# On the Resilience of Hierarchical Graphs.

**Francesc Comellas**

Department of Mathematics

Universitat Politècnica de Catalunya - Barcelona Tech

NORCOM 2022

Tromsø June 6–8, 2022

With the collaboration of former UPC Master and PhD students:

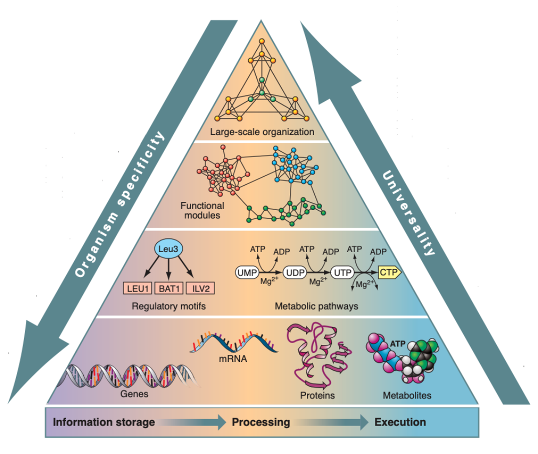
L. Sànchez, S. Aflatounian and J. Vilaltella.

## Overview

- We consider different families of hierarchical graphs introduced as models for complex networks (e.g. power grid and airport networks, human interactome, internet, social networks, ...)
- We compare a standard *degree* model for network cascading failures (Wang and Rong 2009) with modifications based on relevant graph centralities by applying them to different categories of graphs in the context of complex networks.
- In all tested cases, hierarchical graphs are significantly more resilient than the other graphs families.

# Hierarchical graphs

# Motivation: "Life's Complexity Pyramid"

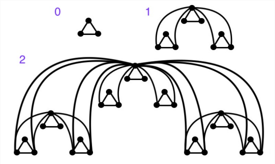


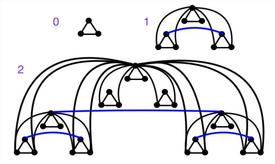

Life's complexity pyramid. Oltvai, Barabasi. Science 298 (2002) 763-764.

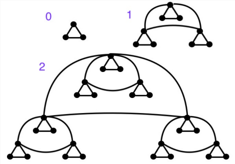
## Hierarchical graphs basic construction.

Consider a rooted graph. Make copies. Connect the copies through the root of the initial graph.

1. Main root joins all vertices of replicas except their roots and vertices of the original graph.
   - Scale-free and hierarchical structures in complex networks. Barabasi, Dezso, Ravasz, Yook, Oltvai, Proc. Modeling of Complex Systems, Granada 2-7 Sept. 2002.
2. As above but roots of replicas are joined by a clique to the main root.
   - Network biology: Understanding the cell's functional organization. Barabasi, Oltvai, Nature Reviews 5 (2004) 101-114.
3. Roots of replicas joined among them and to the main root by some graph (hierarchical product).
   - The hierarchical product of graphs, Barriere, Comellas, Dalfo, Fiol. Discrete Appl. Math. 157 (2009) 36-48.
   - A new graph product and its spectrum, Godsil, McKay. Bull. Austral. Math. Soc. 18 (1978) 21-28.

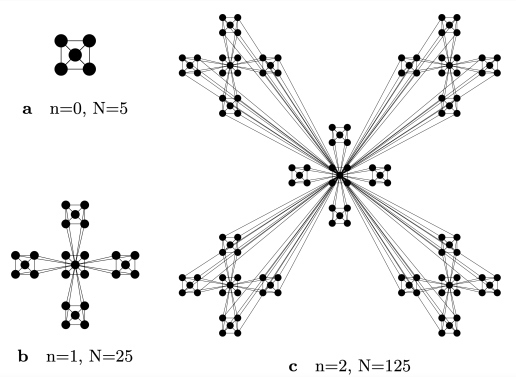






# Case 1. Main root joins all replicas vertices except roots

Initial version of (deterministic) hierarchical graphs

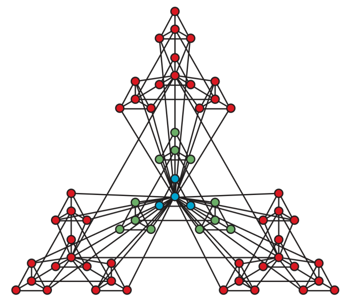



- Scale-free and hierarchical structures in complex networks. Barabasi, Dezso, Ravasz, Yook, Oltvai, Proc. Modeling of Complex Systems, Granada 2-7 Sept. 2002.

# Case 2. Adding to Case 1 a clique connecting roots of replicas .

To account for the coexistence of modularity, local clustering and scale-free topology in many real systems it is assumed that clusters combine in an iterative manner, generating a hierarchical network .

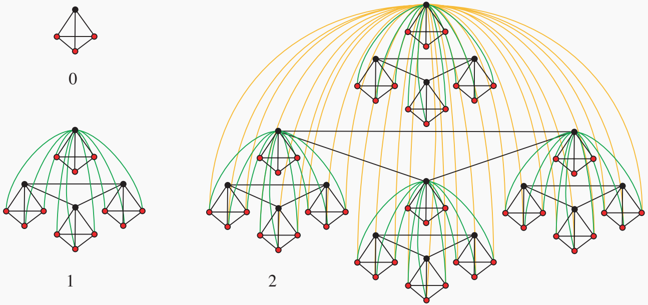
The starting point of this construction is $K_4$ (a 4-clique, the four central blue nodes in this figure). Consider one vertex as main *root.*
Next, three replicas of this module are generated and their three *roots* are joined as a $K_3$. The main root of the original $K_4$ is also joined to all other vertices of the replicas except roots and those related to the initial $K_4$.



- Hierarchical organization of modularity in metabolic networks. Ravasz, Somera, Mongru, Oltvai, Barabasi. Science 297 (2002) 1551-1555.

# Case 2. Adding to Case 1 a clique connecting roots of replicas .

We generalized this hierarchical graph structure, and obtained exact properties: radius, diameter, clustering coefficient and degree distribution




- Deterministic hierarchical networks. Barrière, Comellas, Dalfó, Fiol. J. Phys. A: Math. Theor. 49 (2016) 225202 (16 pp).

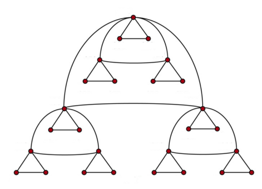
# Some properties of $H_{n,k}$

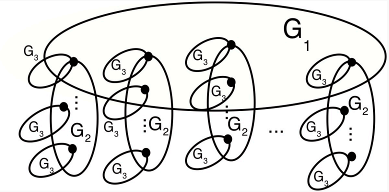
$H_{n,k}$ has order $|V_{n,k}| = n^k$ and size $|E_{n,k}| = \frac{3}{2}n^{k+1} - (n-1)^{k+1} - 2n^k - \frac{n}{2} + 1$.

| Vertex class | Num. vertices | Degree | Clustering |
|---|---|---|---|
| $H_{n,k}$ root | 1 | $\frac{(n-1)^{k+1}-(n-1)}{n-2}$ | $\frac{(n-2)^2}{(n-1)^{k+1}-2(n-1)+1}$ |
| $H^{\boldsymbol{\alpha}}_{n,k-i}$ roots $i = 1,2,\ldots,k-1$, $\boldsymbol{\alpha} \in \mathbb{Z}^i_n$, $\alpha_i \neq 0$ | $(n-1)n^{i-1}$ | $\frac{(n-1)^{k-i+1}-(n-1)}{n-2} + n - 2$ | $\frac{(n-2)^2}{(n-1)^{k-i+1}+(n-1)^2-3(n-1)+1}$ |
| $H_{n,k}$ peripheral | $(n-1)^k$ | $n+k-2$ | $\frac{(n-1)^2+(2k-3)(n-1)+2-2k}{(n+k-2)(n+k-3)}$ |
| $H^{\boldsymbol{\alpha}}_{n,k-i}$ peripheral $i = 1,2,\ldots,k-1$, $\boldsymbol{\alpha} \in \mathbb{Z}^i_n$, $\alpha_i \neq 0$ | $(n-1)^{k-i}n^{i-1}$ | $n+k-i-2$ | $\frac{(n-1)^2+(2k-2i-3)(n-1)+2+2i-2k}{(n+k-i-2)(n+k-i-3)}$ |

# Case 3. Hierarchical product of graphs.

Former cases suggested us to (re)define a hierarchical product of graphs.
Hierarchical product $H = G_N \sqcap \cdots \sqcap G_2 \sqcap G_1$






Instead of adding new edges, root vertices of the replicas are identified with vertices of the "parent" graph.

We can use it to find a recursive relationship for the adjacency and Laplacian matrices of the graphs that can be used to find the number of spanning trees, mean first passage time (for trees)...

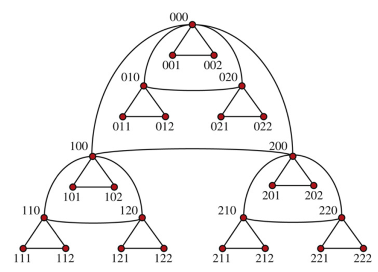
- [BaCoDaFi09] The hierarchical product of graphs, Barriere, Comellas, Dalfo, Fiol. Discrete Appl. Math. v. 157 (2009) pp.36-48.
- [BaCoDaFi09] On the hierarchical product of graphs and the generalized binomial tree. Barrière, Comellas, Dalfó, Fiol. Linear and Multilinear Algebra. 57 (2009) 695

# Hierarchical product of graphs.

The order of $H = G_k \sqcap \cdots \sqcap G_2 \sqcap G_1$ is $n_k \cdots n_3 n_2 n_1$.
If all graphs are the complete graph $K_n$ then the order is $|V(H)| = n^k$.

The number of edges of $G_2 \sqcap G_1$ is $m_2 + n_2 m_1$; the size of $G_3 \sqcap G_2 \sqcap G_1$ is $m_3 + n_3(m_2 + n_2 m_1) = m_3 + n_3 m_2 + n_3 n_2 m_1$; and so on.

For complete graphs $K_n$ the size is $|E(H)| = (n^{k+1} - n)/2$.

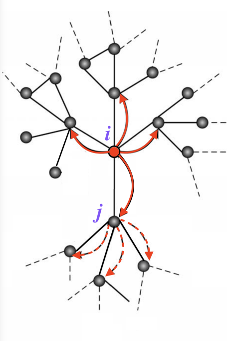



We can find analytically parameters like: eccentricity, radius, diameter, degrees, clustering. The labeling allows to find routing algorithms. Also, recursivity facilitates to find algebraic properties like the adjacency and Laplacian spectrum at each level that allow to find, for example, the number of spanning trees, the Kemeny constant (related to random walks and propagation) etc.

# Cascading Failures

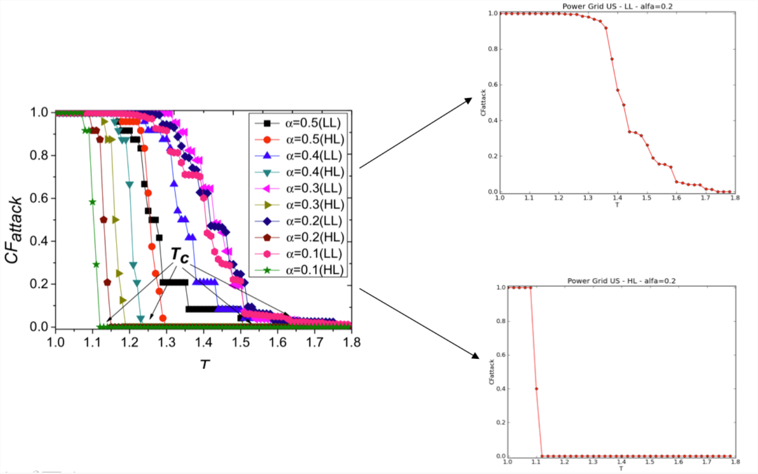
# Wang-Rong cascading failure model for the US power grid

- Power grid west US/Canada: 4941 vertices and 6594 edges, diameter 46. Watts-Strogatz, Nature 1998
- Each vertex has a **load**: $L_i = (k_i \sum_{m \in \Gamma_i} k_m)^\alpha$
- When a vertex fails its load is **distributed** among the neighbors. $L_j^{new} = L_j + L_i \frac{L_j}{\sum_{m \in \Gamma_i} L_m}$ .
- Each vertex has a **capacity**, $C_j = T \cdot L_j$ , based on a **tolerance** $T > 1$. If the capacity is exceeded, the vertex fails.
- Start the **cascading failure process** from each vertex of an initial failing set IFS (e.g. 1 % of $|V|$)
- Check how many vertices fail until the load is absorbed or the whole graph fails: $CF_{\text{IFS}} = \frac{\sum_{i \in \text{IFS}} CF_i}{N_{\text{IFS}}(N-1)}$




- Cascade-based attack vulnerability on the US power grid. J.W. Wang, L.L. Rong. Safety Science 47 (2009) 1332-1336. .

# Cascading Failures in the NW USA Power Grid: Results




Wang-Rong simulations: $N_{\text{IFS}} = 50 \quad ( \, |V| = 4941)$

# NY Times article. March 20, 2010

## Academic Paper in China Sets Off Alarms in U.S.

A Chinese student, Wang Jianwei, above, and his professor, wrote an academic paper on the vulnerability of the American power grid to a computer attack. Scientists said the paper was merely a technical exercise.
Du Bin for The New York Times

Larry M. Wortzel, a military strategist, ı

Larry M. Wortzel, a military strategist and China specialist, told the House Foreign Affairs Committee on March 10 that it should be concerned because "Chinese researchers at the Institute of Systems Engineering of Dalian University of Technology published a paper on how to attack a small U.S. power grid sub-network in a way that would cause a cascading failure of the entire U.S." When reached by telephone, Mr. Wang said he and his professor had indeed published "Cascade-Based Attack Vulnerability on the U.S. Power Grid" in an international journal called Safety Science last spring.

https://www.nytimes.com/2010/03/21/world/asia/21grid.html

# The Wang Rong model vs our models

- ⋆ The load considered is the degree of vertices or degree centrality
- ⋆ The network considered is western US / Canada power grid.
- ⋆ They sort all vertices by their load $L_j$ and select as initial failing sets : vertices with the highest load, with the lowest load and compare results
  - We also consider other centralities: Eigenvector, Information, Betweenness,...
  - We consider different graph families : Hierarchical, Erdős-Rényi, Barabasi-Albert, Random Geometric,...
  - We also consider random choices as initial failing vertices.

# Graph families

# Graph families

- Erdős-Rényi.
- Watts-Strogatz.
- Barabasi-Albert.
- Geographical Threshold.
- Random Degree Sequence.
- Random Geometric.
- Random Partition.
- Random Regular.
- Connected Caveman.

Defined in the Python package NetworkX.
We have used these versions: Pyhon v3.9 and NetworkX v2.8.

# Erdős-Rényi

## erdos_renyi_graph

`erdos_renyi_graph` (*n, p, seed=None, directed=False*)

Returns a $G_{n,p}$ random graph, also known as an Erdős-Rényi graph or a binomial graph.

The $G_{n,p}$ model chooses each of the possible edges with probability `p`.

The functions `binomial_graph()` and `erdos_renyi_graph()` are aliases of this function.

**Parameters:**
- **n** (*int*) – The number of nodes.
- **p** (*float*) – Probability for edge creation.
- **seed** (*int, optional*) – Seed for random number generator (default=None).
- **directed** (*bool, optional (default=False)*) – If `True`, this function returns a directed graph.

**References**

[1] P. Erdős and A. Rényi, On Random Graphs, Publ. Math. 6, 290 (1959).

[2] E. N. Gilbert, Random Graphs, Ann. Math. Stat., 30, 1141 (1959).

# Geographical Threshold

## geographical_threshold_graph

`geographical_threshold_graph`(*n, theta, alpha=2, dim=2, pos=None, weight=None*)
[source]

Returns a geographical threshold graph.

The geographical threshold graph model places `n` nodes uniformly at random in a rectangular domain. Each node $u$ is assigned a weight $w_u$. Two nodes $u$ and $v$ are joined by an edge if $w_u + w_v \geq \theta r^\alpha$

**Parameters:**

- **n** (*int*) – Number of nodes
- **theta** (*float*) – Threshold value
- **alpha** (*float, optional*) – Exponent of distance function
- **dim** (*int, optional*) – Dimension of graph
- **pos** (*dict*) – Node positions as a dictionary of tuples keyed by node.
- **weight** (*dict*) – Node weights as a dictionary of numbers keyed by node

### References

[1] Masuda, N., Miwa, H., Konno, N.: Geographical threshold graphs with small-world and scale-free properties. Physical Review E 71, 036108 (2005)

[2] Milan Bradonjić, Aric Hagberg and Allon G. Percus, Giant component and connectivity in geographical threshold graphs, in Algorithms and Models for the Web-Graph (WAW 2007), Antony Bonato and Fan Chung (Eds), pp. 209–216, 2007

# Random Partition

## random_partition_graph

`random_partition_graph`(*sizes, p_in, p_out, seed=None, directed=False*) [source]

Return the random partition graph with a partition of sizes.

A partition graph is a graph of communities with sizes defined by s in sizes. Nodes in the same group are connected with probability p_in and nodes of different groups are connected with probability p_out.

**Parameters:**
- **sizes** (*list of ints*) – Sizes of groups
- **p_in** (*float*) – probability of edges with in groups
- **p_out** (*float*) – probability of edges between groups
- **directed** (*boolean optional, default=False*) – Whether to create a directed graph
- **seed** (*int optional, default None*) – A seed for the random number generator

**References**

[1] Santo Fortunato 'Community Detection in Graphs' Physical Reports Volume 486, Issue 3-5 p. 75-174. http://arxiv.org/abs/0906.0612 http://arxiv.org/abs/0906.0612

# Properties / Parameters for these Graph families

20 different graphs with the same order $|V| = 343$ and size $|E| \approx 1697, 1497, 1125$ for each graph family – graphs are saved and used for each centrality.–

| graph family | vert. | avg. edges | avg. diam. | avg. dist. | avg. clust. | avg. Q idx. | avg. assort. | avg. GRC | k-cores |
|---|---|---|---|---|---|---|---|---|---|
| Hier BO clique | 343 | 1617 | 5 | 2.39 | 0.84 | -5.44 | -0.12 | 0.411 | 6-8 |
| Caveman | 341 | 1705 | 32 | 16.44 | 0.95 | -1.39 | -0.10 | 0.009 | 9 |
| Rnd Deg Seq | 343 | 1682.75 | 4 | 2.29 | 0.16 | -5.65 | -0.10 | 0.418 | 6 |
| E-R | 343 | 1672.3 | 5 | 2.80 | 0.03 | -1.27 | -1.27 | 0.057 | 2-7 |
| 2D geo thr | 343 | 1679.6 | 13.75 | 5.58 | 0.67 | -3.63 | 0.02 | 0.109 | 1-12 |
| Rnd Geom | 343 | 1687.4 | 18.3 | 7.42 | 0.61 | -2.68 | 0.58 | 0.052 | 1-11 |
| Rnd Part | 343 | 1646.05 | 5 | 2.84 | 0.04 | -1.28 | 0 | 0.058 | 1-7 |
| B-A | 343 | 1690 | 4 | 2.65 | 0.09 | -4.19 | -0.08 | 0.190 | 5 |
| W-S | 343 | 1715 | 5.1 | 3.22 | 0.35 | -1.96 | -0.01 | 0.043 | 6-8 |
| Rnd Reg | 343 | 1715 | 4 | 2.77 | 0.02 | -1.03 | – | 0.006 | 10 |

Basic properties (averaged) of the families of graphs considered

*Q idx* – If $Q < 0$, there is an efficient use of shortest-paths in relation to communicability.

- Complex networks in the Euclidean space of communicability distances. E. Estrada, Phys Rev. 85 (2012) 066122..

*GRC* – Global reaching centrality: $GRC = \frac{1}{N-1} \sum_{i \in V} |C_R^{max} - C_R(i)|$ with $C_R(i) = \frac{1}{N-1} \sum_j \frac{1}{d(i,j)}$

- Hierarchy measure for complex networks. E. Mones, L. Vicsek, T. Vicsek. PLoS ONE 7 (2012) e33799 (10 pp)

# New cascading failure models based on centralities

# New models for different centralities

Just replace $k_i$ (degree of vertex $i$) in the Wang-Rong model for the value of another centrality.

$$L_j = (k_j \cdot \sum_{m \in N(j)} k_m)^\alpha$$

# Possible Centralities

**Degree centrality:** Degree: Number of edges connected to a given vertex. *Measures the ability of a node to communicate directly with others.*

**Eigenvector centrality:** Associated with the largest eigenvalue of the adjacency matrix.

**Betweenness centrality:** Fraction of shortest paths between all vertex pairs that go through a given vertex. *Quantifies the number of times a vertex acts as a bridge along the shortest path between two other.*

**Information Centrality (Closseness Current Flow):** Related to the number of closed walks.

**Comunicability Centrality:** Mean distance (shortest paths) between a vertex and all other vertices reachable from it.

# Communicability (subgraph) centrality

Is the (weighted) sum of closed walks of all lengths starting and ending at vertex $u$.

## Definition

The **communicability centrality** of a vertex $u$ of a graph is

$$CC(u) = \left(e^{A}\right)_{uu} = \sum_{k\geq 0} \frac{1}{k!}\left(A^{k}\right)_{uu}$$

- Subgraph centrality in complex networks. E. Estrada, A. Rodriguez-Velazquez. Phys. Rev. E 71 (2005) 056103.

# Information (Current Flow Closeness) Centrality

Graph $G$ is a resistor network and every edge $e$ has unit resistance.
Let $v_{st}(u)$ denote the voltage of $u$ when a unit current enters the network at $s$ and leaves it at $t$.

## Definition

The current flow closeness $C(u)$ of a vertex $u$ is [BrFl05]

$$C(u) = \frac{|V|}{\sum_{v \in V} (v_{uv}(u) - v_{uv}(v))}$$

Current flow closeness centrality is equivalent to information centrality [StZe89].


- Centrality Measures Based on Current Flow. U. Brandes, D. Fleischer. In 22nd Annual Symposium on Theoretical Aspects of Computer Science, Proceedings (STACS 2005). Springer, 533–544.
- Rethinking centrality: Methods and examples. K. Stephenson, M. Zelen. Soc. Networks 11 (1989) 1–37.
- Ernesto Estrada. The Structure of Complex Networks.Oxford U.P. 2011, ISBN 978-0-19-959175-6. pp 146–147.

# Betweeness centrality (BC)

## Definition

The **betweeness centrality for a vertex** $v \in V(G)$ is

$$B_G(v) = \sum_{s,t \in G, s,t \neq v} \frac{\sigma(s,t \mid v)}{\sigma(s,t)}$$

where $\sigma(s,t)$ is the number of shortest $(s,t)-paths$ and $\sigma(s,t \mid v)$ is the number of those paths passing through $v$.

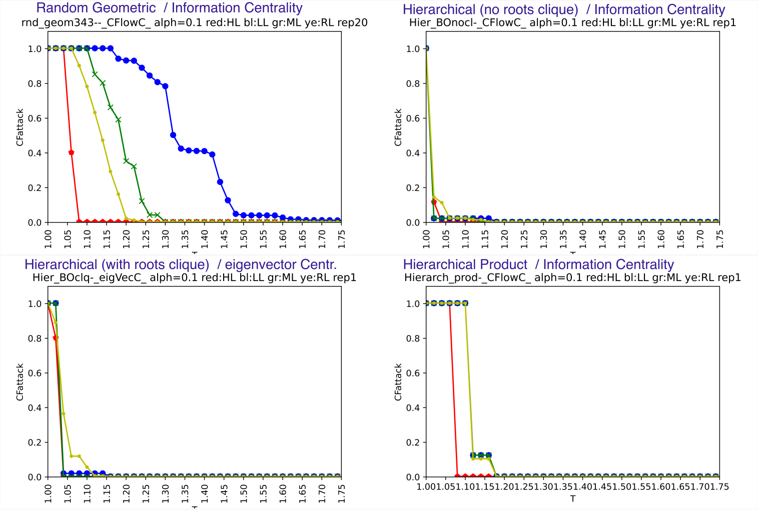
# Methodology

## Simulations done

- Cascading failure methods coded in Python 3.9, with NetworkX package 2.8 (to generate all graphs and call reprogrammed centrality functions).
- For each combination of a graph family, centrality and $\alpha = 0.1, 0.5$ , we run a set of 20 simulations for 20 different graphs of each family with the same order $|V| = 343$ and size $|E| \approx 1617, 1497 or 1197$ than the corresponding hierarchical graph considered.
- Each run involves increasing the load tolerance $T$ parameter from 1 to 1.8 in steps of 0.025.
- Results are averaged for these 20 simulations.

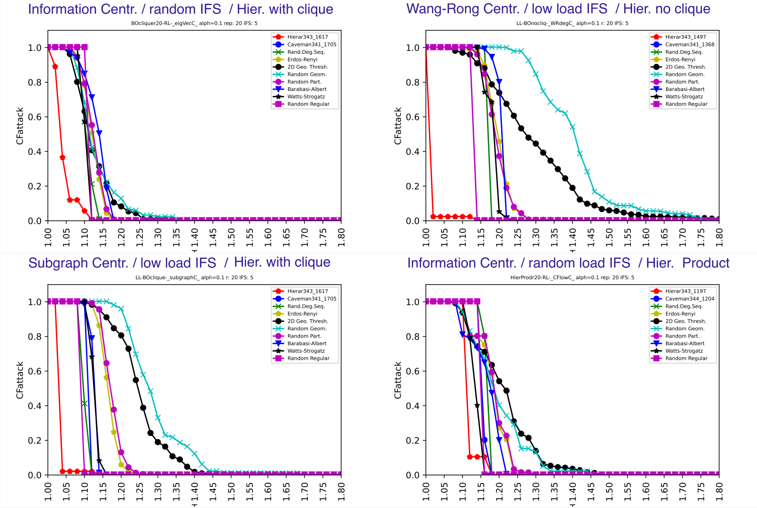
# Results

## Cases High /Mid /Low/ Rand Load for a given family




Comparison for CFlowC / eigVC, $\alpha = 0.1$, high/med/low/random load 5 initial vertices failures.

# Comparison among graph families and centralities




We note not a relevant different behavior with respect the hierarchical graph type and centralities when the initial failing vertices have low load or random load.

# Conclusions

# Conclusions

- All simulations show a higher resilience for hierarchical graphs for an initial failing set (IFS) with low load and random load.
- This resilience is independent of the centrality considered.
- Hierarchical product graphs show less resilience than the other hierarchical graphs.
- The results contradict a recent paper by Robson et al.:
  The structure and behaviour of hierarchical infrastructure networks. Robson, Barr, Ford, James. Appl. Netw. Sci. (2021) 6:65. doi:10.1007/s41109-021-00404-4

  but they measure graph "robustness" in a different way

# Conclusions

To increase a graph resilience, and thus to protect a real life network associated to it, there is need to decide which is the best centrality to model flows in the network and then check all vertices according to it to select those vulnerable (low load). Add edges to them to increase “hierarchy”.

Thank you !