Network Analysis
Network Robustness

Federico Fontana, Lorenzo La Corte

May 24, 2023

Chapter 1

Introduction

A network is robust if it is resilient to fauilures. This report presents multiple experiments where fauilures are simulated removing nodes through different approaches and strategies.

The goal is to test the different combinations of:

1. removal strategies,
2. chosen topologies,

In order to understand their relationship and find the most damaging combinations.

1.1 Datasets

This group of datasets includes three wikipedia page-page networks based on three different topics: chameleons, crocodiles and squirrels. Nodes represent articles from the English Wikipedia in December 2018, edges reflect mutual links between them.

This report aim to compare the robustness of these newtorks among them and in respect with other **synthetic graphs**.

The datasets we will take into considerations are:

	Regular	Erdős-Rényi	Chameleon	Squirrel	Crocodile
Nodes	2277	2277	2277	5201	11631
Edges	31878	31421	31421	198493	170918
Density	0.012	0.012	0.012	0.015	0.003

1.2 Observed Metrics

In order to measure the robustness of a network we will take into consideration two main metrics:

1. the size of the giant component, in terms of nodes and edges,
2. the diameter of the graph,

We will show how they change when the fraction of deleted nodes increases.

Chapter 2

Expected Results

We expect:

1. **Random Graphs**, such as Regular and Erdős-Rényi synthetic ones, to show the same level of robustness against both random and targeted attacks.

2. **Scale-Free Graphs** (our wikipedia graphs) to be robust against random attacks but vulnerable to target attacks.

The following analysis aims to prove these hypotesys.

2.0.1 Molloy-Reed Criterion

We expect graphs to follow the **Molloy-Reed Criterion**. This criterion defines a critical threshold f_c, described as:

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$$

If in a network, we remove a fraction of nodes greater than this breakdown threshold f_c, the giant component will break into subcomponents.

For random graphs we have that this threshold can be simplified as:

$$f_c = 1 - \frac{1}{\langle k \rangle}$$

So, we expect:

1. **Random Graphs** to have a finite f_c.

2. **Scale-Free Graphs** to have a f_c greater than the one of the corrispondent random graph. This would imply enhancenced robustness of our Scale-Free graphs against Random Attacks.

Chapter 3

Empirical Results

Since the chosen datasets have different regimes, we expect them to act differently with respect to different attacks. So, in terms of robustness, we aim to compare the robustness of synthetic graphs against the one of our datsets.

To analyze the changes between other parameters, we fixed:

1. The number of removal per round to 20, with an exception for the largest graph,
2. The number of maximum round to infinite,
3. The fact that we are going always to attack the giant component, computing it again at each round.

3.1 General Observations

Before comparing topologies and strategies, these are some general observations.

3.1.1 Breakdown Threshold f_c

A network displays enhanced robustness if its breakdown threshold f_c deviates from the correspondent random network threshold f_c^{ER}.

	Chameleon	Squirrel	Crocodile	Directed Crocodile
f_c	0.987	0.997	0.997	0.997
f_c^{ER}	0.964	0.987	0.966	0.837

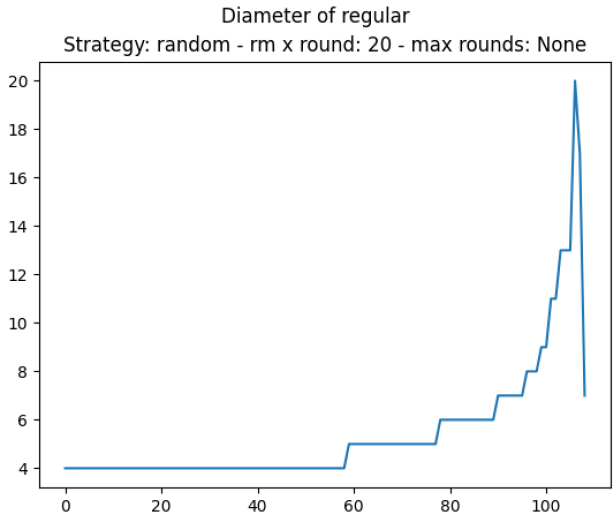
All the datasets show enhanced robustness.

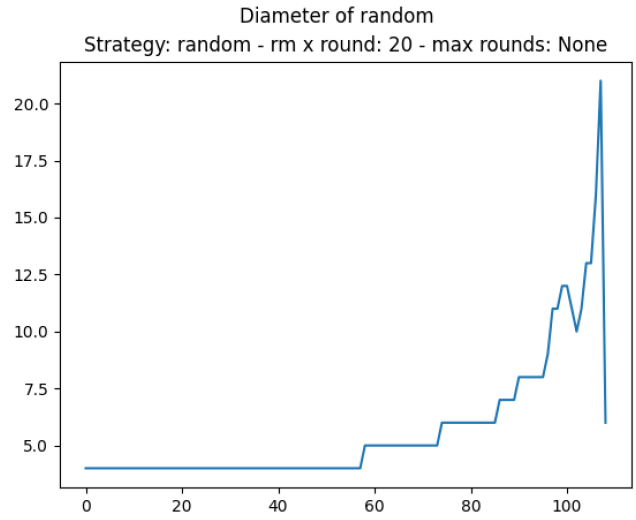
3.1.2 Diameter

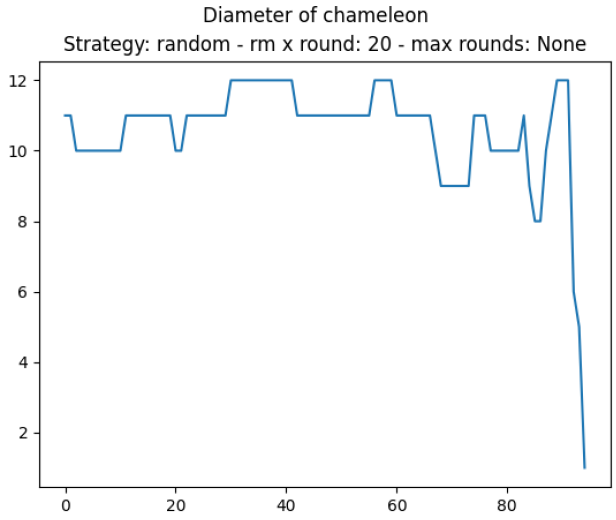
There are many aspects to consider about the evolution of the diameter of a network under attack:

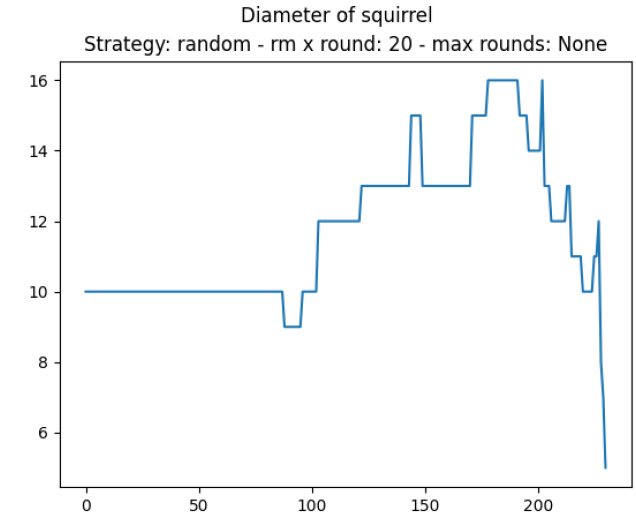
1. **Diameter** value, with respect to the time of the simulation, always goes up before going down, due to the fact that:
   (a) initially, removing nodes means removing edges and so enlarging shortest paths,
   (b) then, once nodes are very few, shortest paths becomes smaller as the size of the graph decreases.
2. **the evolution of the diameter** in random or scale-free graphs:
   (a) doesn't differ significatively in response to a centrality-based attack.
   (b) has a huge difference if the attack is random-based, in particular:
      i. a scale-free network is more resilient against random attacks and so its diameter stays the same until the final part of the simulation.
      ii. a random network responds in the same way to a random or target attack,

Here we can observe an example of the evolution of the diameter with a random attack *(proof of i)*:

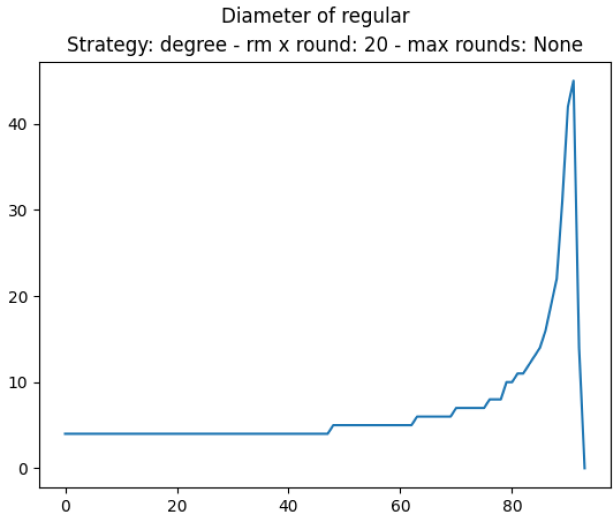


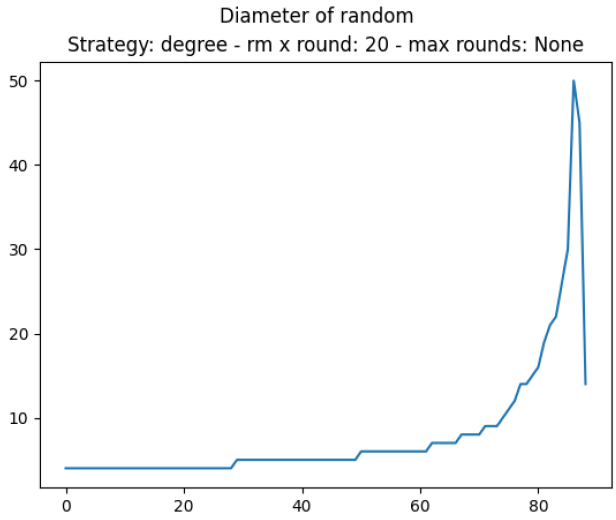


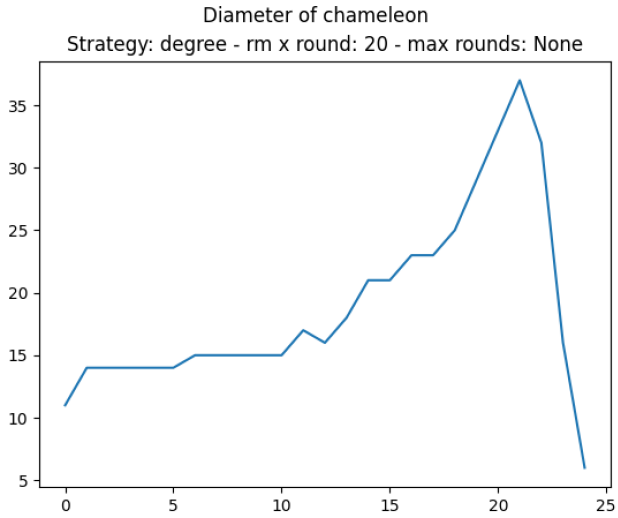


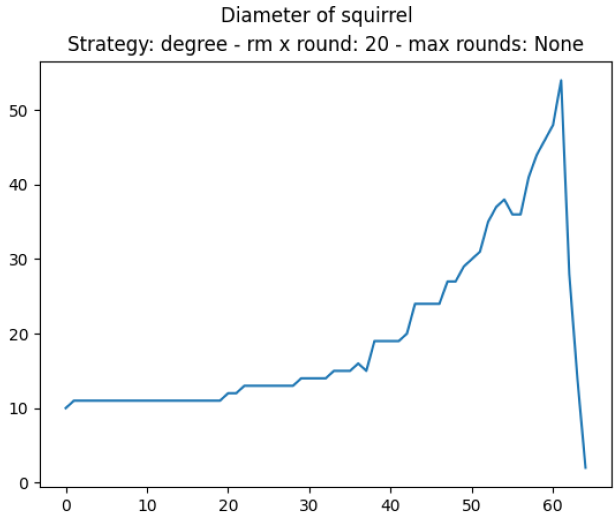


On the other hand, in the plots below we can observe an example of the change of the diameter with a target attack (degree) *(proof of ii)*:







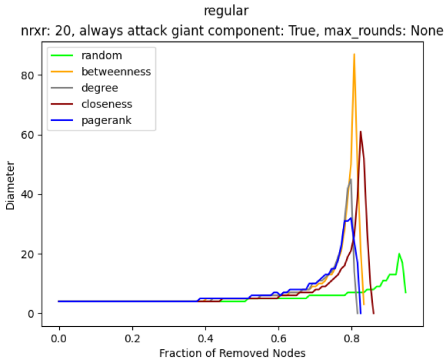


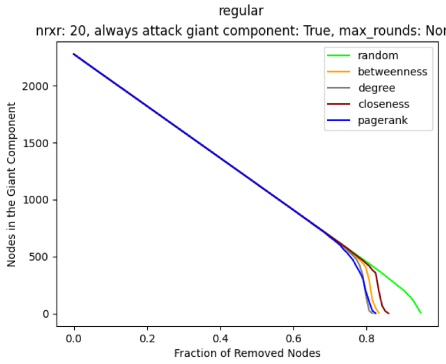
3.2 Synthetic Graphs

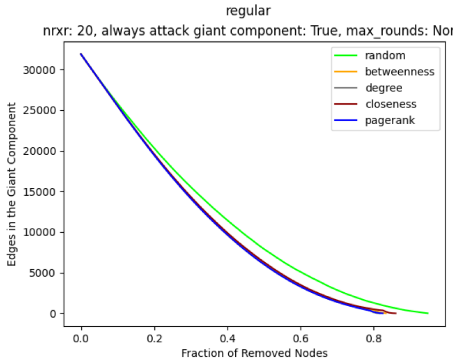
3.2.1 Regular Graph - Regular Version of Chameleon

In this section we are using a regular graph with roughly the same number of nodes and edges of **Chameleon**.

These are the results involving robustness:







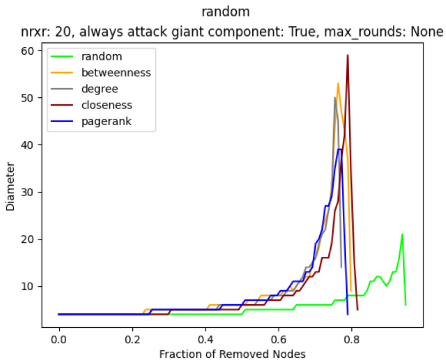
We can observe some interesting facts.

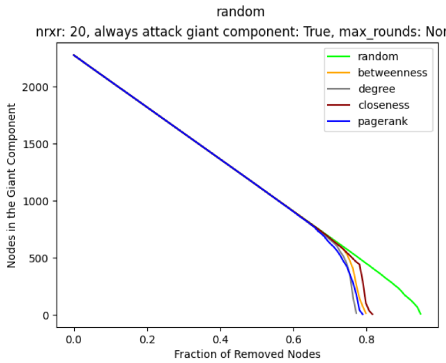
The evolution of the size of the giant component approximately follows a straight line. Although, almost at the end, the network is so broken that doens't act anymore as random and so target attacks are more effective.

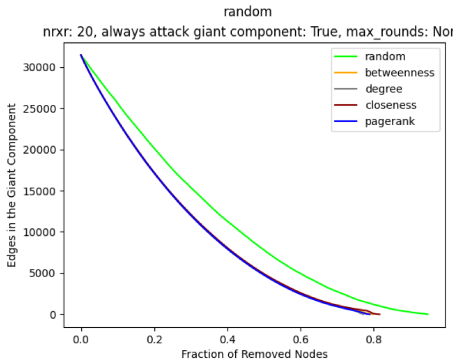
We can also appreciate that the chosen strategy generally doesn't influence the trend of the different curves. There's an exception for the random attack which seems slightly less effective at the end of the simulation.

3.2.2 Erdős-Rényi - Random Version of Chameleon

In this section we are using an Erdős-Rényi random graph, which again, has roughly the same number of nodes and edges of **Chameleon**.





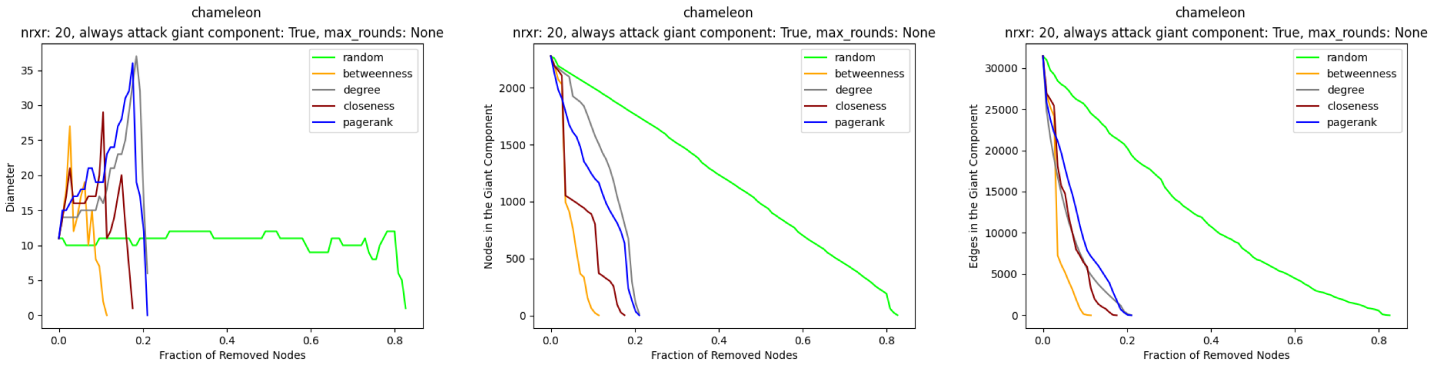


As expected, results regarding robustness are the same as the ones for the regular graph.

3.3 Real Graphs

3.3.1 Chameleon

Our analysis is now focused on **Chameleon**, our smallest graph. This should roughly approximate a scale-free regime, albeit having a small amount of nodes.



As expected, the size of the giant component:

1. decreases linearly using a random-based attack,
2. decreases significatively using a centrality-based attack.

And the same applies for the diameter that:

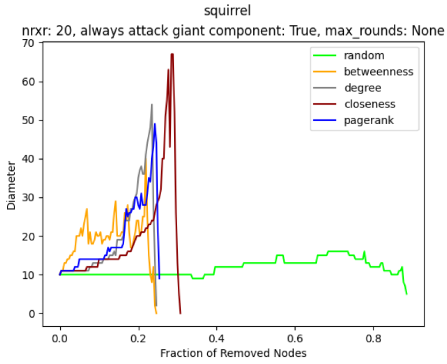
1. is not really affected by random-based attack, until the very end of the simulation
2. has a very rapid evolution with a centrality-based attack, especially betweenness-based attacks

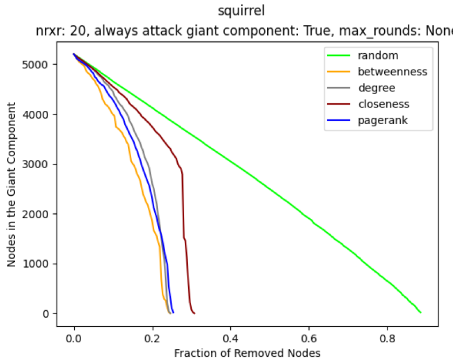
This is the ranking of the damage of the attacks:

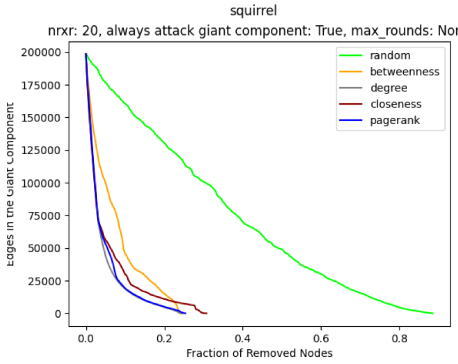
Diameter	Nodes	Edges
Betweenness	Betweenness	Betweenness
Closeness	Closeness	Closeness
Page Rank	Page Rank	Degree
Degree	Degree	Page Rank
Random	Random	Random

3.3.2 Squirrel

Our analysis is now focused on **Squirrel**, which better approximate a scale-free regime, having double of nodes of **Chameleon**.







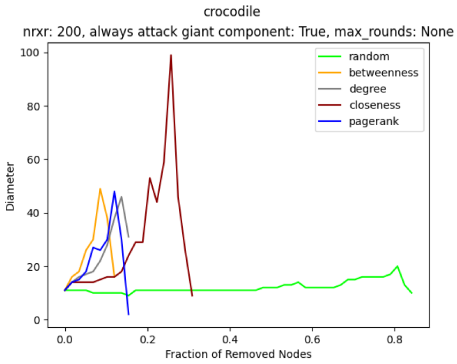
As expected, the size of the giant component and the diamter have an evolution similar to the ones seen for Chameleon.

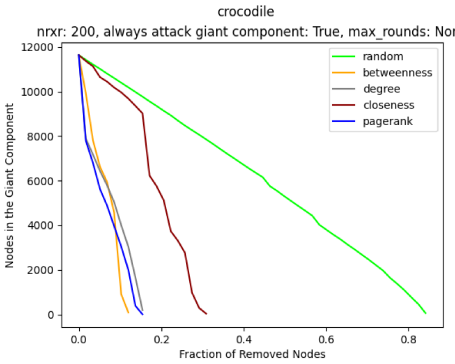
The ranking of the damage of the attacks changes slightly:

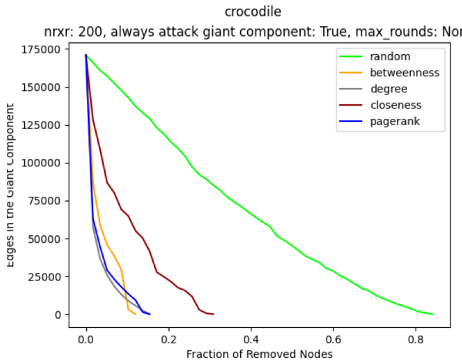
Diameter	Nodes	Edges
Betweenness	Betweenness	Degree
Degree	Degree	Page Rank
Page Rank	Page Rank	Betweenness
Closeness	Closeness	Closeness
Random	Random	Random

3.3.3 Crocodile

Our analysis is now focused on **Crocodile**, the largest of the given graph.







Plots show approximately the same trend as the one seen for chameleon and squirrel, as the scale-free structure is the same. Although, here we can notice that closeness attack is clearly the less effective among the target-based ones.

This is the ranking of the damage of the attacks:

Diameter	Nodes	Edges
Betweenness	Betweenness	Degree
Page Rank	Page Rank	Page Rank
Degree	Degree	Betweenness
Closeness	Closeness	Closeness
Random	Random	Random

3.3.4 Take-Away Points

Our experimental findings provide empirical evidence that supports our hypothesys

With random or regular networks like the synthetic ones we analyzed, the size of giant component and diameter have roughly the same evolution with every kind of attack. With scale-free networks like the wikipedia ones we analyzed, the size of giant component and the diameter are deeply affected, both in terms of edges and nodes, by target attacks.

We can also conclude that, in the wikipedia datasets, as the number of nodes in the datasets increases:

1. random removal strategy impact doesn't change significatively,
2. target attacks are more effective.

This is probably because, as we observed in the first assignment, **as N increases, the network shows more clear evidences of its lack of scale**.

Within the observed experiments, as expected, **betweenness attack is the most effective**, while **random attack is the less effective**. Among the target attacks, for our datasets, the less effective strategy seems to be closeness removal.