ETH zürich

Robustness on the *R&D* alliance network: A multi-layer approach

Jiduan Wu, D-INFK

Supervisors : Dr. Giacomo Vaccario, Dr. Giona Casiraghi, Prof. Dr. Dr. Frank Schweitzer

Outline of talk

Introduction

Method

Results

- Two-layer approximation
- Multi-layer network

Motivation

- Understanding the **robustness** of a network against removing nodes involves analyzing its topology and plays a significant role in improving its resilience.
- Most researches focus on **robustness analysis** in **unipartite** graphs.
- Exceptions for **two-layer network*** reveal how **differentiating** nodes in different layers and **multi-layer** structure influence the robustness of the network.
  - Inter-layer connections.
  - Intra-layer connections.
- What about the **robustness** of **multi-layer** networks with **three or more layers**?

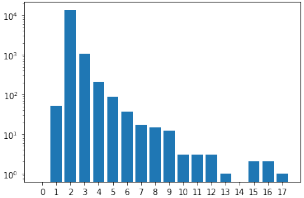
*__casiraghi2019probing__.

Empirical dataset

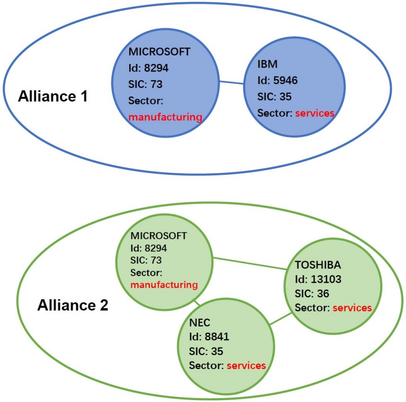
- 14535 valid **firms**, each firm belongs to a sector.
- 10 **sectors**: services, manufacturing, public administration...
- 14881 valid **alliances** with **participated firms** and durations.
  - Each firm can participate in **several alliances**.

	Minimum	Median	Maximum	Mean
#alliances per firm	1	1	284	2.2017

  - Alliance size: **any positive number**.



Alliance sizes distribution



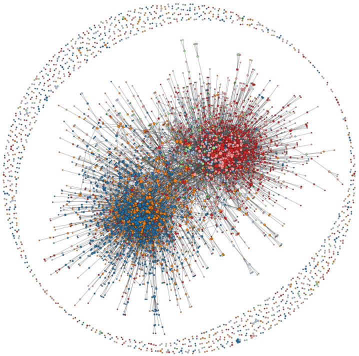
Dataset snapshot

Firm network

- **Node** ↔ **Firm**; **Layer** ↔ **Sector**.
- An **undirected** and **unweighted edge** exists between two firm nodes if they share **common alliance(s)**.

	Minimum	Median	Maximum	Mean
#node degree	0	1	421	2.9643

- **Static network**: **durations** of alliances are **ignored**.
- **Periphery**: not in the largest connected component (LCC) of the network.
  - **Periphery fraction**: 41.07%.



- A **similar topology** to the network[†] on the right.

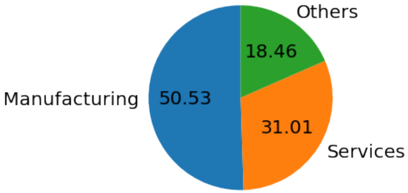
[†] **Giacomo2019**.

Characterising properties

- **Large scale**: 14535 nodes.
  - Inefficient to analyse some properties such as the nestedness.
- **Sparsity**: for matrix M

$$\text{Filling rate}(M) = \frac{\#\text{non-zero entries}}{\#\text{row} \times \#\text{column}} \tag{1}$$

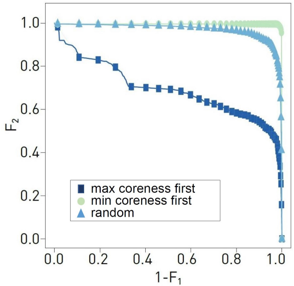
  - **filling rate** of the adjacency matrix is 0.000188.
- **Dominant layers**: **manufacturing** layer and **services** layer.
  - In the whole network, the **fraction of nodes** in these two layers are above 80%.
  - In the LCC, the **fraction of nodes** in these two layers: 84.71%, the **fraction of edges** involving nodes in these two layers: 94.15%.



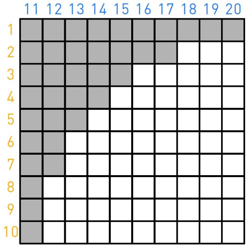
Node-layer distribution

Cascading model[‡]

- **Attacked layer**: its nodes are removed with three strategies.
  - Remove nodes randomly (random), with increasing coreness (min), and with decreasing coreness (max).
- **Synthetic two-layer network**:
  - A node is removed when disconnected with the **attacked layer**.
  - **Coupling: inter-layer** connections.
    - **Nestedness**: "Nestedness can be described as the **tendency** for nodes to interact with subsets of the interaction partners of **better-connected nodes**".
- **Multi-layer network**
  - A node is removed when disconnected with **all the other layers**.



Cascades results



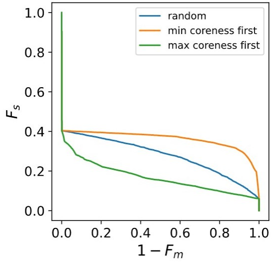
Perfectly nested incidence matrix

"Generalists" to **"generalists"**

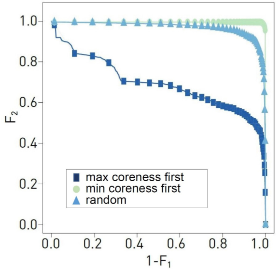
‡**casiraghi2019probing**.

The importance of filling rate

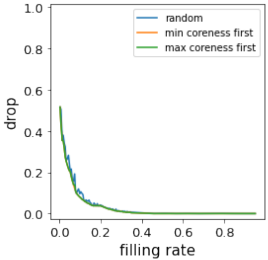
- Big "drops" in the empirical network, not in the **baseline synthetic network**.
  - Empirical network: filling rate of the **incidence matrix** $\rightarrow$ 0.000326.
  - Baseline synthetic network: filling rate of the **incidence matrix** $\rightarrow$ 0.35.



Empirical network (LCC)



Synthetic network

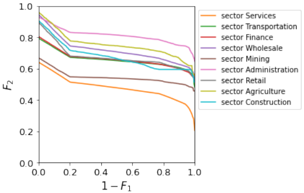


Filling rates and **"drops"**

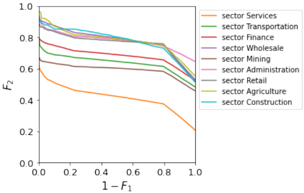
- The drops are caused by the low filling rate the and the assumption that nodes get removed once disconnected with the attacked layer.
- The drop $\rightarrow 0$ when filling rate $\gtrapprox 0.35$.

Cascades in multi-layer network

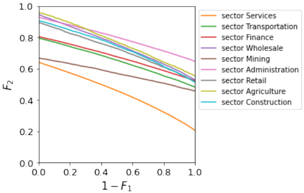
- An example of cascade results on the **whole network**.
  - Attacked layer is the **manufacturing** layer (the **largest layer**, 50.53%).
  - "Drops" at the beginning ↔ Sparsity (**filling rate** of the adjacency matrix is 0.000188).



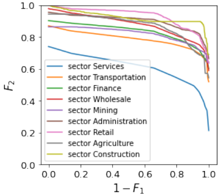
min corness first (whole)



max coreness first (whole)



random (whole)

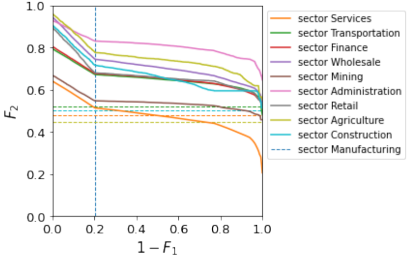


min coreness first (LCC)

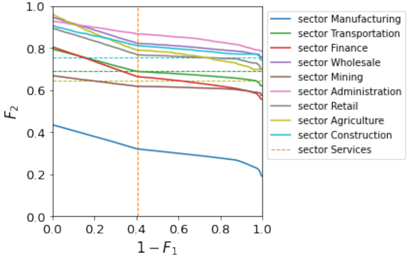
- **Robustness**: even attacking the largest layer (**half** of the nodes) doesn't cause huge cascades.
- **Min coreness first**:
  - **Fragile periphery**: **linear faster decays** at the beginning in the whole network scenario.
  - **Resistance of LCC**: cascades **slow dow** after the **turning point**.

Why is the periphery important?

- **Low-coreness**:
  - A **considerable** part of nodes with **low coreness** in each layer are in the periphery.
    - **Coreness 0**: Manufacturing layer: 55.46%; Services layer: 42.07%.
- **"Well-mixed" periphery**:
  - 92.3% of alliances in the periphery have sizes of 2.
    - 54.8% of 2-alliances consist of two firms from the **same** sector.
    - 45.2% of 2-alliances consist of two firms from **different** sectors.
  - 87.62% of nodes in the periphery have **coreness 1**.
- **Results under "min coreness first" attacks**:
  - Vertical lines: fractions of nodes with **coreness 0**; Horizontal lines: fractions of nodes **connected with** the attacked layer.
  - **Robustness**: most of curves are **close to** their corresponding horizontal lines.



Manufacturing layer

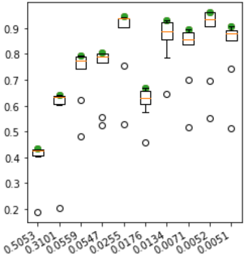


Services layer

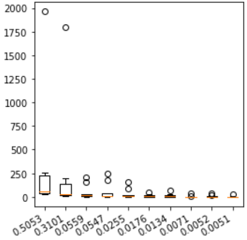
Robustness analysis

Cascade results with different attacked layers:

- **Final fractions of remaining nodes:**
  - x-axis: layer **proportion**, e.g. 0.5053-manufacturing layer; y-axis: **remaining fractions of x**, the attacked layer ∈ the other 9 layers; Green scatters: **real initial fractions** excluding "drops".
  - **Robustness**: removing any layer **won't damage** the **majority** of the **remaining parts** of other layers.
  - **Outliers**: manufacturing layer; services layer.
- **Damages caused by the attacked layer:**
  - x-axis: layer **proportion** of the **attacked layer**; y-axis: #**removed nodes in other 9 layers (drops excluded)**.
  - **More outliers** and **decreasing interquartile ranges**: **small layers** are more **unevenly** connected with other layers.



Final fractions



Damages

Summary

- **Insights into the empirical dataset**
  - **Multi-layer structure** **adds** to the **robustness**.
    - Protect nodes against attacks.
    - Constrain the propagation of cascades.
  - **"Well-mixed" struture** of the periphery → **more fragile** under the "min coreness first" attacks.
    - **Instant huge losses** can be **avoided** if the periphery is "protected".
- **Methodology**
  - In **cascade simulations**, **"drops"** (cascades) can also be caused by the **sparsity** of the network.
    - These **" drops"** should be **differentiated** from standard cascading processes.
  - Focusing on the LCC is not enough: **the periphery can expose unexpected fragilities and weaknesses**.
    - Important information is **maintained** → **LCC approximation**.

References

Giona Casiraghi and Antonios Garas and Frank Schweitzer
Probing the robustness of nested multi-layer networks
arXiv:1911.03277 [physics.soc-ph] (2019).

Giacomo Vaccario
The structure, exchange, and transfer of knowledge in sociotechnical systems
Ph.D. ETH. (2019).

Manuel Sebastian Mariani, Zhuo-Ming Ren, Jordi Bascompte and Claudio Juan Tessone
Nestedness in complex networks: Observation, emergence, and implications
Physics Reports, Volume 813, 2019, Pages 1-90.

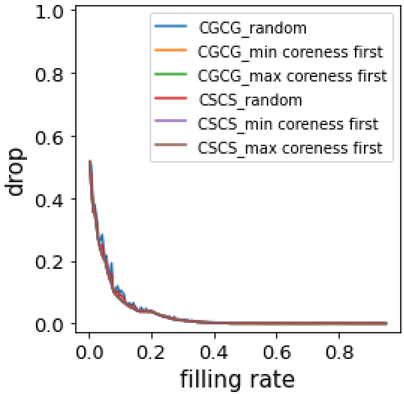
Bascompte, J., Jordano, P., Melián, C. J., & Olesen, J. M.
The nested assembly of plant-animal mutualistic networks
Proceedings of the National Academy of Sciences of the United States of America, 100(16), 9383–9387 (2003) .

Thanks for your time!

Q&A

Slides for possible questions

1. Alliances of size 1 are formed by several subsidiaries of one firm.
2. Filling rates- "drops" relationships under different coupling strategies.
3. "Drops" distribution:
   - 6233 (42.88%) nodes in total.
   - 48.36% of "drops" are in the periphery.
   - 51.64% of "drops" are in the LCC.



4. Nestedness analysis for large-scale and sparse network

- **Two-layer approximation** + **bipartite projection**.
  - Connectance of the bigraph: **0.00015**.
- **Large-scale** empirical dataset: **computational efficiency**. Popular algorithms such as NODF, MT, and BR $\rightarrow$ **computationally inefficient**.
- **Correlation** with other properties:
  - E.g. dissortativity, core-periphery metric ξ, and modularity.
  - **Modularity**: **low-density** networks exhibit a **positive correlation** between nestedness and modularity, while **high-density** networks exhibit a **negative correlation** between them.
  - The **z-score** computed with 50 null models generated by Bascompte's algorithm is -16.6718. The probability to reject the hypothesis that our empirical bigraph consists of "Manufacturing" layer and "Services" layer has modularity is higher than 0.999999.
  - **No modularity w.h.p** $\xrightarrow{w.h.p}$ **no nestedness w.h.p**.