An Experimental Look at the Stability of Graph Convolutional Networks against Topological Perturbations

Author: Yigit Colakoglu (Y.Colakoglu@student.tudelft.nl)
Responsible Professor: Elvin Isufi
Supervisors: Mohammad Sabbaqi; Maoshen Yang

1. Research Question

- How do different properties of a graph impact the stability of its graph convolutional network(GCN) against topological perturbations?
- Which combinations of graph properties can act as reliable indicators of instability in GCNs?

2. Background

- **GCNs:** Models built on a special form of convolution that uses a **shift operator** [1]

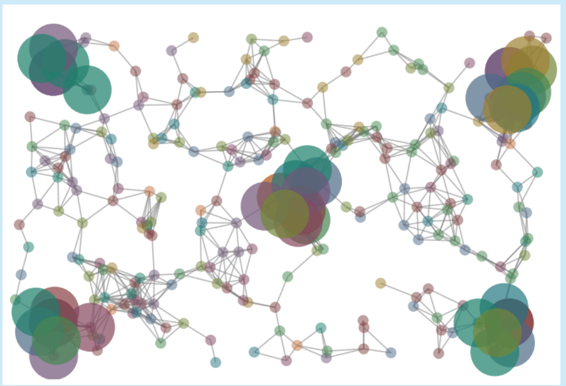

Figure 1: A visualization of the graph convolution operation, when the shift operator is applied twice []

- **Stability:** A GCN's ability to output accurate information when the graph's topology is different, i.e. perturbed
- Theoretical bounds have already been set on stability, but those bounds are often loose [2]
- Relationship between properties of graphs and real stability is still unknown

3. Methodology

- The task under investigation is semi-supervised node classification on **undirected** and **unweighted** graphs
- Topology Adaptive Graph Convolutional Networks (TAGConv) [3] are used as the GCN implementation

- **Measuring Stability:**
  - A TAGConv is trained on a graph **G**, the final layer's output **x** is saved
  - **G** is perturbed and the output of the trained TAGConv for perturbed **G** is stored (**y**)
  - The error is calculated using the relative euclidian distance between **x** and **y:**

$$\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|}$$

- **Challenge:** Existing graph datasets have a limited range in graph properties [4]
- **Solution:** Synthetically generate the datasets using:
  - Stochastic Block Model [5]
  - Lancichinetti–Fortunato–Radicchi Model [6]

- **Size and Types of Perturbations:**
  - The size of the perturbation is 10% of the number of edges in the graph
  - The perturbations should not result in self loops or create parallel edges
  - **Addition:** Add an edge between two random nodes
  - **Deletion:** Delete a random edge
  - **Rewiring:** Add and delete edges in the graph, while keeping the degree sequence untouched

4. Graph Properties Under Investigation

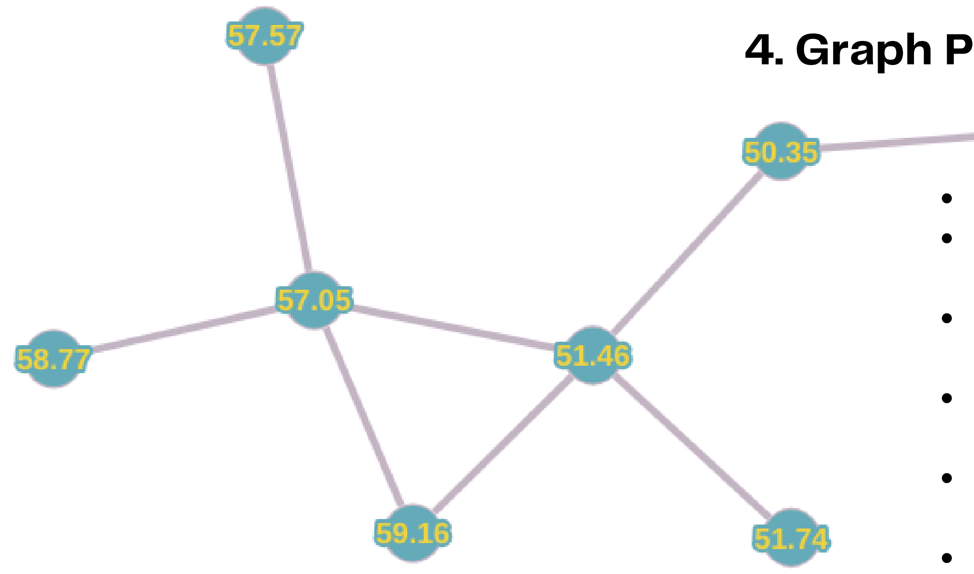

Figure 2: A graph with 8 nodes, generated using the Stochastic Block Model with a diameter of 6, connectivity of 1, clustering of 0.21, assortativity of -0.48 and centrality of 0.5

- **Diameter:** Maximum shortest path between any pair of vertices in a graph.
- **Edge Connectivity:** The minimum number of edges that need to be removed to make the graph disconnected
- **Clustering:** Measure of how much nodes tend to cluster together. Calculated by dividing the number of closed triplets by the total number of triplets (both open and closed)
- **Density:** Density represents how dense the edges are in the graph. It is the ratio of the number of edges to the number of all possible edges
- **Assortativity:** Specifically, nominal assortativity, a measure of how likely similar nodes are to be connected to each other.
- **Closeness Centrality:** The average distance of a node to every other node in the graph. In the context of this research, it is averaged over all nodes.

5. Results

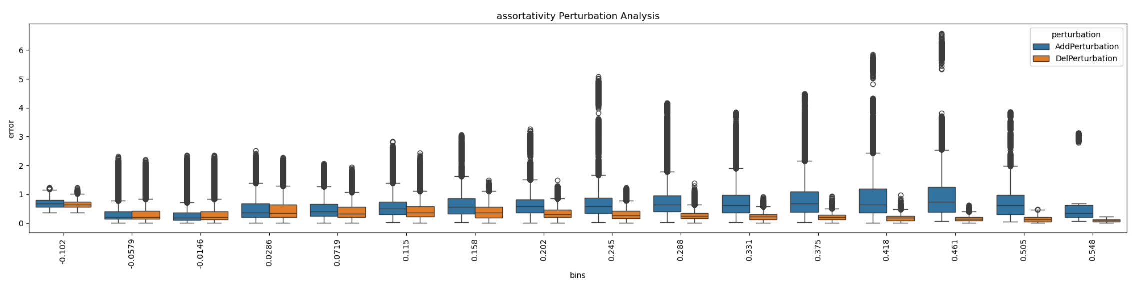

Figure 3: Box plot displaying errors obtained from graphs with different assortativity

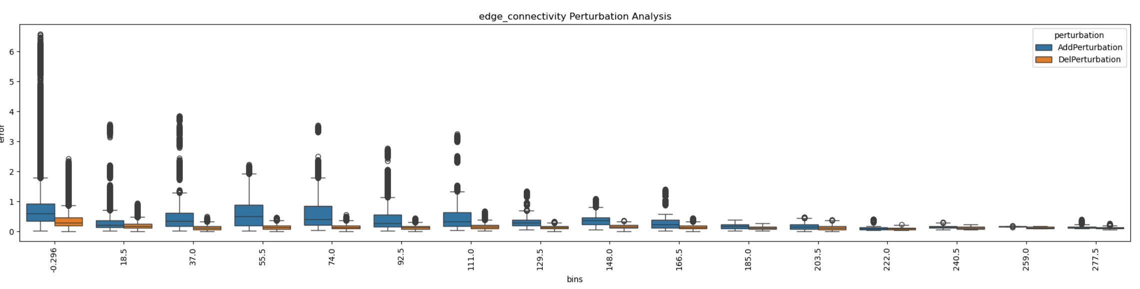

Figure 4: Box plot displaying errors obtained from graphs with different edge connectivity

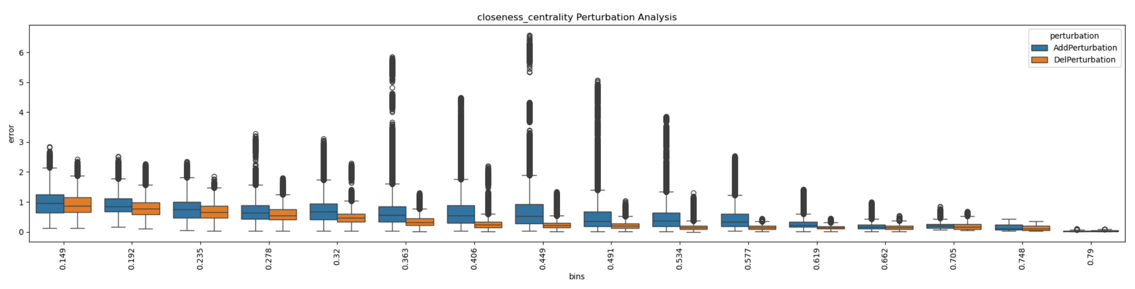

Figure 5: Box plot displaying errors obtained from graphs with different centrality

- **Key Takeaways:**
  - The range of the errors of a GCN obtained from different perturbations is correlated with its vulnerability to **adversarial** perturbations, which are small perturbations introduced explicitly in order to cause a large deviation in output
  - GCNs appear to be a lot less stable against addition perturbations.
    - This is possibly caused by the instability of integral lipschits filters against relative perturbations
  - Stability against addition is inversely correlated with assortativity while deletion is directly correlated.
  - As a general principle, the more connected a graph is the more stable it is, and less vulnerable to adversarial attacks: stability increases with centrality and connectivity

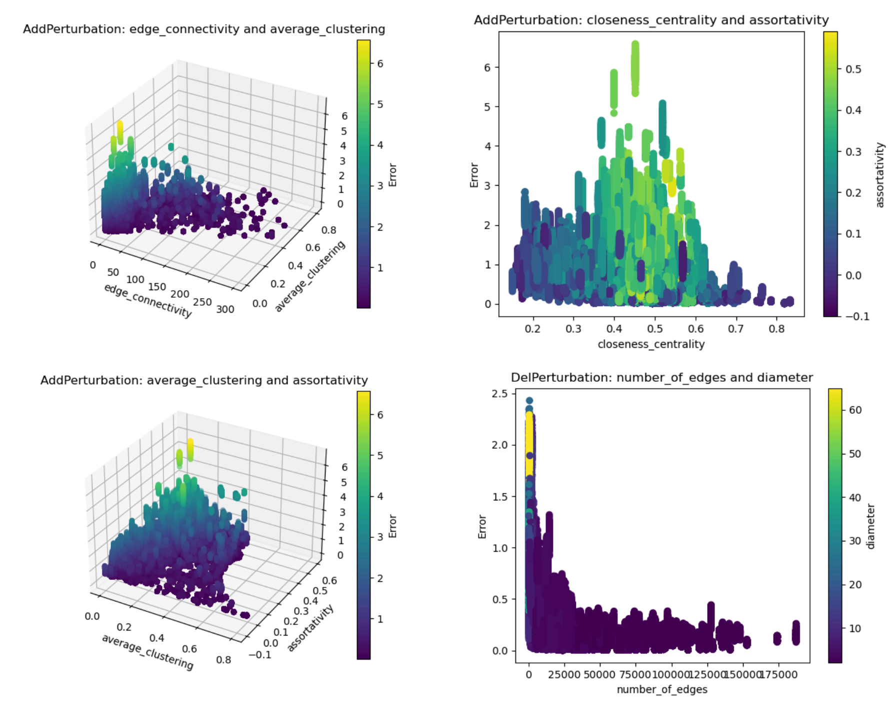

Figure 6: Scatter plotsshowcasing the relationship between 2 different graph properties and the error

- The error spikes at the middle values of closeness centrality in Figure 5 is explained in Figure 6(b). It appears that they are caused by high values in assortativity, and stability is correlated directly to centrality as long as assortativity remains the same
- Very low values in diameter can also cause spikes in error for deletion perturbations when number of edges is low, likely because in such graphs, specificedges play a very crucial role in the graph and their removal causes drastic changes in the graph

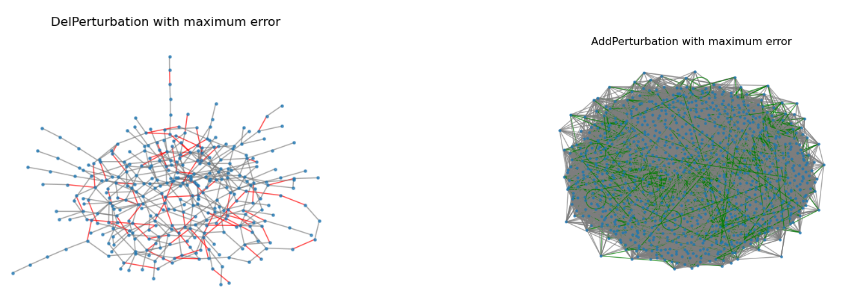

Figure 7: Perturbations that caused the maximum amount of error. Added edges highlighted in green, removed in red

- **Deletion:** It is possible to create many isolated subgraphs by removing few edges. Low connectivity, which results in low stability. The perturbation also leverages this fact.
- **Addition:** The graph is already highly connected. But the perturbation focuses on connecting certain regions, instead of being scattered around the graph, causing a deviation.

6. Conclusion

7. Future Work

8. References

[1] F. Gama, E. Isufi, G. Leus, and A. Ribeiro, 'Graphs, Convolutions, and Neural Networks', CoRR, vol. abs/2003.03777, 2020.
[2] H. Kenlay, D. Thanou, and X. Dong, 'Interpretable Stability Bounds for Spectral Graph Filters', CoRR, vol. abs/2102.09587, 2021.
[3] J. Du, S. Zhang, G. Wu, J. M. F. Moura, and S. Kar, 'Topology adaptive graph convolutional networks', CoRR, vol. abs/1710.10370, 2017.
[4] J. Palowitch, A. Tsitsulin, B. Mayer, and B. Perozzi, 'GraphWorld: Fake Graphs Bring Real Insights for GNNs', in Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022.
[5] B. Karrer and M. E. J. Newman, 'Stochastic blockmodels and community structure in networks', Phys. Rev. E, vol. 83, p. 016107, Jan. 2011.
[6] A. Lancichinetti, S. Fortunato, and F. Radicchi, 'Benchmark graphs for testing community detection algorithms', Physical Review E, vol. 78, no. 4, Oct. 2008.