YZV413E - Graph Theory and Algorithms Project

Mert Gülşen - 150200332, Ahmed Selim Çeleğen - 150200304 12/06/2023

In this project, the paper "Graph Vulnerability and Robustness: A Survey" is studied.

Robustness measures

3 different robustness measures from the paper are selected for analyzing robustness of graphs. There are 3 different categories for robustness measures. These are graph measures, adjacency measures and Laplacian measures. Selected measures from each category is given below:

- Global clustering coefficient
- Spectral gap
- Algebraic connectivity

Global clustering coefficient : The global clustering coefficient (C) is calculated using the formula:

C = 3 * (number of triangles in the network) / (number of connected triples in the network)

In this formula, the number of triangles refers to the number of sets of three nodes that are fully connected, and the number of connected triples represents the number of sets of three nodes that are connected in any way (either directly or indirectly).

Higher global clustering coefficient means better robustness for the graph.

Spectral gap : The spectral gap is a measure used in graph theory to assess the connectivity or the presence of disconnected components in a network. It is calculated as the difference between the first and second largest eigenvalues of the adjacency matrix of the network.

The formula for the spectral gap (Δ) is:

$\Delta = \lambda_1 - \lambda_2$

Where λ_1 is the largest eigenvalue and λ_2 is the second largest eigenvalue of the adjacency matrix. A larger spectral gap indicates a more connected network, while a smaller spectral gap suggests the presence of disconnected components or subgraphs within the network.

Higher spectral gap means better robustness for the graph.

Algebraic connectivity : The algebraic connectivity is a measure used in graph theory to quantify the connectedness or cohesion of a network. It is determined by the second smallest eigenvalue of the Laplacian matrix of the network.

The formula for the algebraic connectivity (λ) is:

$\lambda = \lambda_2$

where λ_2 is the second smallest eigenvalue of the Laplacian matrix. The algebraic connectivity provides information about how well the network is connected, with a larger algebraic connectivity indicating a more connected network. A value of zero suggests the presence of disconnected components or subgraphs within the network.

Higher algebraic connectivity means better robustness for the graph.

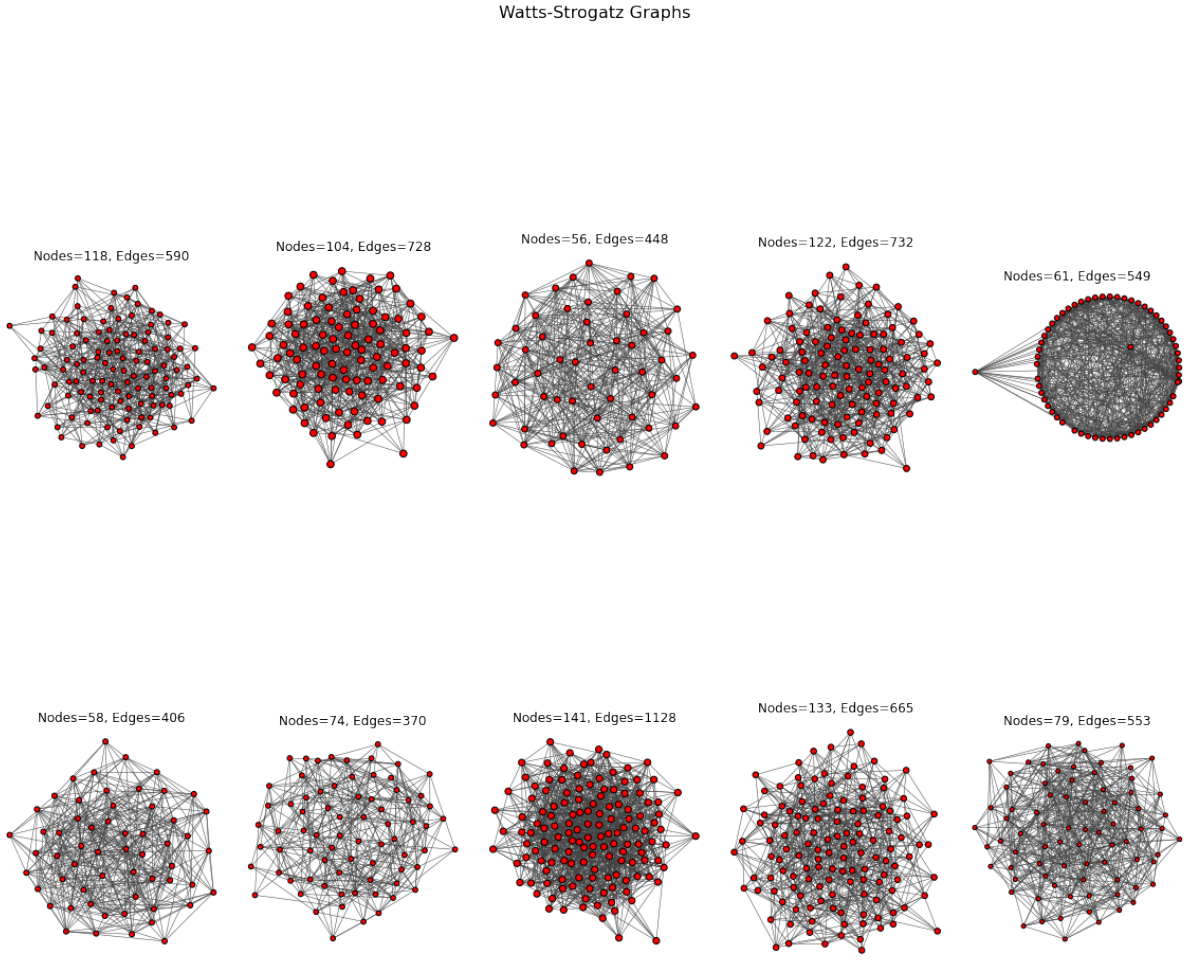
1 Random graph models

To test robustness and do experiments, random graph models are produced. 4 random models are used and 10 graphs are created for each model. These models are:

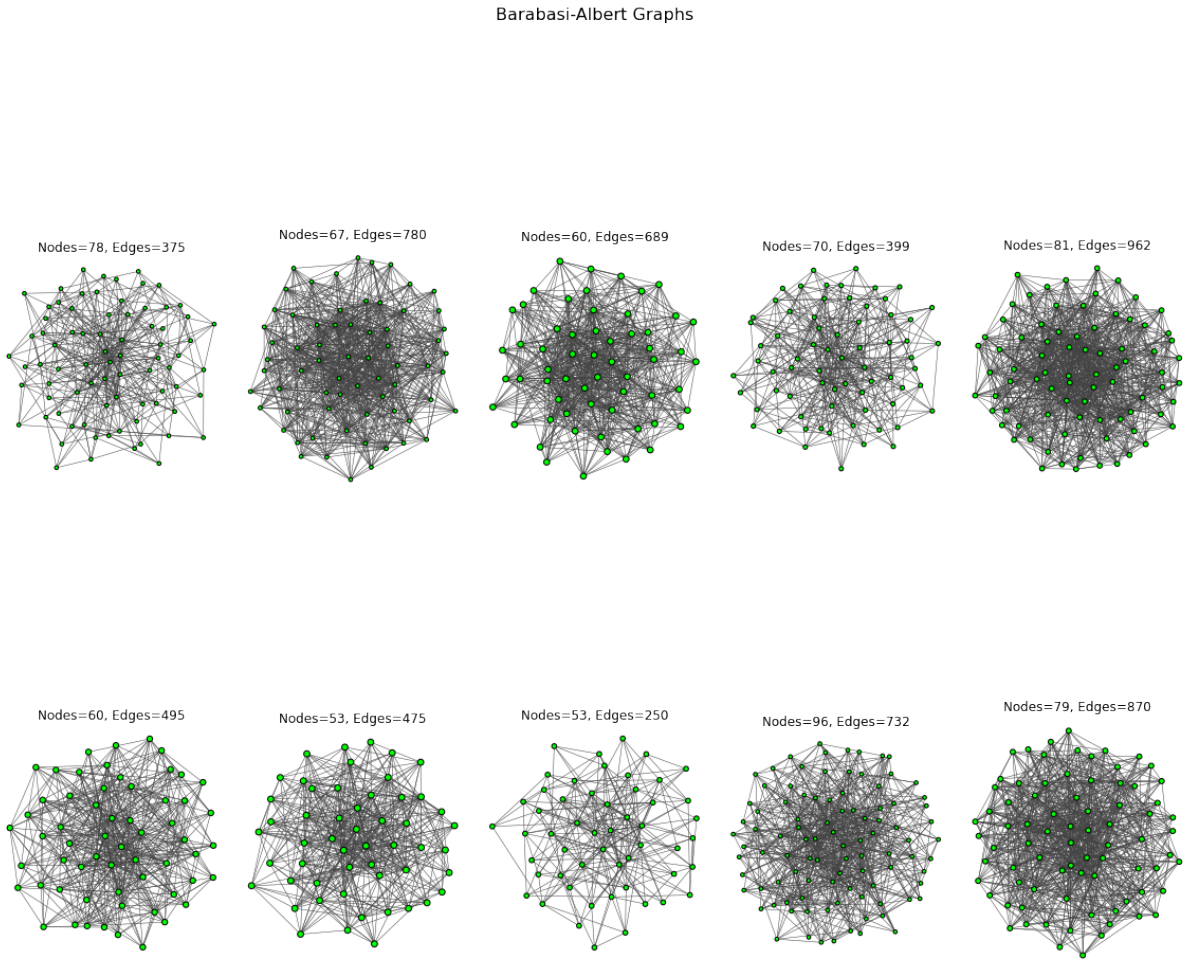
- Erdos-Renyi model
- Watts-Strogatz model
- Barbasi-Albert model
- Random exponential model

Erdos-Renyi model graphs:

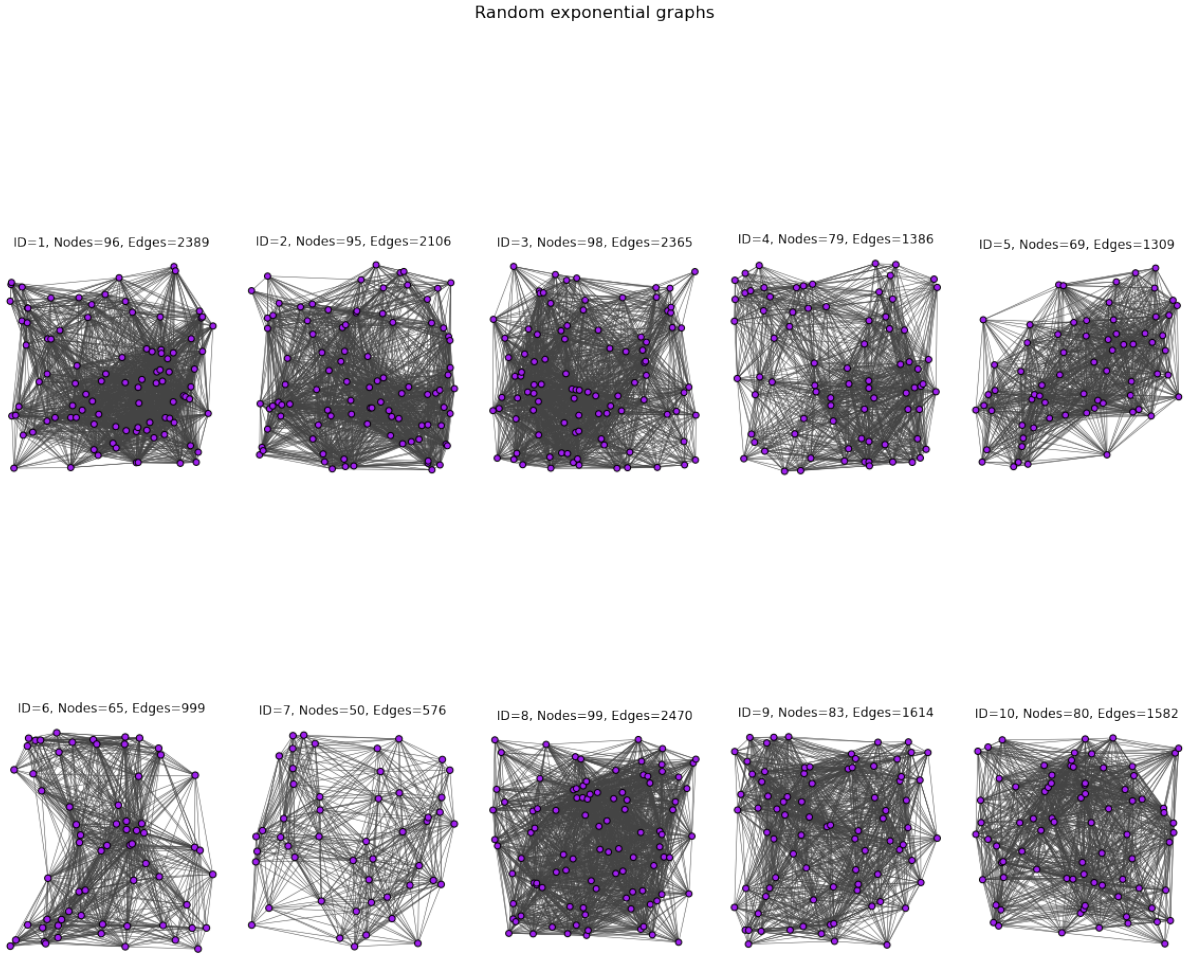
Watts-Strogatz model graphs:



Barbasi-Albert model graphs:

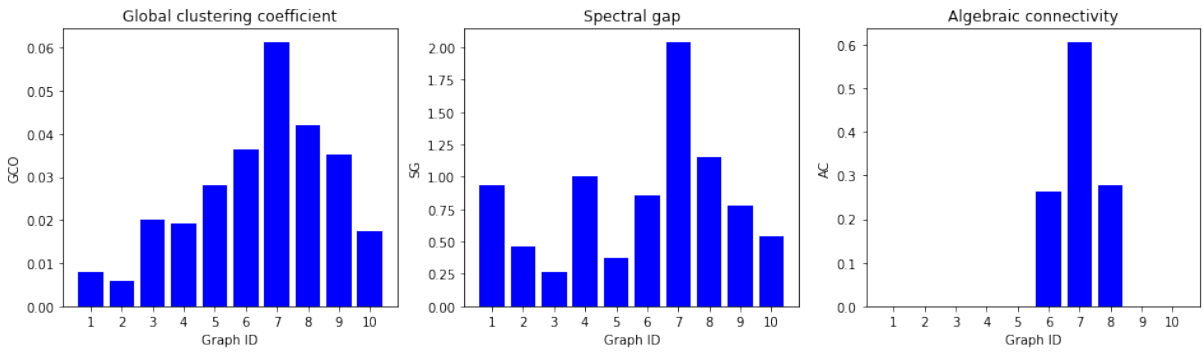


Exponential model graphs:

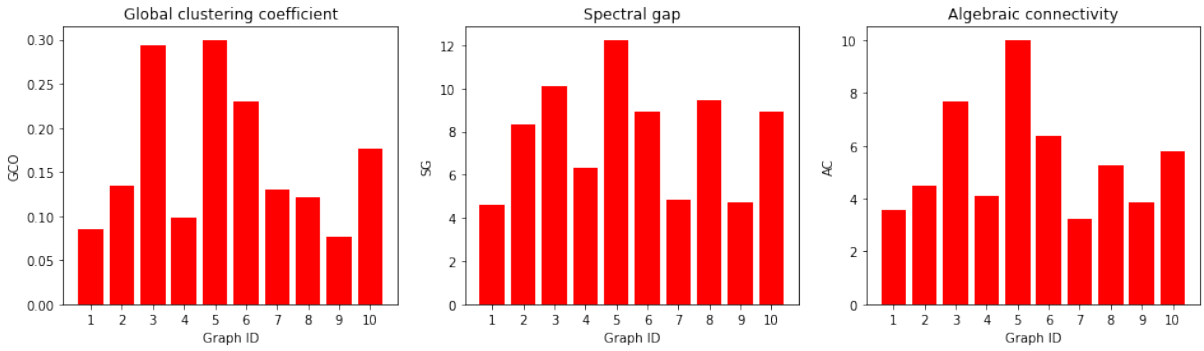


Previously selected robustness measures are calculated on these graphs and results are plotted:

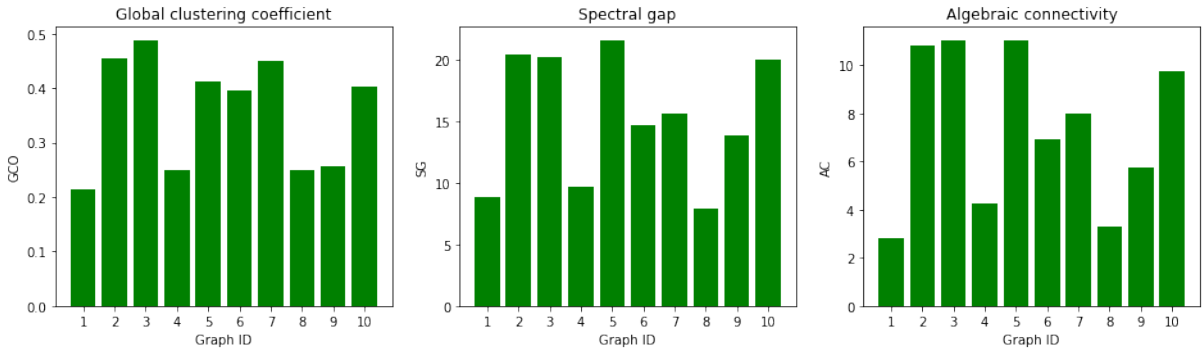
Erdos-Renyi graph robustness measures:



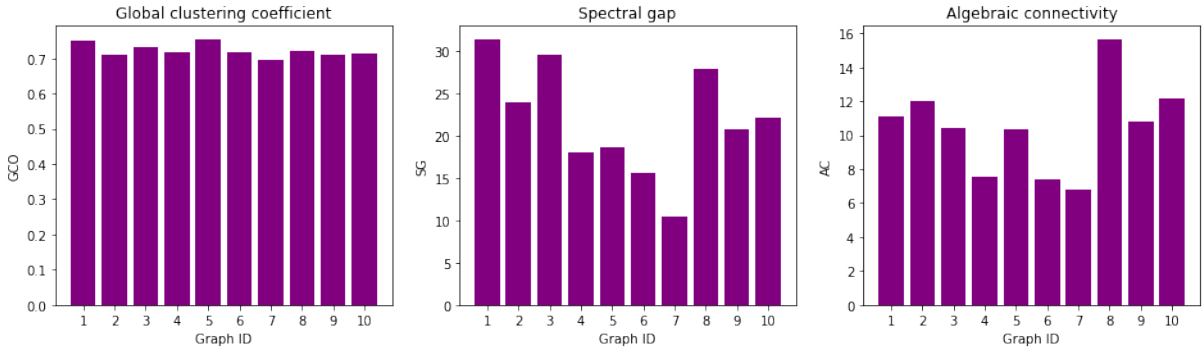
Watts-Strogatz graph robustness measures:



Barbasi-Albert graph robustness measures:



Exponential graph robustness measures:



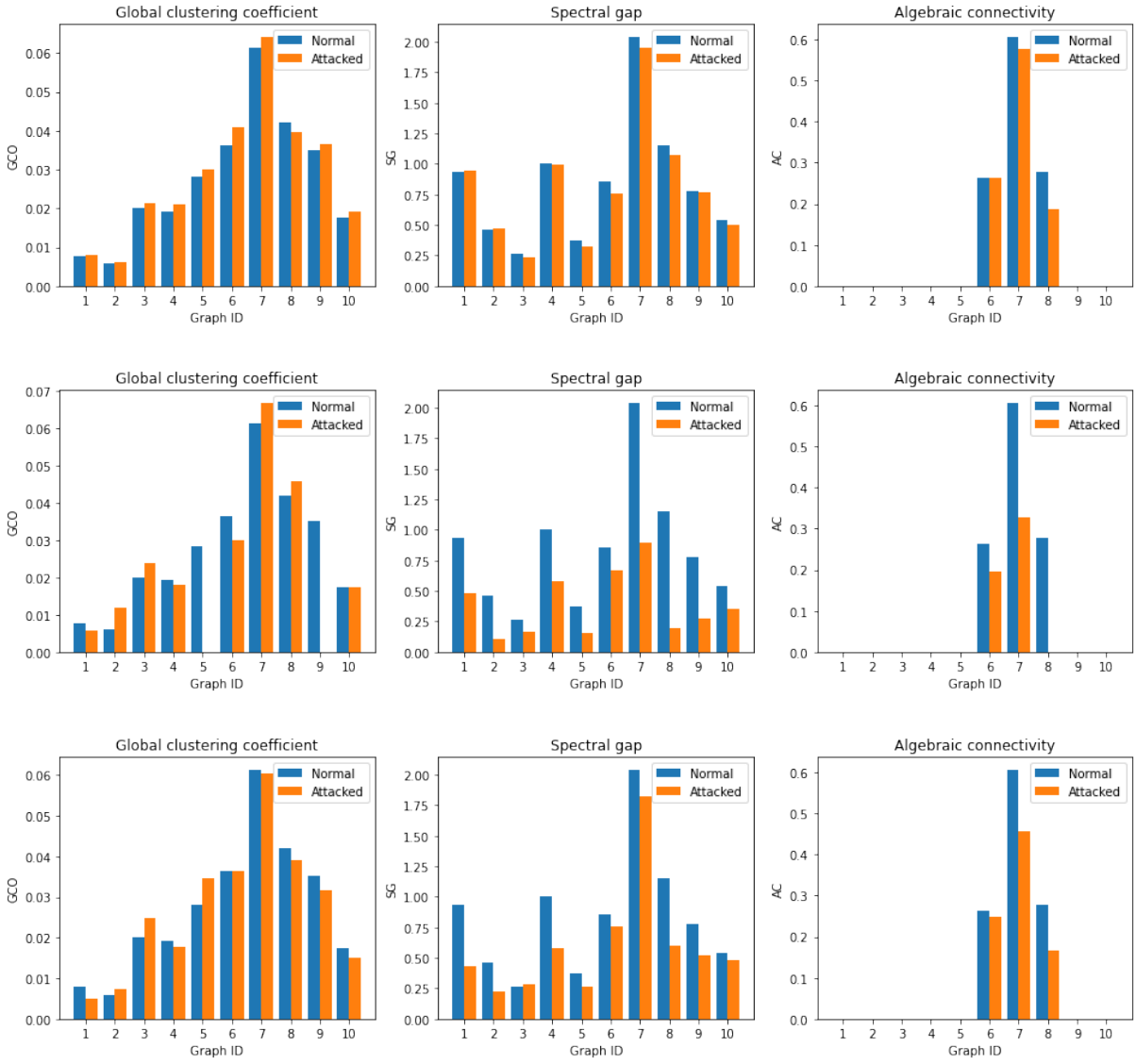
2 Attacks on random model graphs

3 different attack strategies are selected for testing the robustness of graphs. 2 of these attacks are targeted attacks and other one is a random attack strategy. These attacks are:

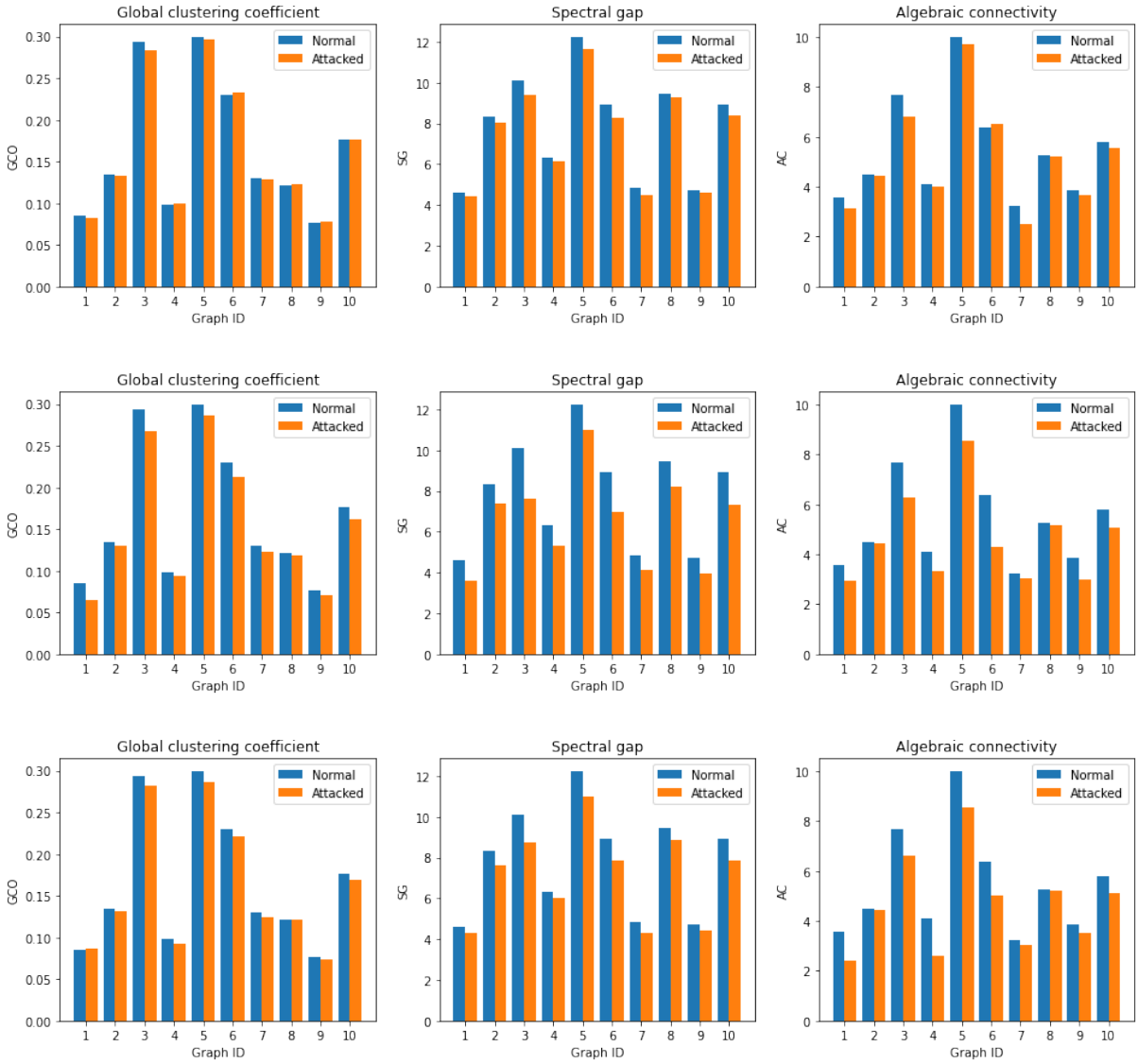
- Random attack In this attack we choose a random node, remove it and its edges from the graph.
- Recalculated degree removal attack In this attack we choose the node with highest degree value, remove it and its edges from the graph. If multiple nodes will be removed then degree calculation made each step.

- Recalculated betweenness removal attack In this attack we choose the node with highest betweenness value, remove it and its edges from the graph. If multiple nodes will be removed then betweenness calculation made each step.

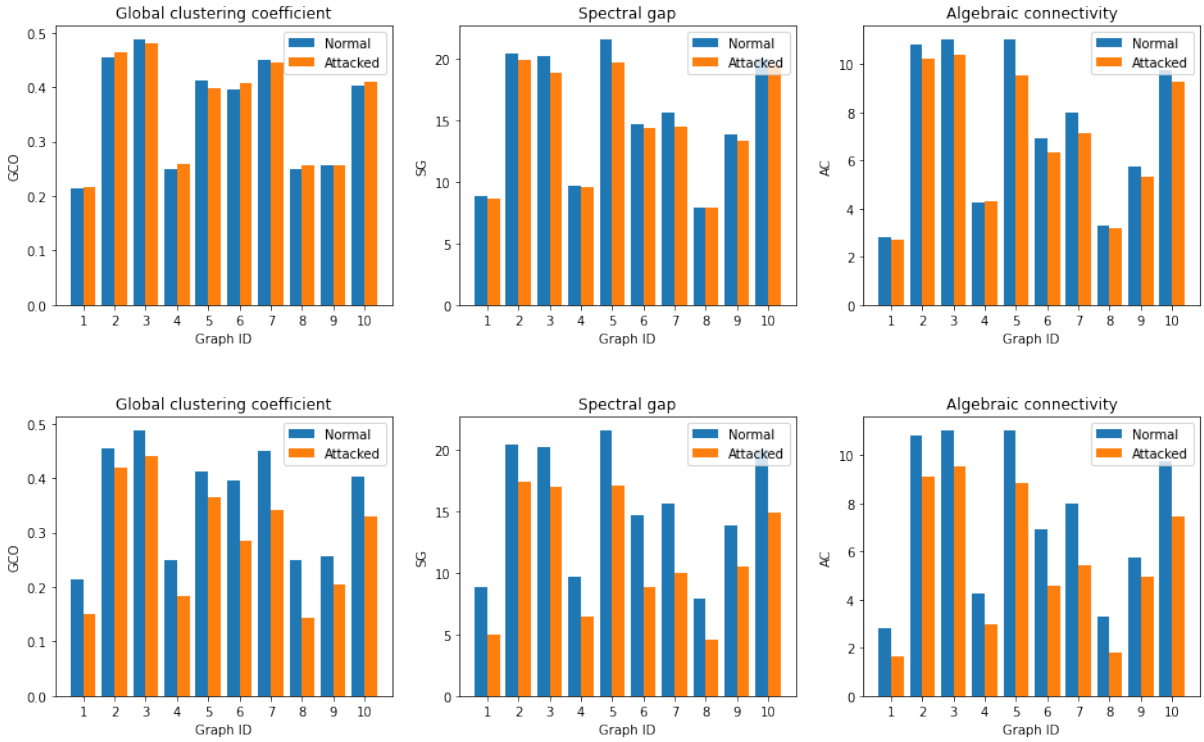
Random, degree and betweenness removal attack results for Erdos model, respectively:



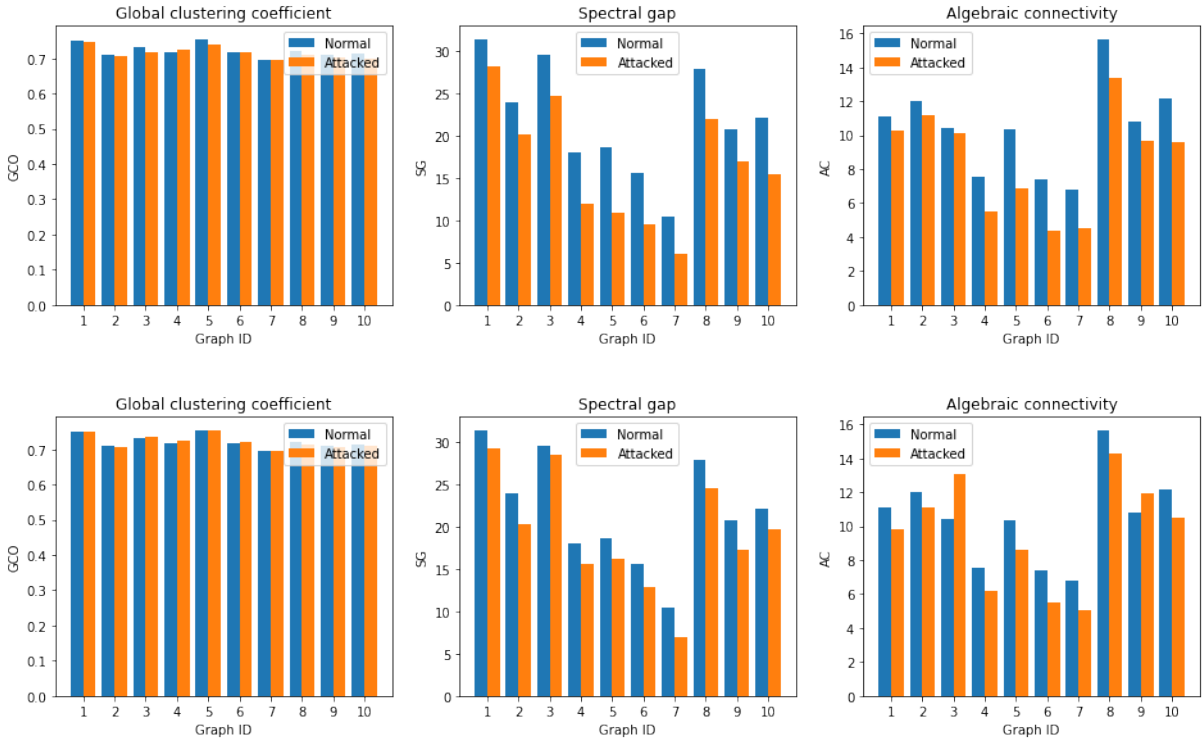
Random, degree and betweenness removal attack results for Watts model, respectively:



Random and degree removal attack results for barbasi model, respectively:



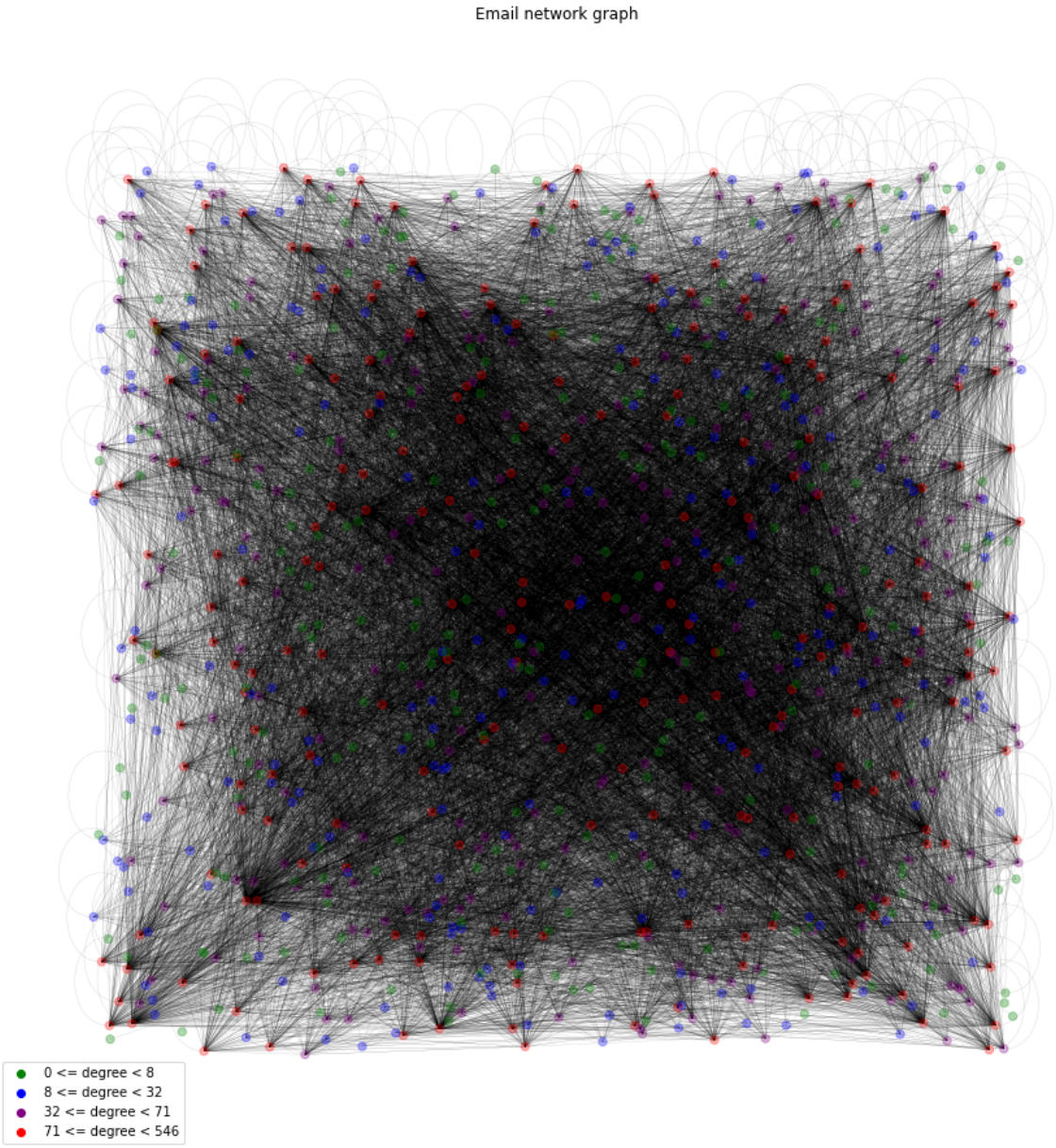
Degree and betweenness removal attack results for exponential model respectively:



3 SNAP library graph datasets

To test robustness of real-world graphs, 2 different graph datasets from SNAP library are used. These are email network dataset and facebook social network dataset.

Email network graph visualization and properties:



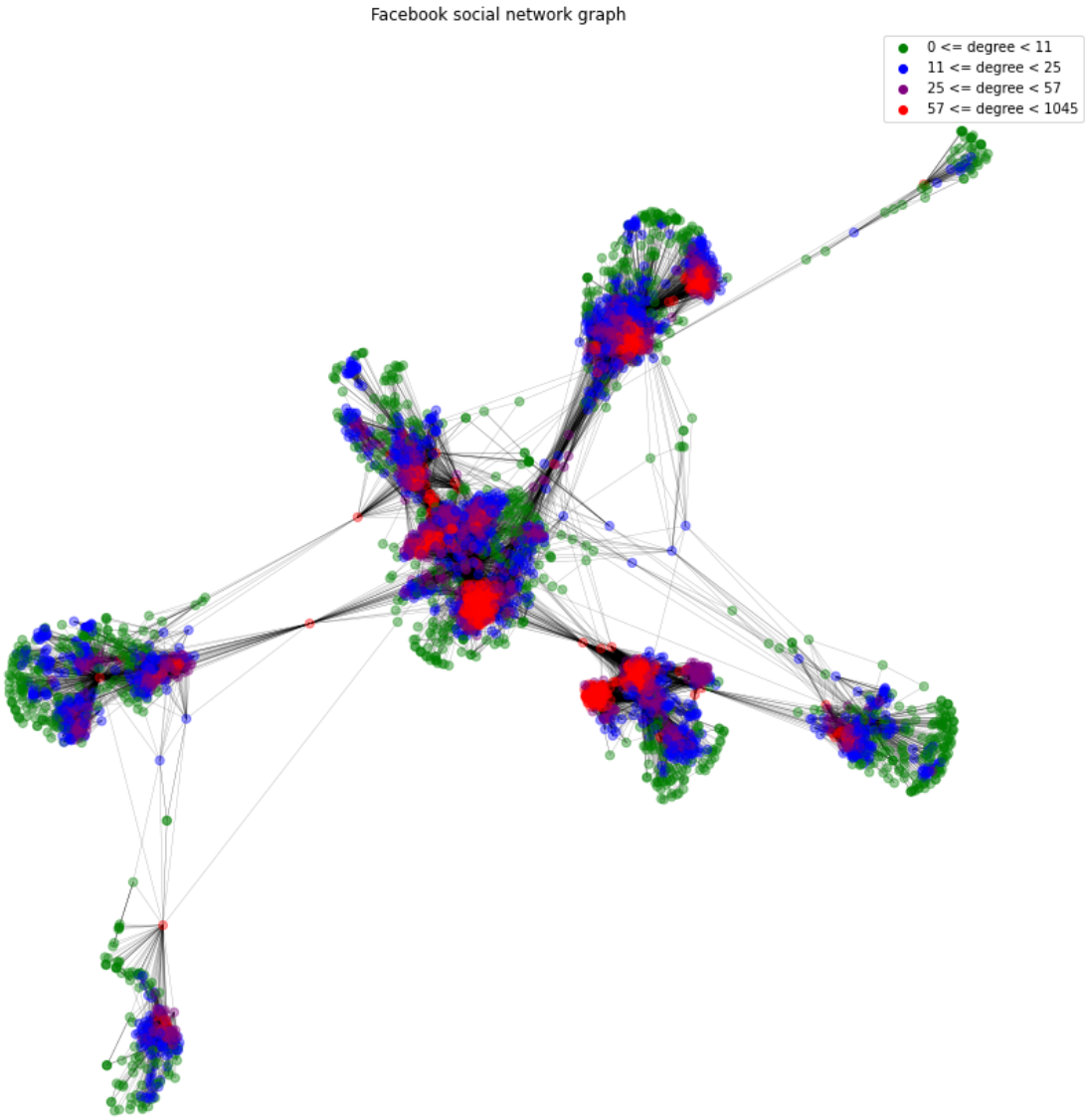
Robustness measures of graph:

Global clustering coefficient : 0.26739242877040204

Spectral gap : 62.02552757374946

Algebraic connectivity : 0.6313733566198104

Facebook social network graph visualization and properties:



Robustness measures of graph:

Global clustering coefficient : 0.5191742775433075

Spectral gap : 36.8807403746528

Algebraic connectivity : 0.028988033385845668

4 Attacks on SNAP library datasets

Previously specified 3 attack methods are applied to email network graph and robustness measures are recalculated.

SNAP Email network graph attacked measures:
Global clustering coefficient : 0.29649143573117465
Spectral gap : 7.681430531491266
Algebraic connectivity : 0.6207994751926129

SNAP Email network graph normal measures:
Global clustering coefficient : 0.26739242877040204
Spectral gap : 62.02552757374946
Algebraic connectivity : 0.6313733566198104

It is observed that clustering coefficient decreases as expected but there is an increase in spectral gap measure. This can be explained by the definition of spectral gap. Spectral gap is the ability of identifying bottlenecks in the network. By removing high centrality nodes, bottlenecks are reduced therefore spectral gap is higher after attack.

3 attacks are also applied to Facebook social network graph and measures are recalculated:
Facebook network graph attacked measures:
Global clustering coefficient : 0.6108977373560811
Spectral gap : 37.68676938602566
Algebraic connectivity : -9.464761410502781e-14

Facebook network graph normal measures:
Global clustering coefficient : 0.5191742775433075
Spectral gap : 36.8807403746528
Algebraic connectivity : 0.028988033385845668

As expected, global clustering coefficient and spectral gap measures decreased therefore robustness is lower. Algebraic connectivity was inconclusive in this case.

Defense Strategy

We have implemented a modified version of Preferential rewiring defense strategy. In our implementation we chose the edge to remove which has the maximum value for multiplication of degrees on ends of the edges, as edge to remove. To find where to add a new edge instead of the one we removed, we divided our graph into communities and found the 2 community which has the minimum value for their degree multiplication and lastly we found the 2 minimum

degree nodes within the communities, then added an edge between them. This way we aimed to distribute edges between communities and between edges uniformly.

We could have tried to add the edge between the two communities which has the least number of edges between them but that might have been inefficient if there was already a path through some other community which both of them are strongly connected. This way we aimed to connect the less connected communities into the rest of the graph.

This algorithm also has some possible problems. If the edge removed is a bridge, then our removal might turn our graph into disconnected two components and next iterations of our edge rewiring might happen inside the components which may cause our graph to never become connected again.

Possible solution to this problem might be adding a special case to never remove bridges and give priority to add an edge between communities which are not connected or give priority to edges between communities which doesn't have an edge between regardless of their connectedness.

Email network graph is defended by the previously defined defense method and results are reported:

Email network graph defended attack measures
Global clustering coefficient : 0.2673505176343738
Spectral gap : 38.72677832282112
Algebraic connectivity : 0.2498652554875013

Email network graph normal attack measures
Global clustering coefficient : 0.29649143573117465
Spectral gap : 7.681430531491266
Algebraic connectivity : 0.6207994751926129

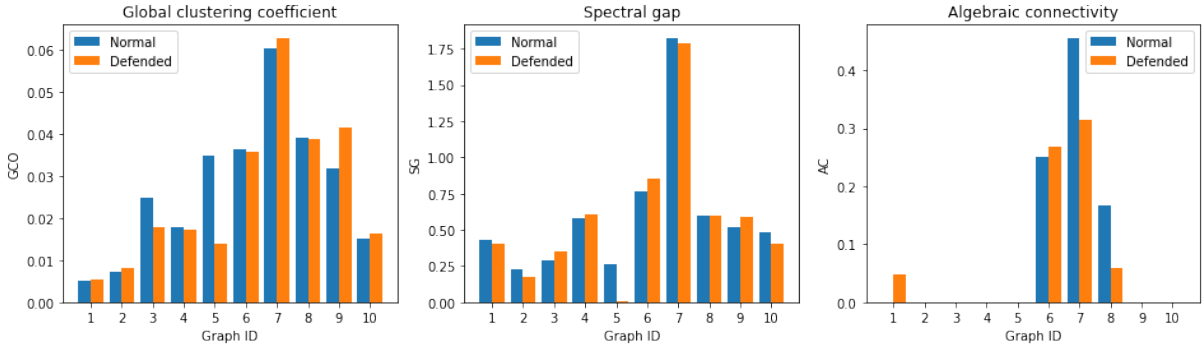
Facebook social network graph is also defended by the previously defined defense method and results are reported:

Facebook network graph defended attack measures
Global clustering coefficient : 0.5893721055056294
Spectral gap : 40.680121059053306
Algebraic connectivity : -2.832866114037121e-14

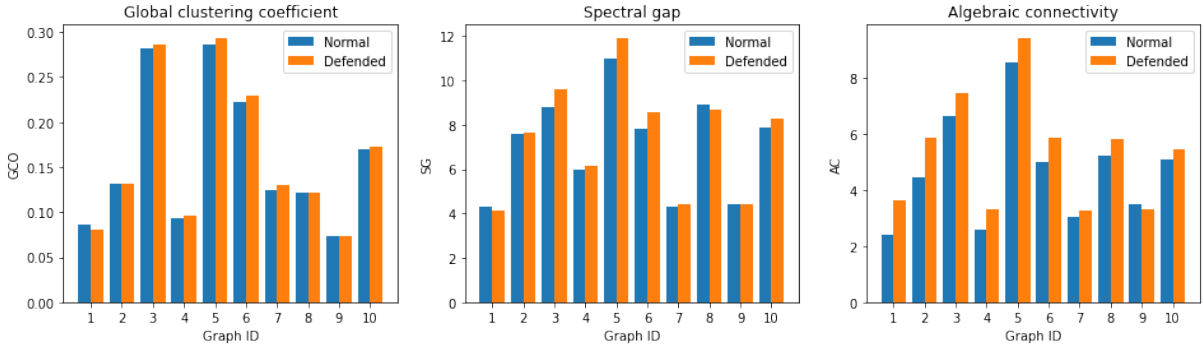
Facebook network graph normal attack measures
Global clustering coefficient : 0.6108977373560811
Spectral gap : 37.68676938602566
Algebraic connectivity : -9.464761410502781e-14

We can see a trade-off between different measures when graphs are defended.

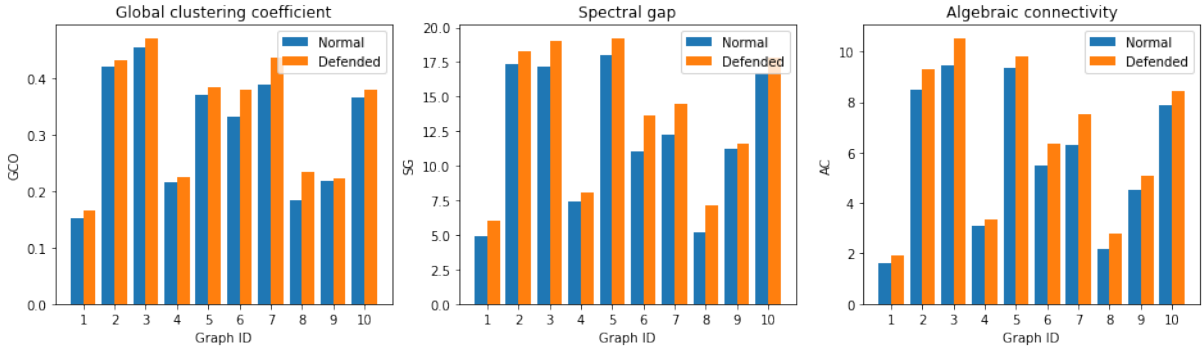
Robustness measure comparison of defended and normal attacks on Erdos model graph:



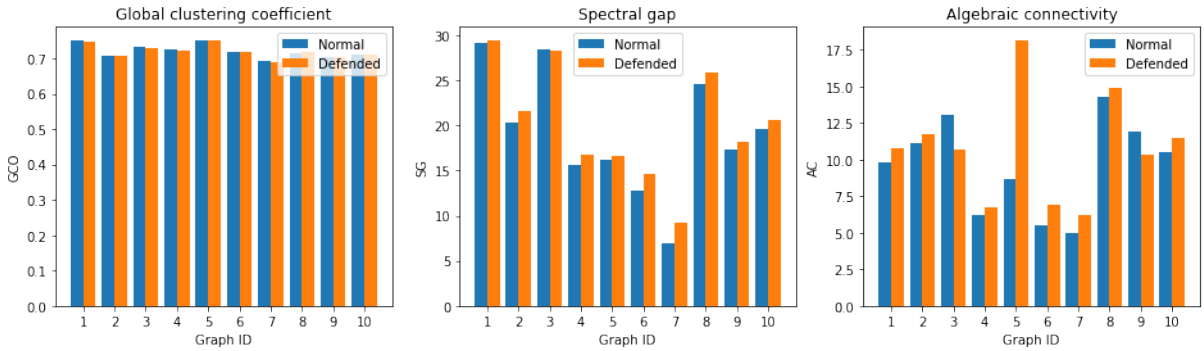
Robustness measure comparison of defended and normal attacks on Watts model graph:



Robustness measure comparison of defended and normal attacks on Barbasi model graph:



Robustness measure comparison of defended and normal attacks on Exponential model graph:



5 Conclusion

We have observed including real life price measure and robustness measures for a graph measure, there is no absolute strategy which definitely improves all of the measures but rather there can be seen a trade-off between different measures.

3 key conclusions from the project:

1. It can be concluded that robustness of graphs can be analyzed by using certain measures.

2. Attacks on graphs can impact robustness a lot even if it is performed on small group of vertices which makes graph vulnerability an important research field

3. Random graph models can be defended successfully in most cases. But defending real-world graphs is more complex.