Network Analysis: Assignment 2

Second assignment of the network analysis course

Part 1

For the first part of the assignment we performed the experiments of the option 1, which plans to perform *robustness* test on small synthetic graphs. The tests of the robustness are performed simulating different types of attack, removing nodes according to the listed critera:

- **Higest betweenness**.
- **Highest degree nodes**.
- **Higest pagerank**.
- **Random Nodes**.

The robustness is given by the current size of the diameter and the giant component, which are measured every time a node is removed from the network

Karate Club Graph

The first graph we treated is the already known **Zachary's Karate Club** graph, which represent a social network of a university karate club. This network is composed of 34 nodes (member of a karate club), and there are links between pairs of members who interacted in and outside the club.

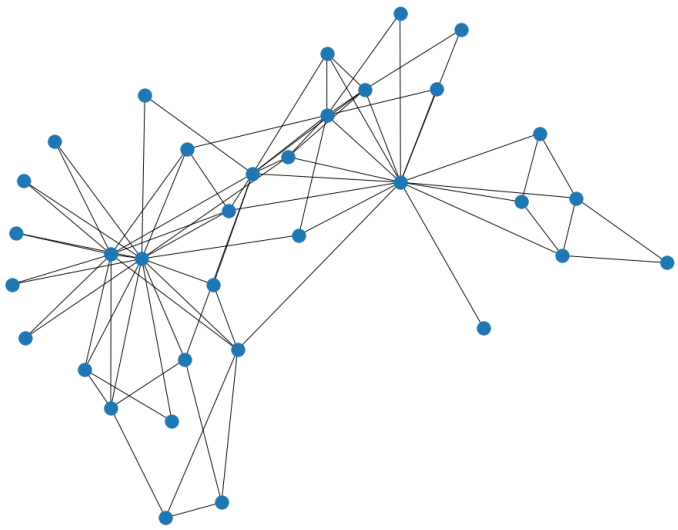


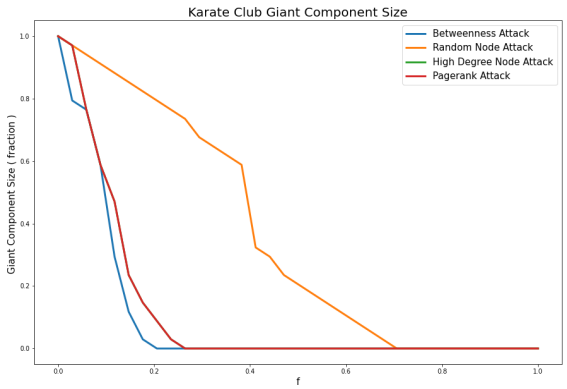
Figure 1 - Karate Club Graph

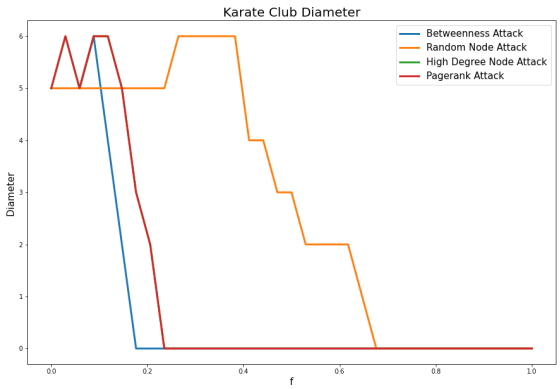
Dataset statistics	Values
Nodes	34
Edges	78
Average Degree	4.58
Average clustering	0.57
Global clustering	0.25
Nodes giant component	34
Diameter	5
Average shortest path	2.40
Density	0.13
Assortativity	-0.47

The figure below on the left shows the relative size of the **giant component** (computed by dividing the size of the giant component with the size of the original graph, since the network is connected) as the nodes of the graph are sequentially removed. The networks is more robust against *random failure*, and in such case the critical threshold f_c is more and less 0.7.

By contrast the network seems to be very vulnerable to all the *target attacks* with a $f_c = 0.2$. With high degree, betweenness and Pagerank attacks the behavior is the same, and they also overlap in some part.

The figure on the right shows how the **diameter** changes depending on the type of network attack. Again, the targeted attacks breaks apart the network in a few steps, we have a small increase in diameter at the beginning, but then decrease at $f \simeq 0.2$, going from 6 to 0 very rapidly. The random attack takes more steps to disaggregate the network, in which we have an increase of the diameter before $f \simeq 0.7$ and then there is a drastic decreasement of the diameter which goes rapidly to zero. The initial increase is due to selection of the largest subgraph when the network is not connected anymore.





Even if it is a very small network it shows a behavior as it is a power-law one, in fact it is dissassortative and to connect the nodes relies on a few hubs.

Erdős–Rényi Random Graph

The second synthetic graph is a random graph, generated with the the Erdős–Rényi model.

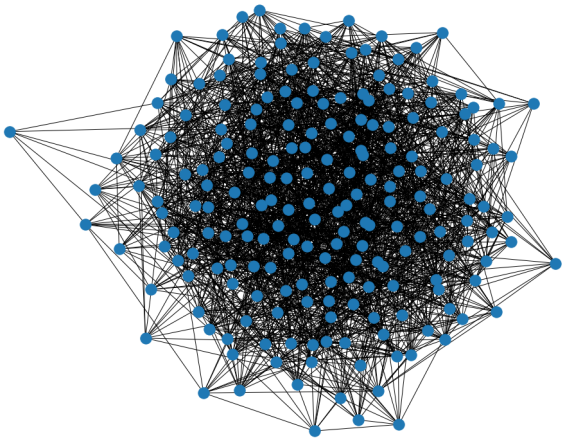


Figure 2 - Erdős–Rényi Random Graph - N = 200, p = 0.1

Dataset statistics	**Values**
Nodes	200
Edges	2017
Average Degree	20.17
Average clustering	0.09
Global clustering	0.10
Nodes giant component	200
Diameter	3
Average shortest path	2.01
Density	0.10
Assortativity	0.03

This kind of network behaves very differently: The plot below of the **giant component** (on the left) shows that the Erdős–Rényi random graph we made is robust against failures of all the types we tested. The size of the giant component decreases monotonically with all the attacks, but also here the targeted attacks break apart the giant component more rapidly.

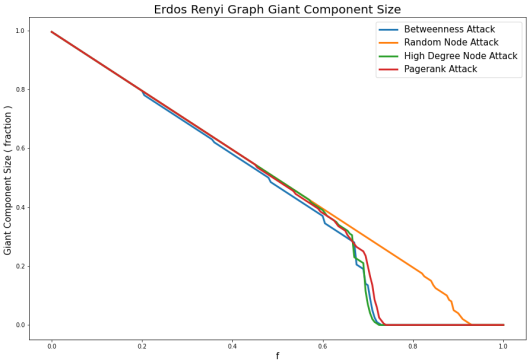
As in the case of the Karate Club network, the network is more robust against the random attack with a critical threshold f_c equal to more or less 0.9. With the target attacks, instead, we have at most $f_c = 0.7$. But it requires that most of nodes should be removed.

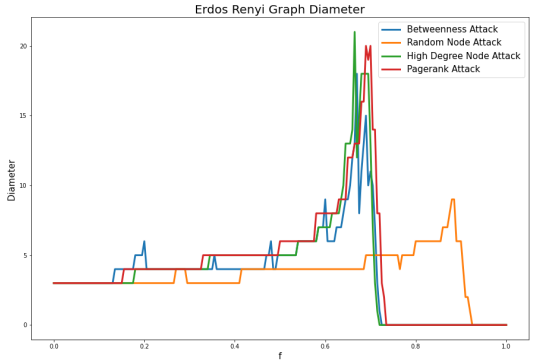
The plot on the right represents the **diameter** variations: also here it's clear that the network is more robust against random attack, in fact the diameter doesn't change too much until the majority of the nodes are not

removed, it oscillates until $f \simeq f_c = 0.9$ where the network is completely destroyed.

On the other hand, with the target attacks the diameter increases a lot until the critical point $f \simeq 0.8$, where the network break apart and the diameter goes quickly to 0. Since the diameter characterizes the ability of two nodes to communicate with each other, the target attacks, affect the interconnectedness of the network by increasing a lot the length of the path between two nodes that want communicate.

The lines are chainsaw shaped because each time a node is removed the biggest connected subgraph is selected and the new diameter can be particularly different.

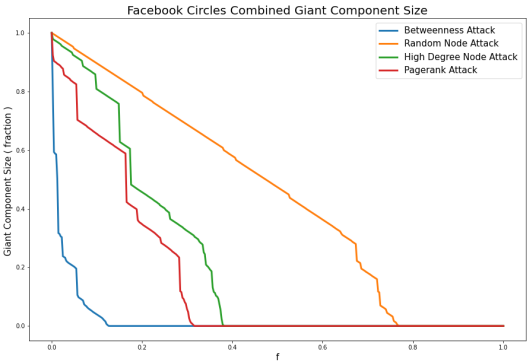


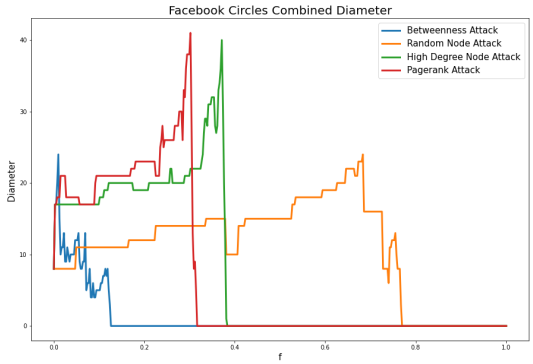


Part 2

For the second part of the assignment we performed the same experiments on the Facebook dataset. Since this network is very large the experiments are made removing 10 nodes at times, otherwise the execution would take too much.

Facebook Circles Combined





This network is scale-free and hence it should be more vulnerable to target attacks.

Especially with the betweenness attack because the graph has very few nodes with an high betweenness, the giant component will break apart in a very few steps. the critical threshold f_c, is between 0.1 and 0.2.

With the pagerank and high degree node attacks are respectively, more or less, 0.3 and 0.4. But as can be expected the network is more robust against random attacks: in order to destroy the giant component it's necessary to remove about the 80% of the node in the network ($f_c = 0.8$).

From point of view of the diameter, the pagerank and the high degree have a significantly impact on the diameter: diameter increases a until the network is made up of isolated nodes.

As said before the networks has very few nodes with a relevant betweenness so it is necessary to remove some of them to remove completely the communication between the nodes, the diameter doesn't even increases as the other attacks. This means that the network relies on very few nodes to keep together the various communities.

the a random attacks are the one that affect less the diameter, it doesn't even increase too much but it is necessary about 80% of the nodes to destroy the network. As expected, the network is robust against random attacks because it follows a power law distribution.