Data literacy

Dr. Nino Antulov-Fantulin[1]

[1]Computational Social Science, ETH Zurich, Switzerland

December 3, 2019

ETH zürich

Network dismantling – case study 1

Network Dismantling

NEWS & VIEWS

For News & Views online, go to nature.com/newsandviews

NETWORK SCIENCE

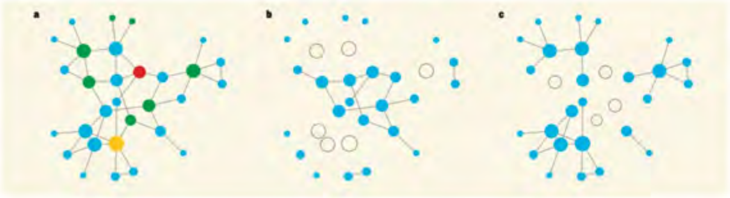
Destruction perfected

Pinpointing the nodes whose removal most effectively disrupts a network has become a lot easier with the development of an efficient algorithm. Potential applications might include cybersecurity and disease control. SEE LETTER P.65

ISTVÁN A. KOVÁCS & ALBERT-LÁSZLÓ BARABÁSI

which deletion would cause maximum damage is a non-deterministic polynomial-time hard

(known as influencers) and recalculating the collective influence of the rest following each



2. Cost of attack

Current assumption*:
cost of removing any node in a network is the same.

* Braunstein A, Dall'Asta L, Semerjian G, Zdeborová L (2016) Network dismantling. Proceedings of the National Academy of Sciences 113(44):12368–12373.
Morone F, Makse HA (2015) Influence maximization in complex networks through optimal percolation. Nature 524(7563):65–68
Kovács IA, Barabási AL (2015) Network science: Destruction perfected. Nature 524(7563):38–39.
Zdeborová L, Zhang P, Zhou HJ (2016) Fast and simple decycling and dismantling of networks. Scientific Reports 6.
Mugisha S, Zhou HJ (2016) Identifying optimal targets of network attack by belief propagation. Phys. Rev. E 94(1):012305.
Morone F, Min B, Bo L, Mari R, Makse HA (2016) Collective influence algorithm to find influencers via optimal percolation in massively large social media. Scientific Reports 6(1). Tian L, Bashan A, Shi DN, Liu YY (2017) Articulation points in complex networks. Nature Communications 8:14223.
Zhou HJ (2013) Spin glass approach to the feedback vertex set problem. The European Physical Journal B 86(11):455.

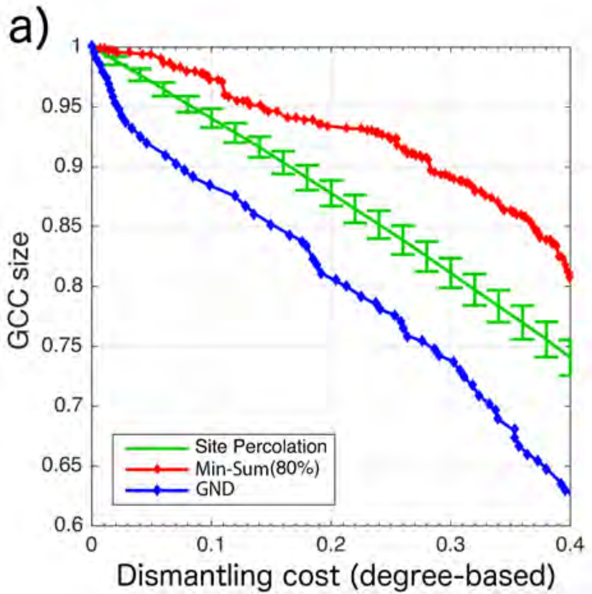
2. Cost of attack

Our assumption**:

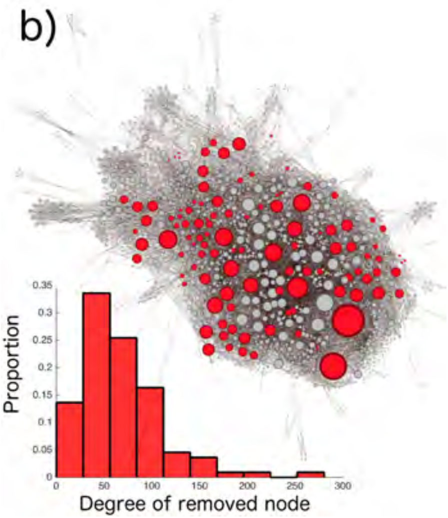
cost of removing any node in a network is **not the same**. It can be arbitrary non-negative weight e.g. degree.

**Weights taken into the account:
Patron A, Cohen R, Li D, Havlin S (2017) Optimal cost for strengthening or destroying a given network. Phys. Rev. E 95(5):052305.
Ren XL, Gleinig N, Tolic D, Antulov-Fantulin N (2017) Underestimated cost of targeted attacks on complex networks, arxiv:1710.03522

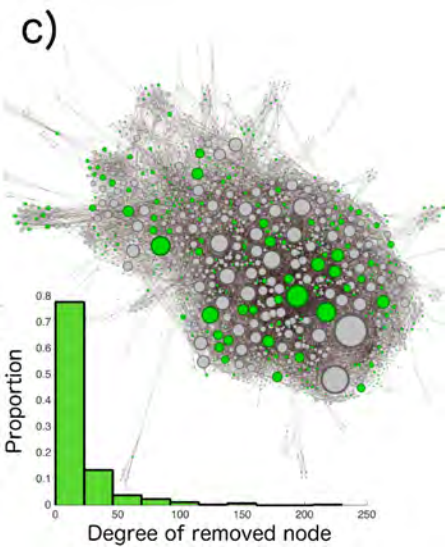
Example 1: Online social network



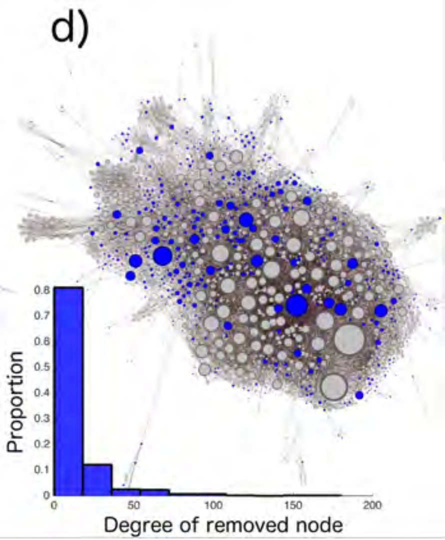
Example 1: Online social network



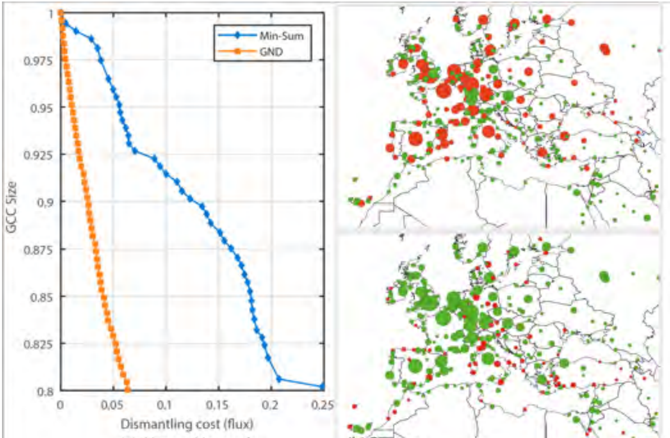
Example 1: Online social network



Example 1: Online social network



Airports



(a) Airport Network (b) Europe

Publication

PNAS

Generalized network dismantling

Xiao-Long Ren[a,1], Niels Gleinig[b,1], Dirk Helbing[a], and Nino Antulov-Fantulin[a,2]

[a]Computational Social Science, ETH Zürich, 8092 Zürich, Switzerland; and [b]Department of Computer Science, ETH Zürich, 8092 Zürich, Switzerland

Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved February 15, 2019 (received for review April 12, 2018)

Finding an optimal subset of nodes in a network that is able to efficiently disrupt the functioning of a corrupt or criminal organization or contain an epidemic or the spread of misinformation is a highly relevant problem of network science. In this paper, we address the generalized network-dismantling problem, which

find good approximations of the optimal dismantling strategy. For example, novel approximations (17–23) have been proposed based on spin-glass and optimal percolation theory. However, all these methods make the implicit assumption that the cost of removing nodes is the same. Only recently have people been

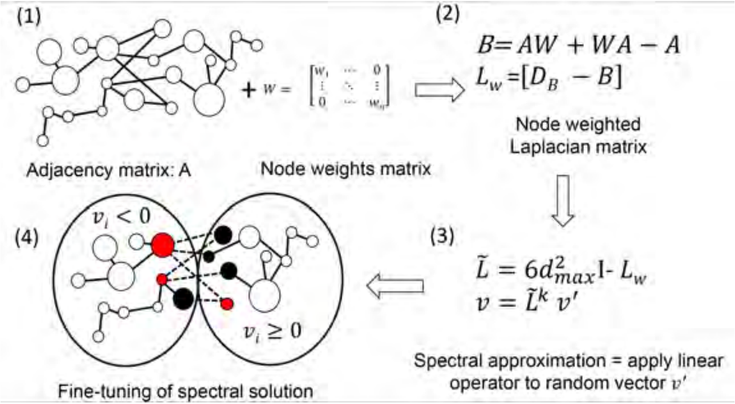
3. New approach – objective function

Then the upper bound for the cost of removing a subset of nodes that are adjacent to the edges separating clusters M and $\overline{M}$ is:

$$\frac{1}{2}\sum_{i,j} -\frac{1}{2}\left(v_i v_j - 1\right) A_{i,j}\left(w_i + w_j - 1\right), \qquad (1)$$

where the matrix A denotes the adjacency matrix of the network.

Overall



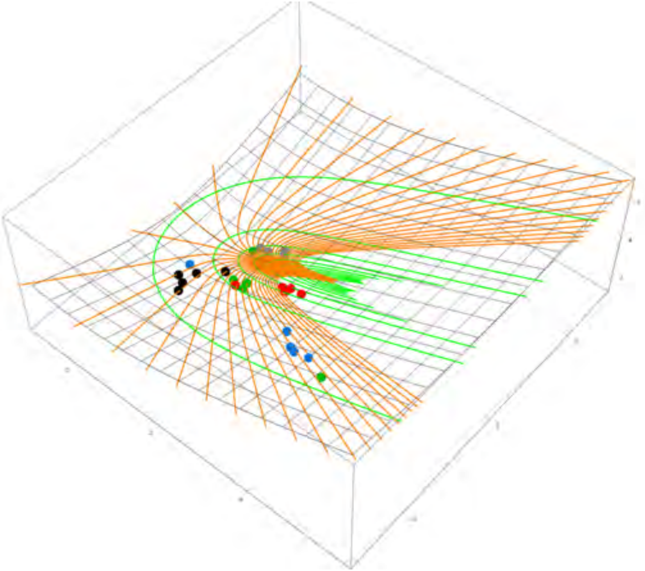
Visualize math!

Statistical manifold embedding

Circle Limit III by M. C. Escher. Wood engraving, 1959.

Statistical embedding

Statistical embedding



Dynamical process – Spreading process

In this talk, we will focus on the class of the dynamical process that are **compartmental**, **stochastic** and have **monotone** dynamics.
Generalized SIR process:

- $X \to Y \to Z$ (monotone dynamics)
- $X + Y \xrightarrow{\psi(\tau)} Y + Y$ (network modulated)
- $Y \xrightarrow{\phi(\tau)} Z$

For Poisson process, this process has exponential inter-event time distributions:
$\psi(t) = \beta e^{-\beta t}$ (spread)
$\phi(t) = \gamma e^{-\gamma t}$ (recovery).

Dynamical process – Spreading process

Again art – not to be boring!

Source Inference

For a given snapshot r^* of a spreading process at time t_0 on a network, how does one detect the source node that generated the snapshot?

Figure: Visualization of aggregated empirical temporal network (≈ 3500 nodes) of sexual contacts in Brazil, used for source detection (Phys. Rev. Lett. 114, 248701)

Dynamical process – Spreading process

https://www.youtube.com/watch?v=U32j1gd_cDw

https://psmag.com/environment/a-better-way-to-find-patient-zero

Every lecture I try to learn something new !

Do not fool yourself.