Social (complex) Networks Analysis

CENTRALITY MEASURES

Rushed Kanawat

kanawati@sorbonne-paris-nord.fr

https://www.kanawati.fr

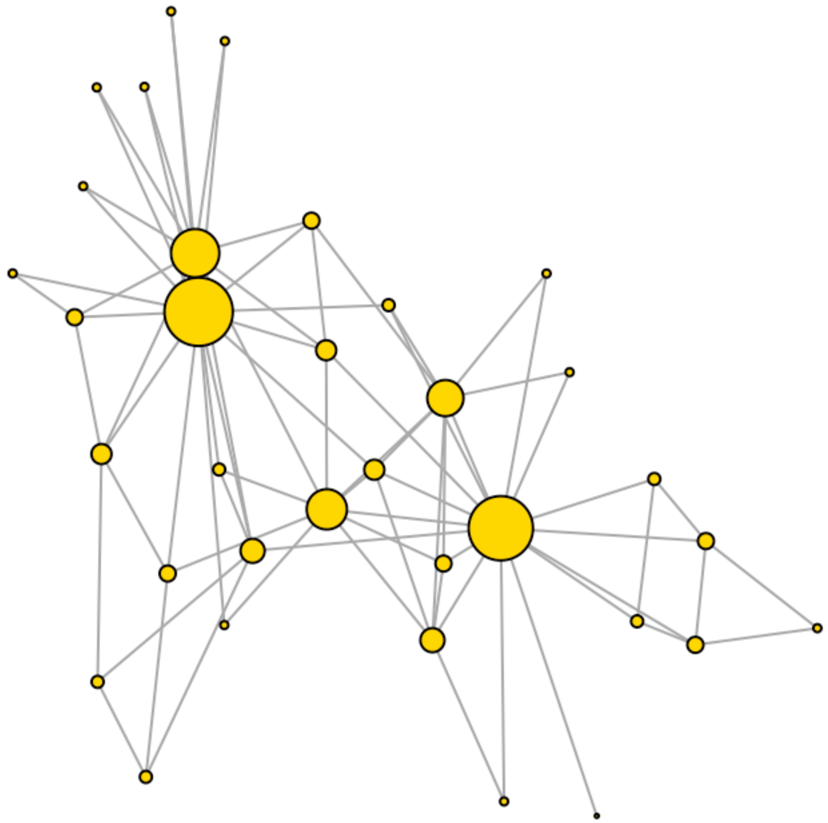
Université Sorbonne Paris Nord

Plan

1. INTRODUCTION
2. CENTRALITY METRICS
3. CENTRALITY CORRELATION
4. RANK AGGREGATION
5. ADVANCED TOPICS

1

Introduction



Centrality ?

Centrality: A measure of the relative **importance** of a node (or an edge) in a (complex) network.

- **Influential nodes**
- **Vulnerability nodes**
- **Control nodes**
- **…**

Centrality: some applications

Viral marketing

Influential nodes in social networks

Ranking web pages returned by a search engine

Influential nodes in web graphs

Ranking researchers

Influential nodes in citation networks

Where to vaccinate in priority to stop pandemic ?

Control nodes in human contact networks

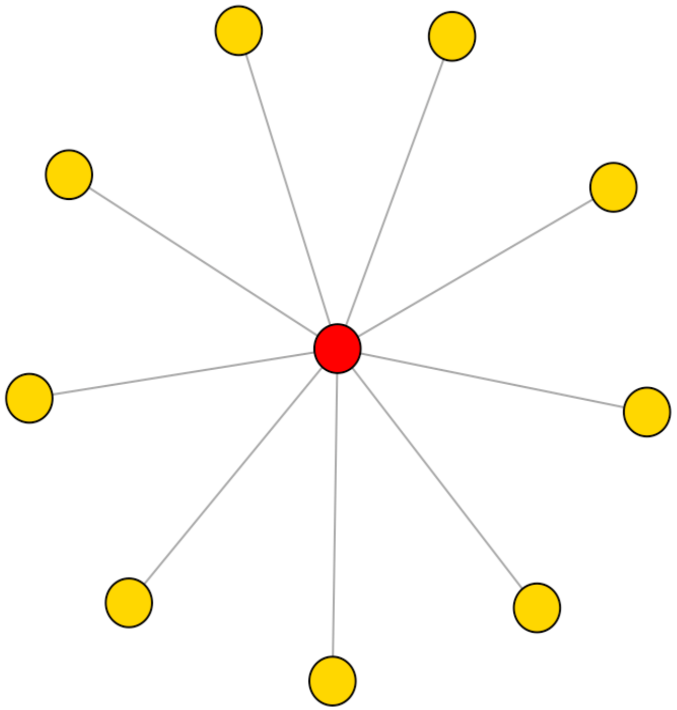
Which Internet router to attack ?

vulnerable nodes intenet graph

Intuitive example

Why is the central node in a star is the most important node ?

- It has the largest degree
- it has the smallest average distance to other nodes
- It is at the intersection of all shortest paths in the network
- It is the node that maximizes the dominant eigenvector of A_G
- ….



Centrality types

- Degree-based
- Distance based
- Path based
- Spectral measures
- …

Centrality ≠ Role

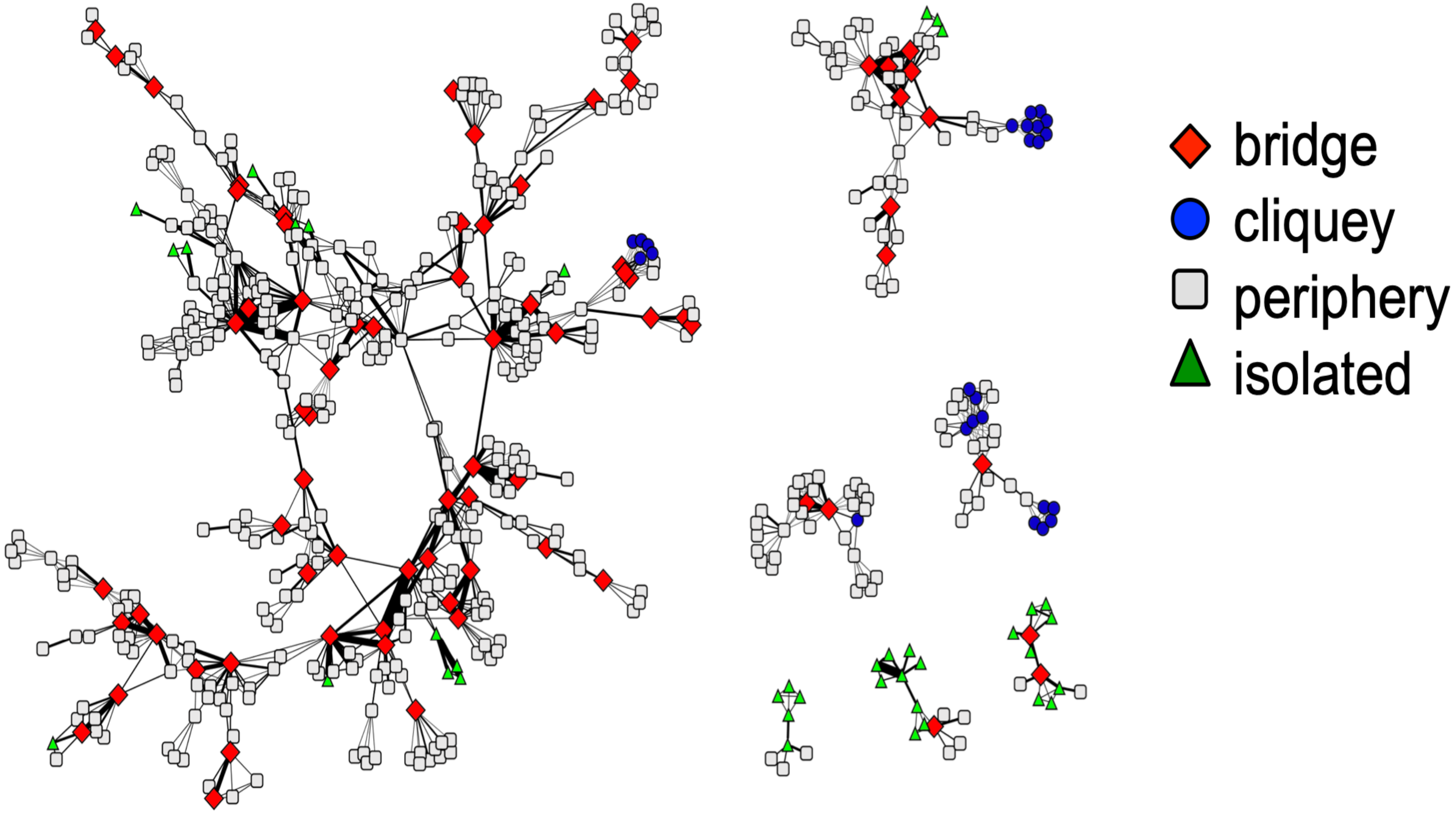
Role : Function of nodes in a network measured by structural behaviours.

Examples :

- Centers of stars
- Members of cliques
- Peripheral nodes
- …

Roles: examples

[D. Koutri et. al. *Node & Graph Similarity Tutorial, ICDM'2014]*

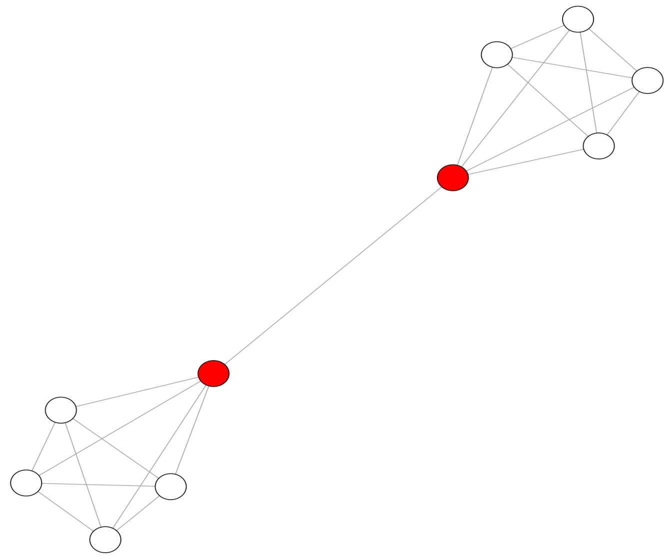


Example: Articulation point

Articulation point (aka cut nodes) are nodes whose removal increases the number of connected components in a graph.

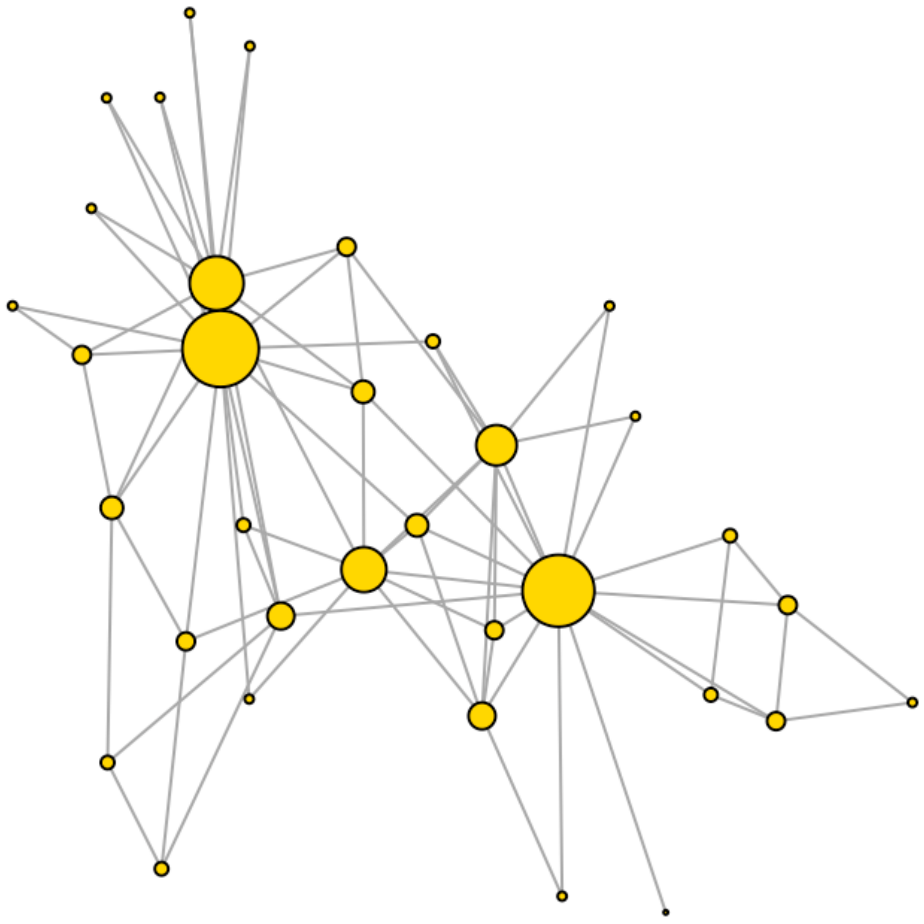
```
> V(g)[V(g) %in% articulation.points(g)]$color<-
"red"

> plot(g,vertex.label=NA)
```

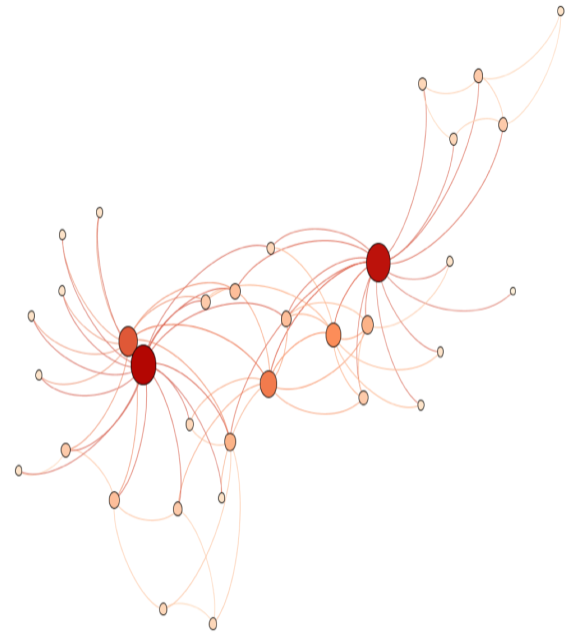


2

Centrality metrics



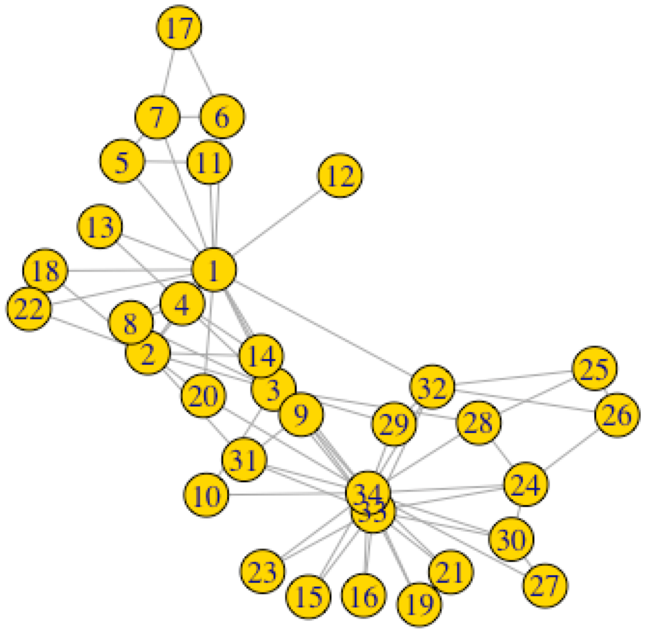
Degree-based centralities

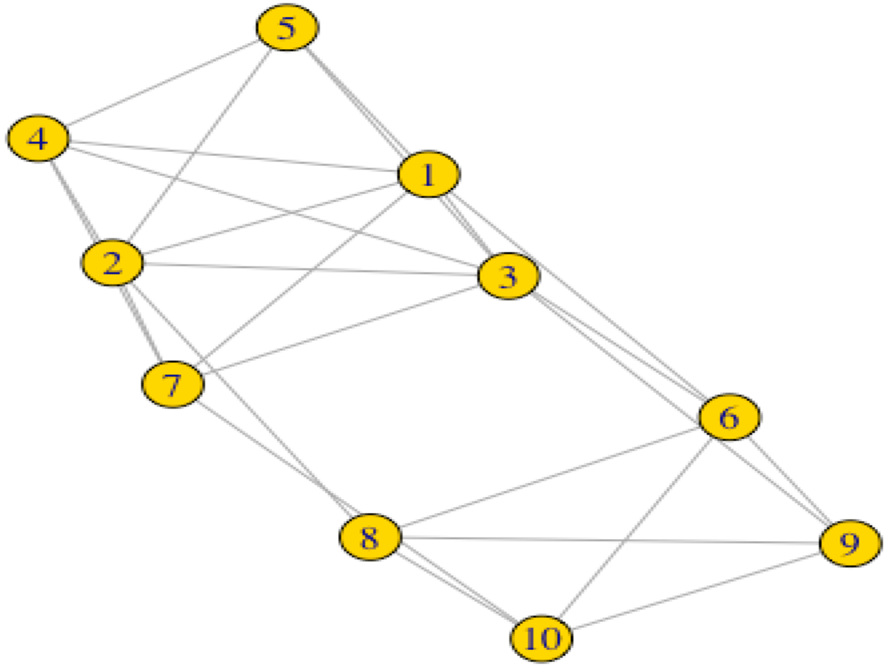


Degree centrality $C_d(v) = \frac{d_v}{n_G - 1}$. Complexity $O(m_G)$

K-core centrality : K-core is a connected maximal induced subgraph which has minimum degree **greater than or equal to k**.

K-core : illustration

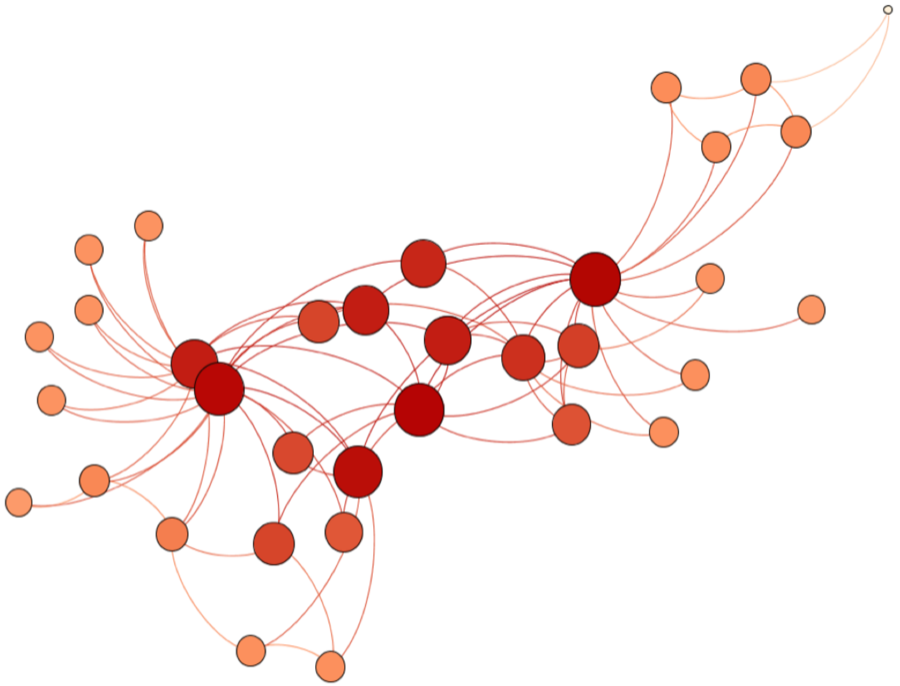




4-core

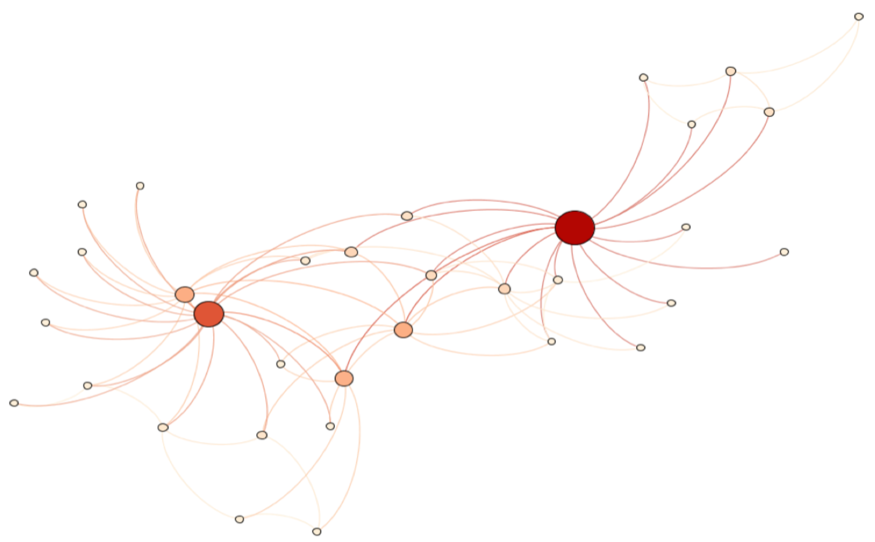
Closeness centrality

Closeness centrality. $C_c(v) = \frac{1}{\sum_{u \in V(G)} d(u,v)}.$



Complexity for binary networks $\mathcal{O}(n_G + m_G)$

Betweeness centrality



- $C_i(v) = \sum_{s,t \in V, stv} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$
- $\sigma_{s,t}(v)$: number of shortest paths linking s to t that include v
- $\sigma_{s,t}$: total number of shortest paths linking s to t
- Complexity = $\mathcal{O}(n^3)$

Eigen centrality

A node is central is it is connected to central nodes.

$$C(v) = \frac{1}{\lambda}\sum A_{i,j}C(x_j) \,:\, AX = \lambda X.$$

Hits

- HITS : Hyperlink Induced Topic Search
- IBM Web search engin Clever
- Two types of web pages
  - **Authority** : good information source
  - **Hub** : good index of authority pages

Hits

Soit G un graphe connexe, z le vecteur unité de $\Re^n$

$x_0 \leftarrow z$

$y_0 \leftarrow z$

Répéter jusqu'à convergence ou au max k fois :

- $x_i^{<p>} = \sum_{\forall q : q \rightarrow p} y_{i-1}^{<p>}$
- $y_i^{<p>} = \sum_{\forall q : p \rightarrow q} x_{i-1}^{<q>}$