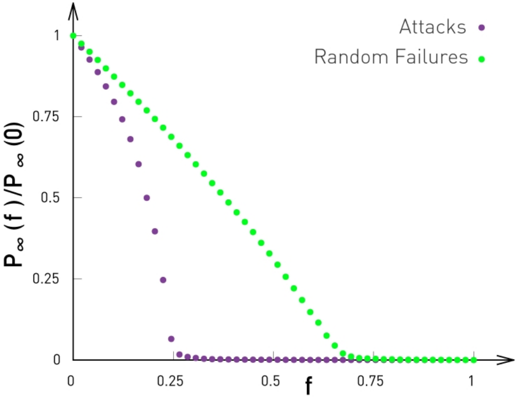
How does f_c change?

Network	**Random**(pred.)	**Random**(real)	**Attack**
Internet	0.84	0.92	0.16
Power Grid	0.63	0.61	0.20
Email	0.69	0.92	0.04
Protein	0.66	0.88	0.06

Attacks: Scale-Free Networks



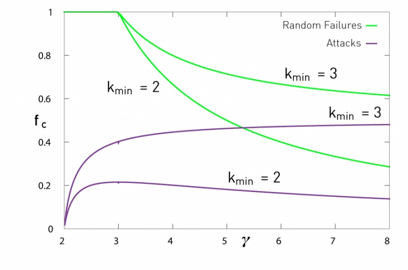
Attacks: Random Networks



Critical Threshold Under Attack

Using the fact that, for large γ the scale-free networks resemble random networks, so random failures and targeted attacks are indistinguishable when $\gamma \to \infty$:

$$f_c \to 1 - \frac{1}{k_{\min} - 1}. \tag{3}$$



Cascading Failures

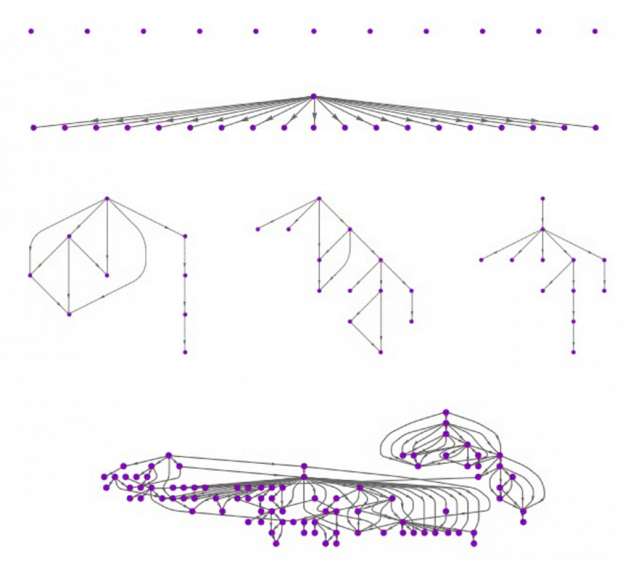
Once an attack is perpetrated, some failures are **cascading**: the neighbours of the attacked node can fail, which triggers cascades on their neighbours etc.

Examples of cascading failures:

- **blackouts** on power grids
- **denial of service attacks**
- **information cascades** in social networks, viruses
- **financial crises**

Common characteristic: all the cascading failure follow **power laws**.

Information Cascades



Acknowledgments

Figures in slides 4, 7, 10, 13, 16, 20, 21, 22, and 24 taken from the book "Network Science" by A.-L. Barabási. The contents is partly inspired by the flow of Chapter 8 of the same book.

`http://barabasi.com/networksciencebook/`

References i

Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. (2011).
Everyone's an influencer: Quantifying influence on twitter.
In *Proceedings of the ACM International Conference on Web Search and Data Mining*, pages 65–74.

Bollobás, B. and Riordan, O. (2003).
Robustness and vulnerability of scale-free random graphs.
Internet Math., 1(1):1–35.

Callaway, D. S., Newman, M., Strogatz, S. H., and Watts, D. J. (2000).
Network robustness and fragility: Percolation on random graphs.
85(25):5468–5471.

References ii

Molloy, M. and Reed, B. (1995).
A critical point for random graphs with a given degree sequence.
Random Struct. Algorithms, 6(2-3):161–180.