Marco Grassia

Real-world Network Science

1st year Ph.D. student

Introduction

Network Science is the academic field which studies complex networks, i.e., networks (graphs) with non-trivial topological features that are used to **model complex systems**. **Each element** of the system **is** represented as **a node** (vertex) and **interactions** between those elements **are represented as links** (edges).

Long Distance Links [1] [2] [3]

We found that creating new links between the farthest nodes in the network (i.e., shortest path length is maximum):

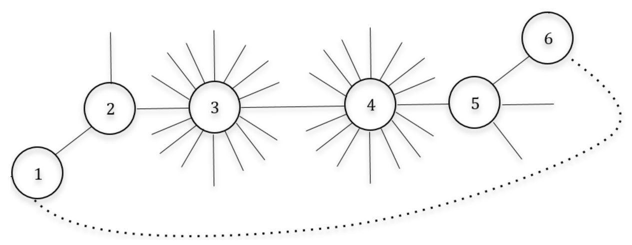

Figure 1: Example of long distance link between nodes 1 and 6.

- Improves **Network Robustness** to targeted node attacks;
- Improves the **PageRank** of the target node, outperforming most heuristics;
- Can tackle the ***In-Link Building*** multi-objective problem, allowing nodes to climb the (PageRank) rankings with less in-links, which translates to smaller costs.

Network Robustness [4]

We enhance network robustness using Long Distance Links and compare this strategy with other popular ones. Our results show that it grants a **peak increase** of the **Largest Connected Component size** of **more than** 17% **with** the addition of **just** 2.5% **new links**.

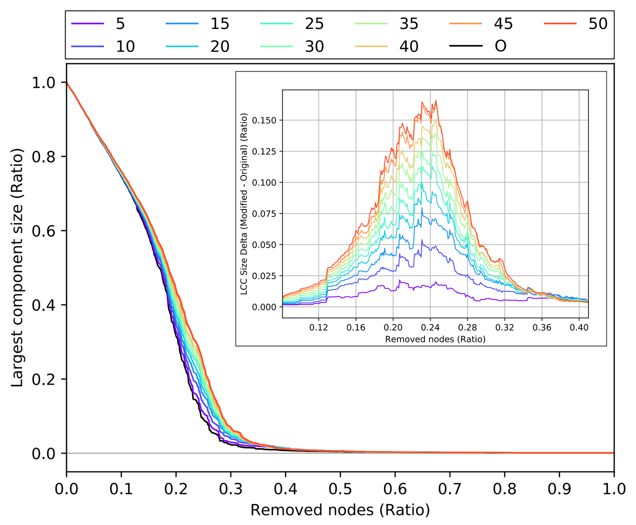

Figure 2: Largest Connected Component size as edges are added

For instance, in figure we compare this strategy with a similar one that connects nodes with lowest degree and that grants better performances during the initial stage of the attack. although there is some overlap between the two, the Long Distance based tends to connect nodes that belong to the peripheral zones of the network, keeping them together as the attack progresses.

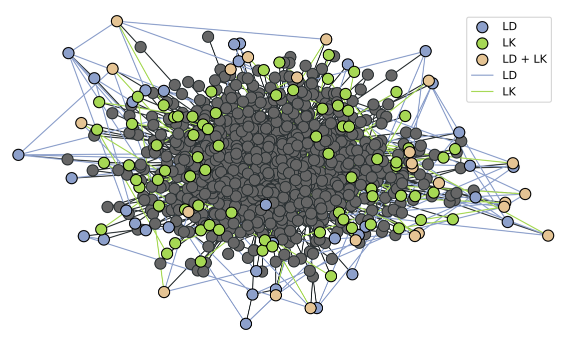

Figure 3: Low Degree and Long Distance comparison

Dipartimento di
Ingegneria Elettrica Elettronica
e Informatica

Tackle hard problems with Machine Learning

The main research topic of my Ph.D. is the application of **Machine Learning techniques to tackle computationally hard problems** on networks, namely **NP-hard** problem. My efforts on this topic began with my internship at Nokia Bell Labs Dublin, where I worked on the search space reduction for the **Maximum Clique Enumeration (MCE)** [5] problem as part of my Master's thesis.

Network Dismantling (under review)

Robustness is the **ability of a network to withstand failures** in its structure. Considering the importance of many real-world networks and their natural resilience to random failures due to their degree distribution, in our work we focus on targeted attacks, which could disrupt the networks easily. Specifically, we aim to find the critical subset of nodes that cause the sudden disruption of the Largest Connected Component when removed. Of course, removing a node comes with some cost, so the smaller the subset, the better. However, the problem of finding the optimal subset of nodes is an NP-hard problem.

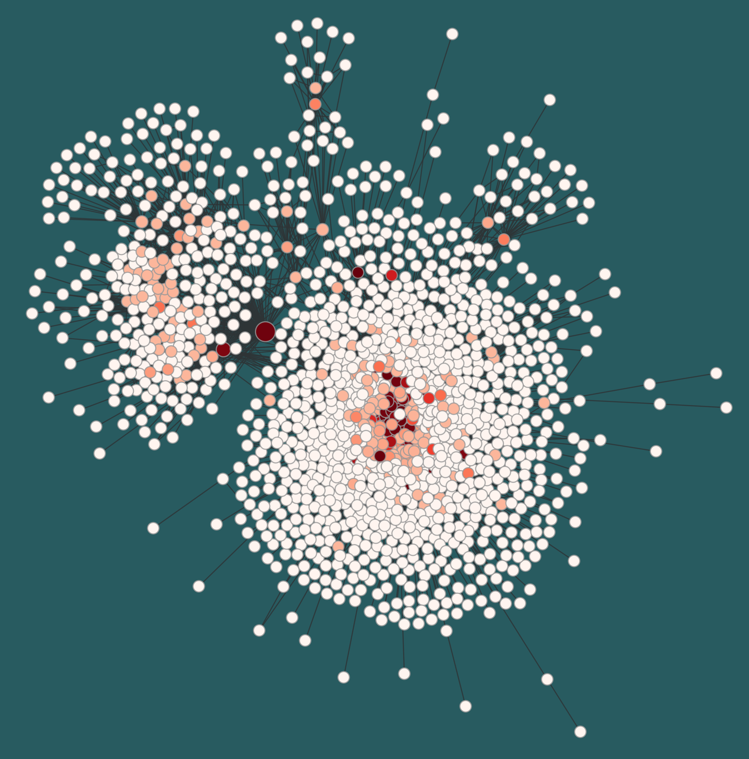

Figure 4: Predictions on US Airports network

Our goal is to find a **low complexity strategy** that can **outperform other popular heuristics**, like the removal of nodes in descending degree or betweenness order, using Geometric **Deep Learning**. We test our approach on real-world networks from various domains (e.g., social, infrastructure and technological networks). Our results show that our approach has lower cumulative Area Under the Curve (AUC) of the dismantling curve than the other heuristics, meaning that, on average, it outperforms them. Another advantage of our approach is that it relies on features that are computationally cheap, so it is expected to scale well. Moreover, we stress that we learn how to attack networks directly from the raw data.

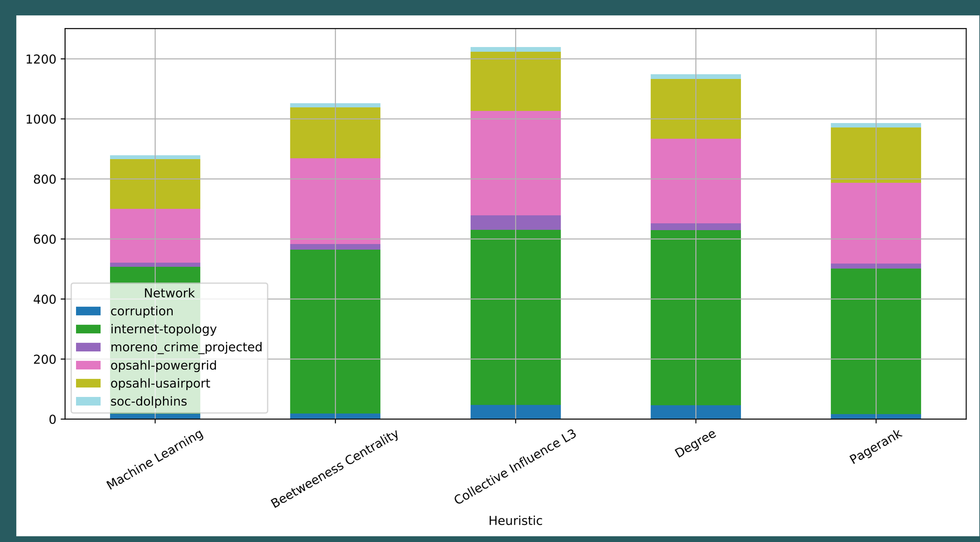

Figure 5

[1] Carchiolo, V., Grassia, M., Longheu, A., Malgeri, M. & Mangioni, G. Long distance in-links for ranking enhancement. In *International Symposium on Intelligent and Distributed Computing*, 3–10 (Springer, Cham, 2018).

[2] Carchiolo, V., Grassia, M., Longheu, A., Malgeri, M. & Mangioni, G. Climbing ranking position via long-distance backlinks. In *International Conference on Internet and Distributed Computing Systems*, 100–108 (Springer, Cham, 2018).

[3] Carchiolo, V., Grassia, M., Longheu, A., Malgeri, M. & Mangioni, G. Strategies comparison in link building problem. In *International Symposium on Intelligent and Distributed Computing*, 197–202 (Springer, 2019).

[4] Carchiolo, V., Grassia, M., Longheu, A., Malgeri, M. & Mangioni, G. Exploiting long distance connections to strengthen network robustness. In *International Conference on Internet and Distributed Computing Systems*, 270–277 (Springer, Cham, 2018).

[5] Grassia, M., Lauri, J., Dutta, S. & Ajwani, D. Learning multi-stage sparsification for maximum clique enumeration. *arXiv preprint arXiv:1910.00517* (2019).

Econophysics (WIP)

We also employ Network Science to study the **international relations between countries**. In fact, ***weighted multiplex* networks** can model the **trade of goods and services, the flow of money and people**, and other deals between countries **across the world**.

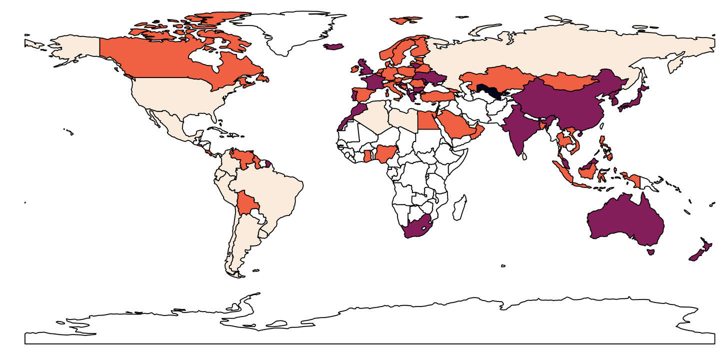

Figure 6

For this work we cooperate with a well known econophysics from Sant'Anna School of Advanced Studies. After a **preliminary statistical study**, we find the **communities** in the world network in two different years (**pre and post economic crisis**), study the cross-layer **correlations** and also their correlation with the geographical communities.

Black Hole metrics (in press)

We define a **PageRank based** approach **to measure the cohesion between groups** in a network. **Given a set of** N **groups, we define** $N(N-1)$ **dummy Black Hole nodes** BH_{ij} such that they **capture the mistrust** ratio between groups i and j, and define the $Cohesiveness_{ij}$ as $1 - BH_{ij}$. If $i = j$, we call it **internal cohesion**, or **external cohesion** otherwise. This metric is well suited for **trust networks**, where the weight of each edge has a lower and upper bound. The main advantage of a **dynamic measure** is that the resulting value of cohesion is affected by each node differently, i.e., the more important a node is, the larger is its impact.

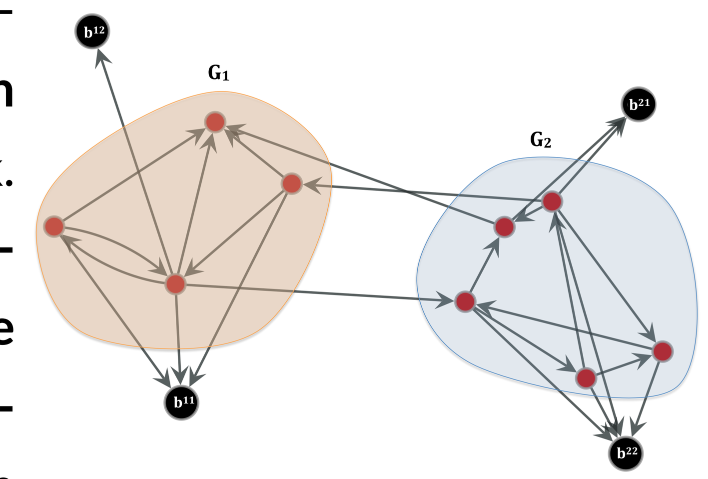

Figure 7

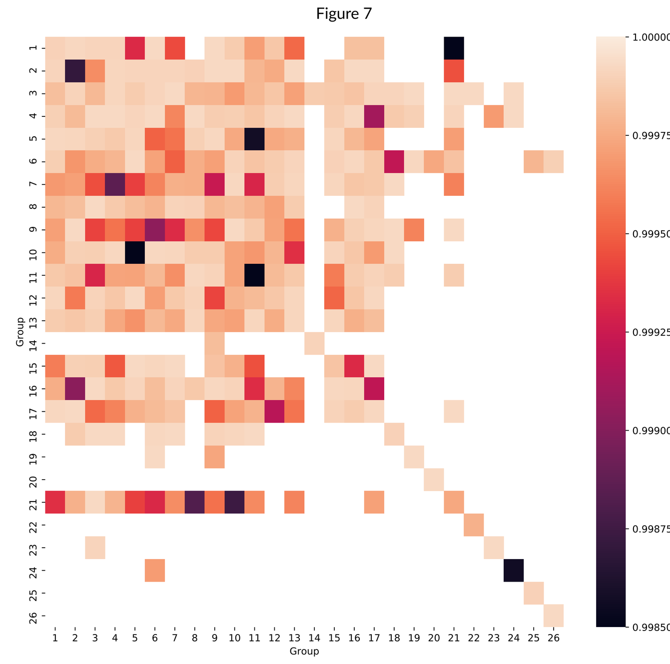

Figure 8

Scan for **more about my research activity**!