

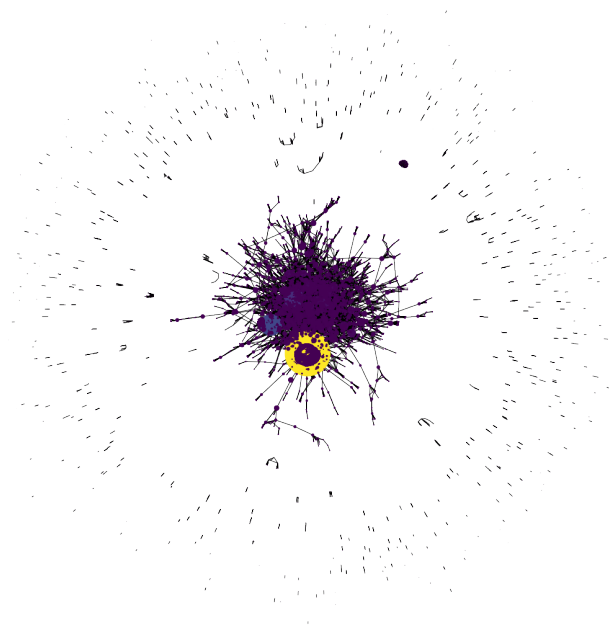
## Homework 1<sup>1</sup>

### Initial Analysis

The dataset<sup>2</sup> contains a set of nodes, usernames on Reddit, with edges linking to other users if they have commented on the same post. Figure 1 illustrates the network graph, with nodes with higher edge degrees and different *Betweenness Centrality* (measure of centrality in a graph based on shortest paths<sup>3</sup>) being colored differently. A list of key features about the network can be found below:

- ☐ Number of nodes: 25304
- ☐ Number of edges: 28611
- ☐ Average degree: 2.2614
- ☐ Clustering coefficient: 0.0864

Upon initial analysis, we can see that graph is most unconnected, meaning that many of the nodes have only one neighbor. There is a large cluster which could indicate that a lot of users tend to comment on similar top posts, thus causing a sort of echo chamber.



**Figure 1**

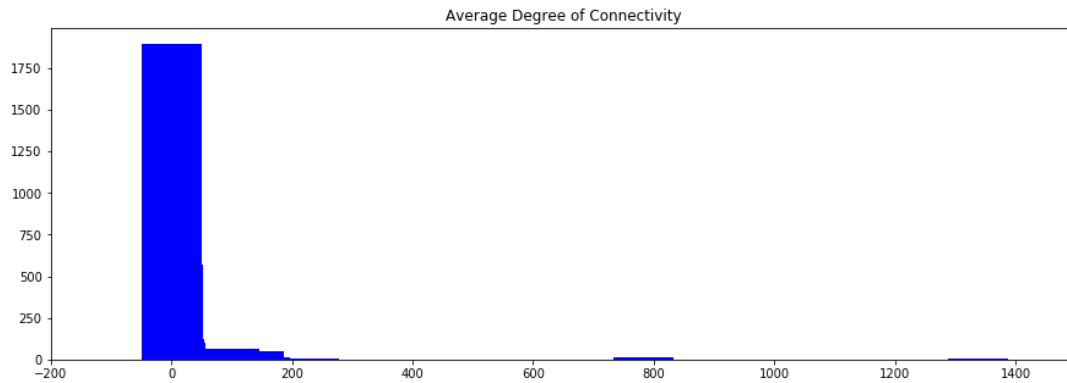
Representation of the graph shown using nodes and edges colored by density. Notice that there is a cluster of nodes with many peripheral nodes that have no edges leading outwards.

<sup>1</sup> All of the code for this assignment can be found at: <https://github.com/karkipra/CS-4440/blob/master/Homework%201/Homework1.ipynb>

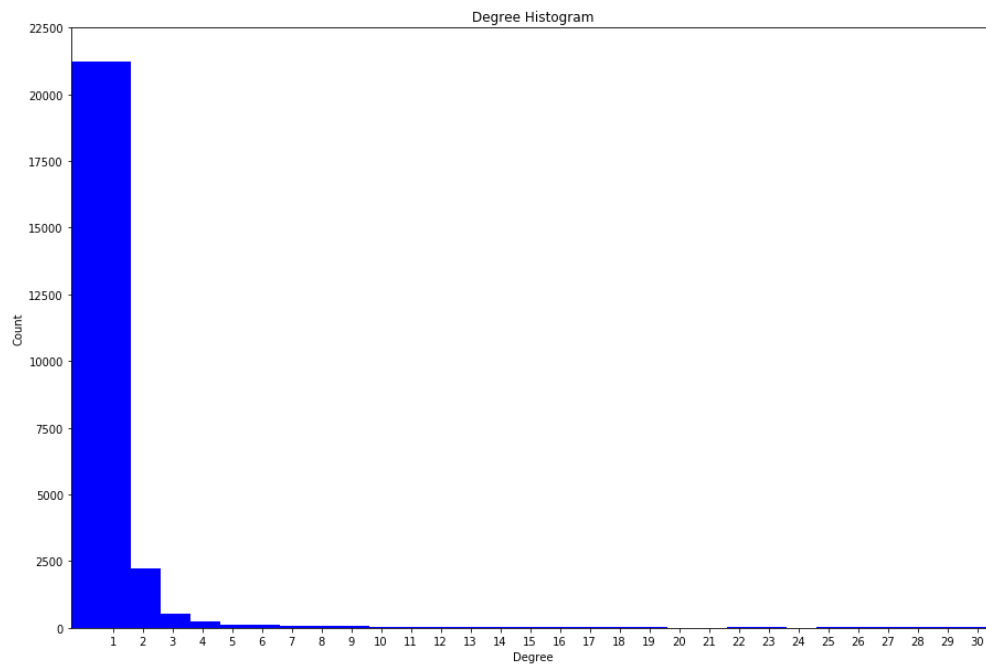
<sup>2</sup> Dataset provided for the course: networkDataset\_v2.psv

<sup>3</sup> [https://en.wikipedia.org/wiki/Betweenness\\_centrality](https://en.wikipedia.org/wiki/Betweenness_centrality)

Figure 2 shows the Average Degree of Connectivity<sup>4</sup>, which is a measure of average nearest neighbor degree of nodes with degree  $k$ . This analysis shows that the network does not have high connectivity amongst the neighboring nodes and that the small world phenomenon may not be present. Similarly, Figure 3 shows the Degree Distribution, which shows that power law is present, such that each node has an exponential effect on other subsequent ones.



**Figure 2**  
Average Degree of Connectivity of the graph. Observe that degree of connectivity is clustered around 0.



**Figure 3**  
Degree Distribution of the network. Observe that the distribution follows the pattern associated with the Power Law.

<sup>4</sup> [https://networkx.github.io/documentation/networkx-1.9.1/reference/generated/networkx.algorithms assortativity.average\\_degree\\_connectivity.html](https://networkx.github.io/documentation/networkx-1.9.1/reference/generated/networkx.algorithms assortativity.average_degree_connectivity.html)

### Power Law

I calculated the power law coefficient of the graph to be 2.657, which also matches the degree distribution found in Figure 3. As you can see, the degree is clustered around 0-1, before following a pattern like the power law distribution. The long right-skewness indicates that there is an exponential relationship between quantities within the network.

### Small World Phenomenon

From the graph and attempting to compute the small world Omega and Sigma characteristics<sup>5</sup>, I observed that the structure of the dataset does not indicate that there is a small world phenomenon. I found the methodology in the PLOS Journal to calculate a small world coefficient<sup>6</sup>. However, given the low average connectivity and that the graph isn't fully connected, the methodology could not be used in this network. This further proves that the network does not display such properties.

---

<sup>5</sup> <https://networkx.github.io/documentation/stable/reference/algorithms/smallworld.html>

<sup>6</sup> <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002051>