Using network organization to hinder propagation in structured populations

Laurent Hébert-Dufresne, Antoine Allard, Jean-Gabriel Young and Louis J. Dubé

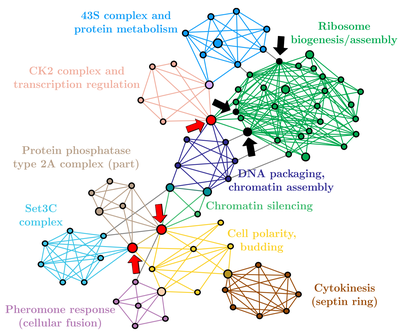
Département de Physique, de Génie Physique, et d'Optique, Université Laval, Québec, Canada

Objectives

Which nodes should be **immunized** in a network to minimize the spread on an epidemics? We seek a recipe to identify **influential spreaders** that should

- depend on a **local measure**, easily estimable in practice;
- work for all networks in all epidemics scenarios.

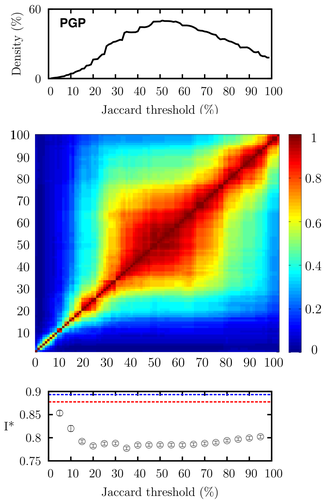
Why not simply remove the hubs?



Protein interactions of *S. cerevisiae* (subset). The three nodes in **black** correspond to the ones with the most links, and in **red** are the ones belonging to the most communities/complexes. The latter appears more structurally important.

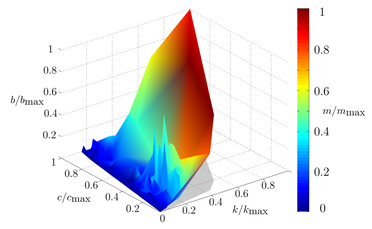
Case study: PGP network

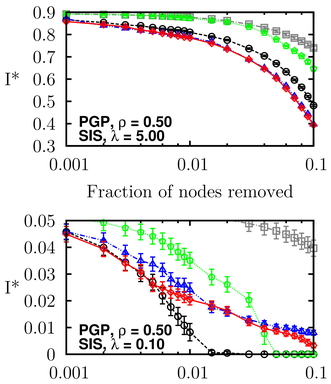
Network of users of the **Pretty-Good-Privacy** algorithm for data exchange (10680 nodes, 24316 links)



Results of SIS epidemics ($\lambda = 5$) with top 1% structural hubs removed for different Jaccard threshold (dots) compared with no removal (blue) and random removal (red).

- **Community Structure.** The community structure is extracted using a link community algorithm [Ahn *et al.*, Nature 2010] that groups links connecting similar neighbourhoods. Links are considered to be in a same community when the similarity of their respective neighbourhood (Jaccard coefficient) exceeds a certain threshold. This threshold acts as a resolution, enabling to look at different levels of organization. On the **left figure**: the average density of communities (ρ), the similarity between the structural hubs (top 1%) selected with different threshold, and the efficiency of their removal on the SIS dynamics. The identification of structural hubs is **robust around the maximal density** of communities. This coincides with the **maximal effect** on the epidemics.
- **Correlations.** The **top right figure**, a triangulation surface on the nodes in the space (k,b,c,m), illustrates how nodes of high membership numbers can be found even at relatively small degree, coreness and centrality.
- **Immunization scenarios.** In epidemic models, networks feature an epidemic threshold; i.e. a critical virulence of the disease (T_c or λ_c) below which it has a null chance of infecting a macroscopic fraction of the networks. Epidemics near or far from the threshold need to be treated differently. The **two bottom right figures** present the prevalence (infectious fraction I^*) for an SIS dynamics after the immunization of a fraction ϵ of the nodes through different targeting methods for a disease (first) well above epidemic threshold and (second) near the threshold. Color code of targeting methods: **degree**, **betweenness centrality**, **memberships**, **coreness** and **random**.





Discussion and conclusion

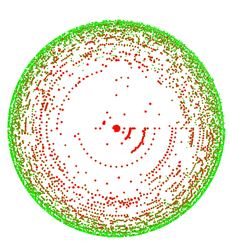
Importance of memberships in virulent epidemics.

A virulent disease will travel most links of the networks. A modular structure with dense communities implies that the disease will often follow links leading to nodes already visited/infected instead of reaching new susceptible nodes. Hence in this case the important nodes are not those with the most links (**hubs**), but those connecting the most different neighbourhoods (**structural hubs**).

Importance of degree near the epidemic threshold.

For a small infection probability ($T \sim T_c$ and $\lambda \sim \lambda_c$), the path used by the disease is much less likely to be affected by the clustering of the network and will instead follow an almost tree-like structure. It is therefore better to simply remove as much links as possible. Note that the efficiency of immunization based on centrality is the most sensitive to this change in the virulence.

Ineffectiveness of coreness: danger versus inflence.



This figure (k-core of PGP under SIS with $\lambda = 5$) shows that core nodes (high c, center) are most at risk of being infected (red) because of the core's edge density. This density implies redundancy. Hence, the core nodes are highly at risk of being infected, but their removal has little effects because of alternative paths within their neighbourhood.

Practicality of local, coarse-grained measures.

The exact network structure of a whole system is rarely available: it is too large and/or it evolves too quickly and/or links are ill-defined. Thus a *local and coarse-grained measure like the number of memberships offers a practical and efficient alternative.* Consider how easier it is to enumerate your social groups than the totality of your acquaintances.

Methods

The effectiveness of each measure of a node influence is quantified through immunization scenarios. How much will the removal of a fraction ϵ of nodes, targetted by this measure, affects epidemics?

Local measures

Degree (k). Nnumber of links attached to a node. The highest degree nodes are the **hubs** of the network.

Community memberships (m). Number of communities (the colored groups in the figure above) to which a node belongs. The nodes with the highest m are the **structural hubs**.

Global measures

Betweenness centrality (b). Sum of a node's contributions $1/\ell$ to all shortest paths of length ℓ between all pairs of nodes.

Coreness (c). Highest integer c such that the node is part of the set of all nodes with at least c links shared within the set.

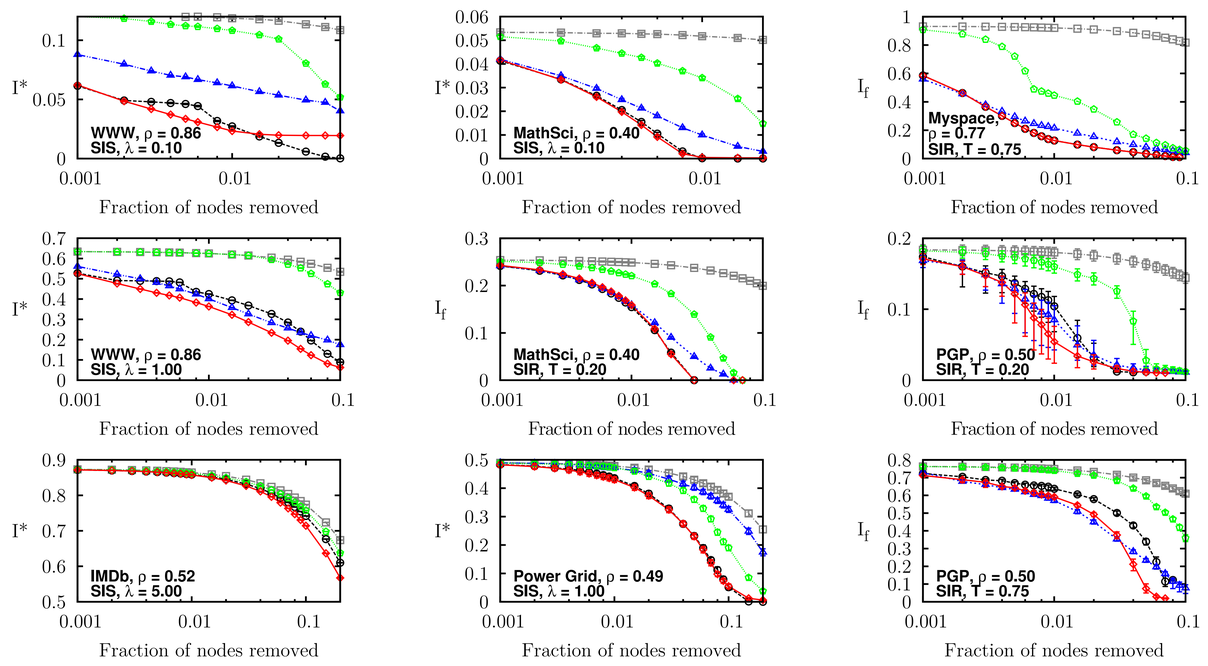
Epidemic models

Susceptible-Infectious-Removed (SIR). An infections node transmits the disease to its susceptible neighbours with probability T, and ends in a recovered state (through death or immunity). We measure the outcome in term of I_f, the final fraction of recovered (infected) nodes, as a function of T.

Susceptible-Infectious-Susceptible (SIS). An infections node transmits the disease to its susceptible neighbours at a rate α and recovers, returning to the susceptible state, at a rate β. We measure the outcome with I^*, the fraction of infectious nodes at equilibrium, as a function of $\lambda = \alpha/\beta$.

Results

Color code of targeting methods: **degree**, **betweenness centrality**, **memberships**, **coreness** and **random**.



For efficient immunization:

1. From the degree distribution, estimate the **virulence** of the disease in relation to the epidemic threshold T_c.
2. If virulent, evaluate the network's **community structure**; otherwise, go to 4.
3. If the community density is high ($\gtrsim 33\%$), immunize nodes according to their **memberships**; otherwise, go to 4.
4. For a disease near the epidemic threshold of the system, or for less modular networks, immunize according to node **degree**.

Acknowledgements

Calcul Québec

CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

NSERC CRSNG