Assignment No. 3

Network Robustness

Network:

- Twitter Social Circles

Group Members:

- Parsa Moslem (5755015)
- Amir Mohammad Azimi (5795736)

Course:

- Network Analysis

The course is part of the below degree:

- Computer Science – Software & Security Engineering

Academic Year: 2023 - 2024

UNIVERSITÀ DEGLI STUDI
DI GENOVA

Dibris

Table of Contents

Introduction

This assignment consisted of three steps.

1. The first step involved investigating the impact of various attack strategies on two well-known graph models: the ***Erdós-Rényi*** model and the ***Barabási-Albert*** model. These models, previously introduced in coursework, were readily available through the *NetworkX* library. Four distinct attack methods were employed:
    - **Random node** attack
    - **Highest degree node** attack
    - **Highest PageRank node** attack
    - **highest betweenness node** attack
2. The second step focused on the significantly larger network previously utilized in Assignments 1 and 2 and attacking the mentioned methods on the network.
3. By Leveraging the results obtained in step 2, we further investigated the attack method that exhibited the most significant influence on network robustness in step 3. Finally, efforts were directed towards enhancing the overall robustness of the network.

The code consists of two files:

- **helper.py**: it is used for developing the helper functions which are needed through this assignment.
- **network_robustness.ipynb**: This file is used to work on the assignment itself.

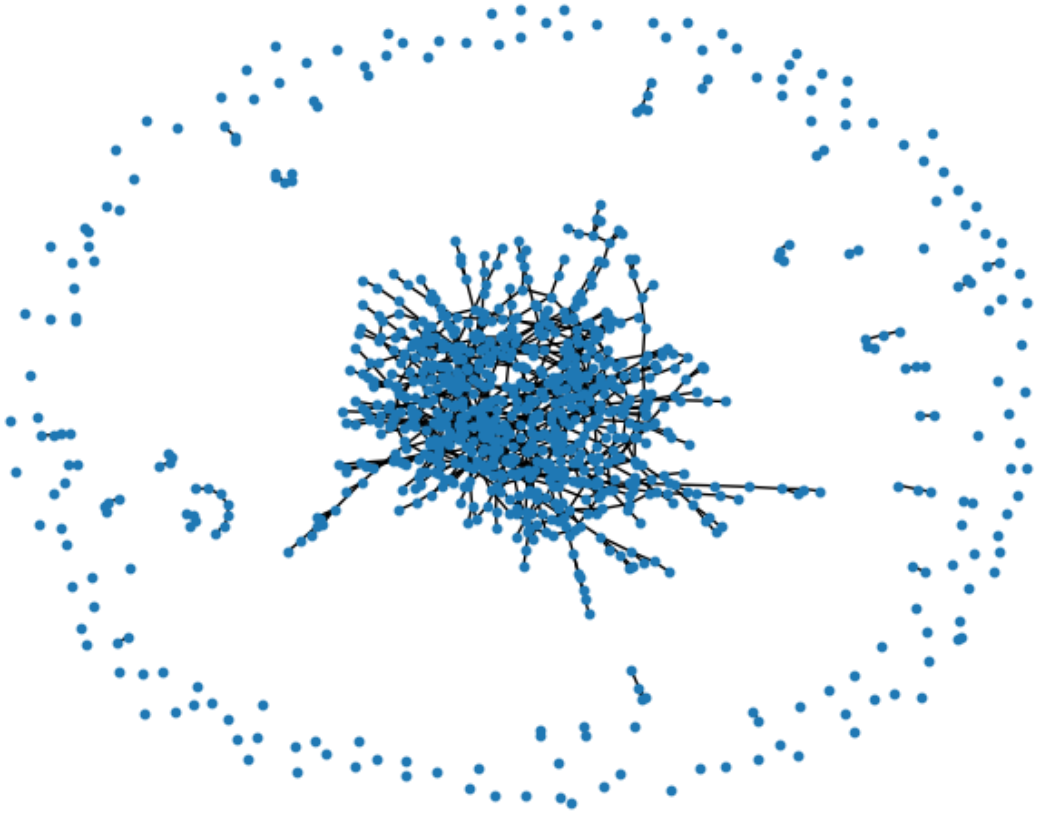
Step 1: Attack on two famous random graphs

The four attack methods are declared in the helper.py file and they are used in “perform_attacks” function which is declared in .ipynb file, and this function is written to handle the logic of step 1 and the visualization of attacking process effect on the graph.

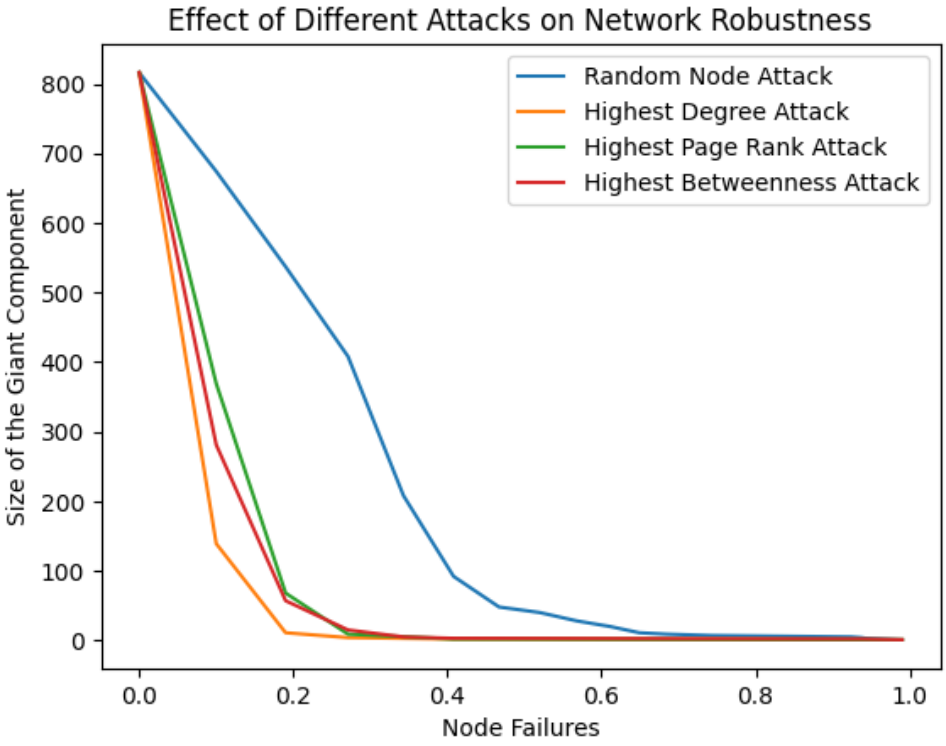
First graph: Erdos-Renyi graph

This random graph is generated using **1000 nodes** and **0.002 probability** of edge creation.

Graph visualization



Performing the attacks

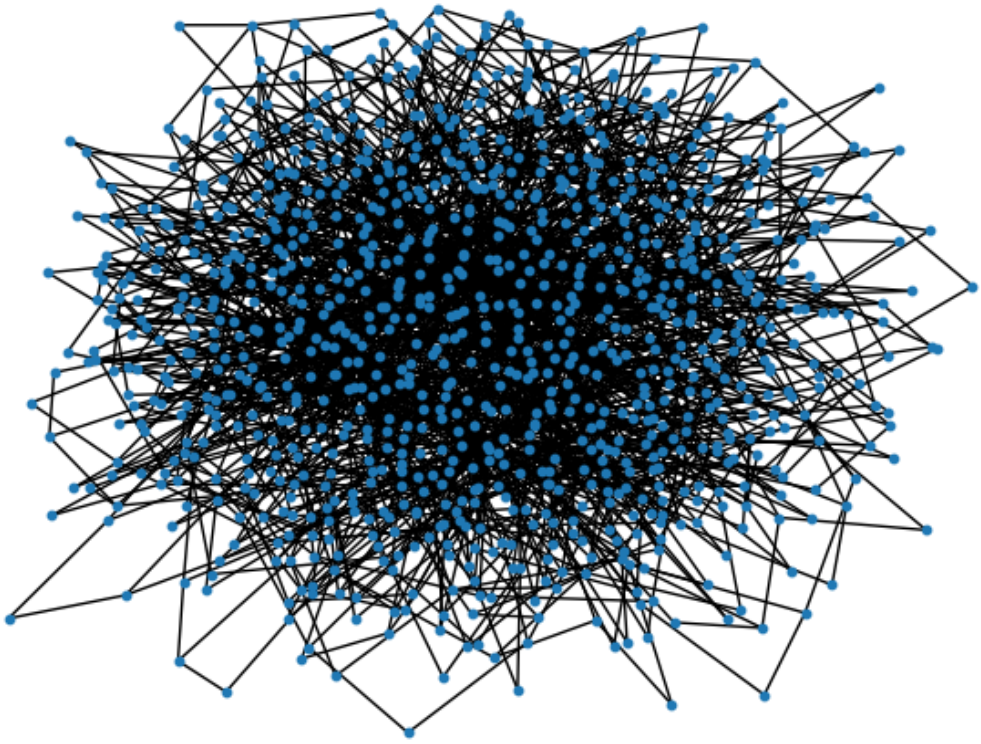


Our analysis revealed that the "**Highest Degree Node**" attack strategy had the most significant impact on the network's robustness by disrupting its giant component. This finding suggests a heightened vulnerability within the network structure. Specifically, the presence of highly connected nodes ("hubs") occupying central positions appears to be a critical factor. The removal of these hubs during the attack process resulted in a fragmentation of the giant component. Furthermore, the removal of a single high degree node can trigger a cascading effect, where the disconnected neighbors, who might themselves be crucial connections for other nodes, are subsequently isolated, leading to further fragmentation.

Second graph: Barabási-Albert graph

This random graph is generated using 1000 nodes and 2, as the number of edges to attach from a new node to existing nodes.

Graph visualization



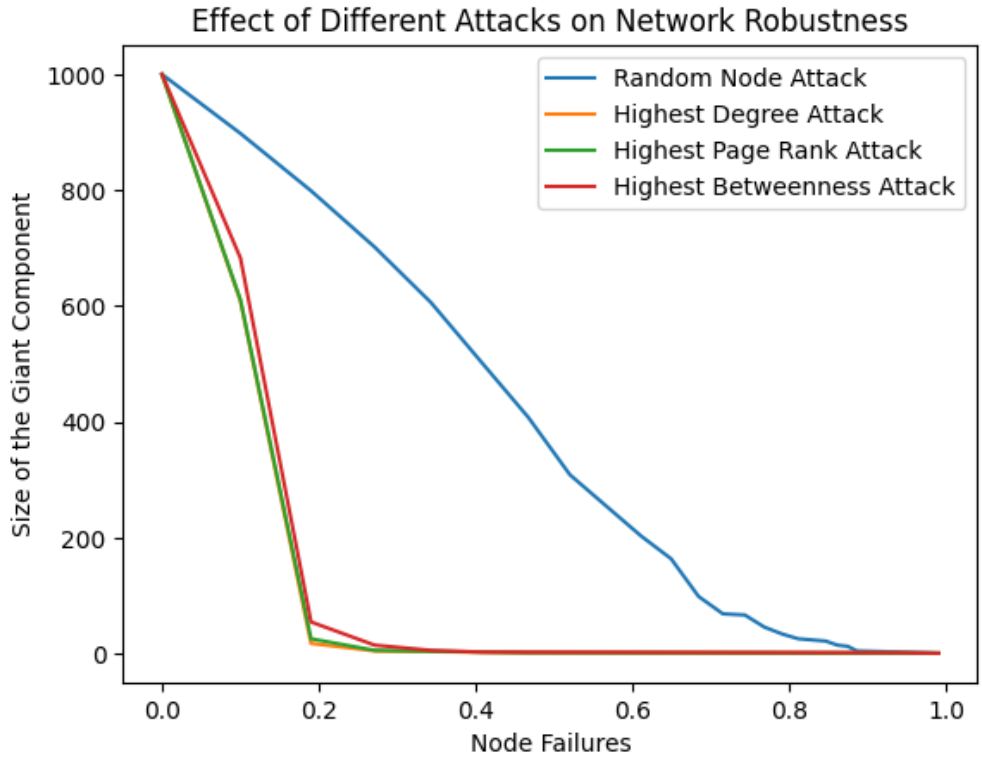
By initial analyzing of this graph, it is achievable that all the three attack methods have had the same effect on the network. As this type of graph is scale-free, meaning that it has small number of nodes have a very high number of edges while most nodes have relatively few connections. Because of this specification of the graph, we have some hubs that are acting as bridges between different parts of the network. But why these three attack methods still have shown almost the same behavior?

All the three attack methods tend to target the same nodes, which are the hubs. In the below section, it is written how each attacking method made an effect on the graph:

- **Highest Degree Attack:** This attack directly removes nodes with the highest number of connections, which can rapidly and easily break the network and its giant component into smaller parts.
- **Highest PageRank Attack:** As it was mentioned in the course, this is a centrality measure (which initially introduced by Google), considers both the number and quality of node's edges. In Barabási-Albert graph, hubs naturally have high PageRank due to their extensive connections with other nodes. So, this attack also targets hubs, just like the Highest Degree Attack.
- **Highest Betweenness Attack:** This centrality feature measures a node's importance in bridging paths between other nodes. In this graph, the hubs are

often lie on the shortest paths between many other nodes. So, this attack also tends to target hubs.

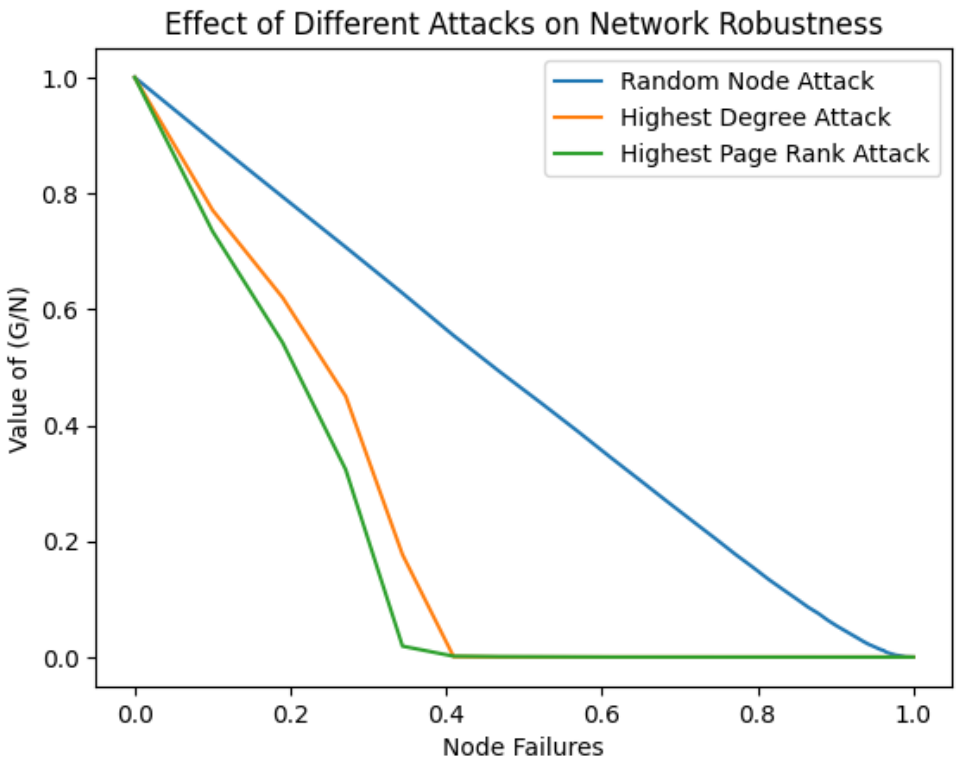
Performing the attacks



Just as the first graph (Erdos-Renyi), this attacks on the Barabási-Albert also might have cascading effect that whenever a hub is removed, it might affect other connected nodes and make them also isolated components that will be ignored in the second attack (if the giant component is considered in the attack algorithm).

Step 2: Attack on the network of Twitter's Circles

This stage of the analysis excluded the highest betweenness centrality attack due to its computational demands associated with betweenness calculations. Consequently, only three alternative attack scenarios were simulated on the network. As previously observed, the random node attack exhibited the minimal disruptive impact on network robustness compared to attacks targeting nodes with the highest degree centrality and PageRank centrality.



As random node attack has the **least impact**, it suggests the network has some level of robustness, which is also mentioned in the course that **scale-free networks are generally robust against the random node attacks**, with the $N \to \infty$, f_c is equal to 1, which means the network is extremely robust against random attacks. Even if random users are removed, the core functionality might not be significantly affected. The consideration of alternative attack scenarios targeting nodes with high degree centrality and PageRank centrality suggests the presence of critical nodes within the Twitter Circles network, which are probably the creators of circles. Nodes exhibiting high degree centrality likely possess a significant number of connections to other users, functioning as hubs within the information flow. Conversely, nodes with high PageRank centrality, while potentially possessing fewer connections, exert a strong influence due to the prominence of users connecting to them. Targeted attacks on these critical nodes would demonstrably disrupt network functionality to a greater extent compared to attacks on randomly chosen nodes.

Step 3: Building robustness on the network of Twitter's Circles

In the final step, as we understood that the highest PageRank attack is the most effective one on the disruption of the network, we will work on this area to make the network more robust against this type of attack.

Algorithm

This algorithm designed to improve the resilience of complex networks by strategically connecting highly influential nodes. The core principle lies in leveraging the PageRank metric to identify critical nodes within the network and subsequently augmenting their connectivity. This approach aims to bolster the network's robustness against potential disruptions, such as highest PageRank targeted attack, which had the most influence on the network robustness.

The algorithm commences by calculating the PageRank score for each node within the network. Subsequently, these scores are compiled and arranged in descending order, effectively forming a ranked list that prioritizes nodes based on their influence. The average degree of the network is then computed, which serves as a parameter for the critical threshold calculation (denoted as f_c).

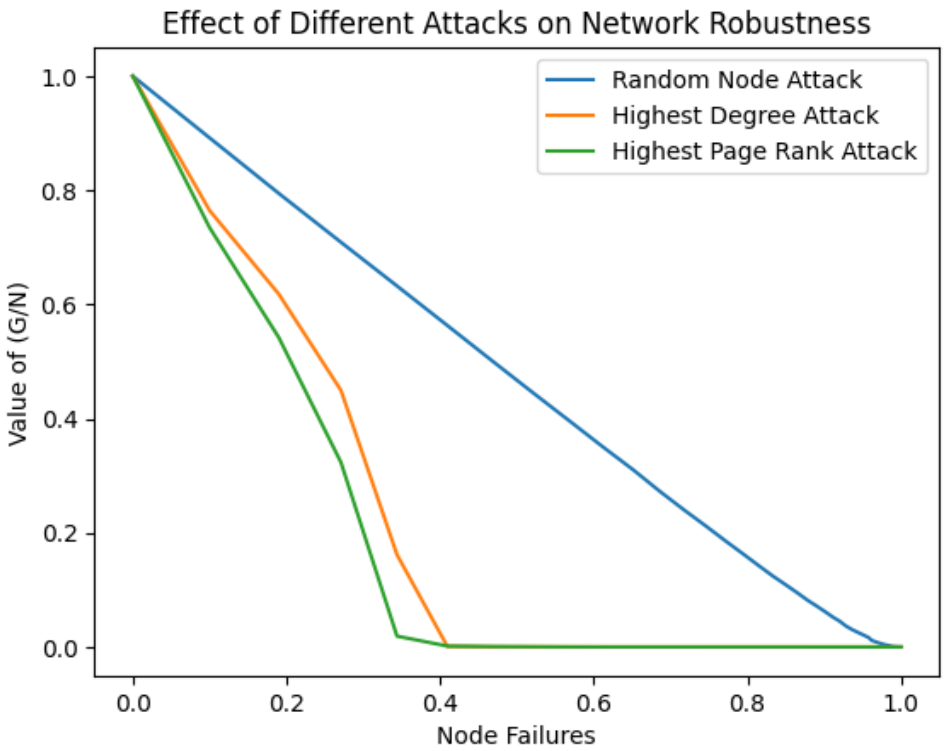
The **critical threshold**, a pre-defined value representing the network's susceptibility to disruptions before the robustness enhancement (**0.9967** in this instance), serves as a benchmark for evaluating the effectiveness of the algorithm.

Following the identification of critical nodes based on the top 'f' fraction of the ranked PageRank list (f = 0.1 in this specific case), the algorithm establishes connections between these influential nodes and a maximum of five non-connected nodes within the network. This targeted edge creation aims to strategically increase the connectivity of these critical nodes, potentially enhancing the network's overall robustness. However, the selection of five connections remains a limitation of the current implementation, as it relies on a pre-determined value rather than a dynamically determined threshold based on network properties. Future iterations of the algorithm should incorporate a more data-driven approach for selecting the optimal number of connections.

The **critical threshold** is then recalculated after implementing the network modifications (**0.9976**). The critical threshold (f_c) for the emergence of a giant component in the network can be elevated by $\Delta \boldsymbol{f_c} = \mathbf{0.0009}$. This improvement is particularly relevant for **large-scale (N >> 1) scale-free networks**, where f_c is known to asymptotically approach unity.

Finally, the algorithm returns the modified network structure, enabling visualization of the changes. However, the minimal improvement observed in this specific case highlights the need for further refinement to ensure a more visually discernible enhancement in network robustness.

The visualization is illustrated in the below section:



We now investigate the critical threshold (f_c) of the network in relation to another model: a random network with an identical number of nodes and edges. The following results are presented:

Network	***Random Failure (Circles of Twitter)***	***Random Failure (Random Network)***
Circles of Twitter	0.9976	0.9714

As the critical threshold for the network of Twitter's circles is bigger than random network threshold, **the network shows enhanced robustness ($f_c > f_c^{ER}$)**, meaning that the network can maintain its functionality and integrity even when faced with disruptions or attacks.

Conclusion

This assignment provided valuable insights into the impact of diverse network attacks, encompassing both random and targeted variants, on network structures. The observations demonstrate a strong correlation between network vulnerability and its underlying architecture. Furthermore, the assignment explored mitigation strategies,

equipping us with methodologies to enhance network robustness against specific attack types that demonstrably inflict significant disruption.