# Information Retrieval

# Assignment 5

Prabhat Kumar

MT17036

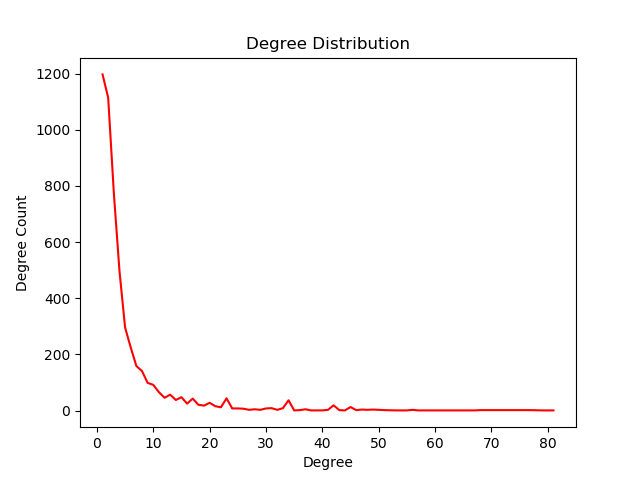
## Assumptions

* In case, for calculation of Clustering Coefficient, the degree of the neighbor node comes out to be 1 or 0, the coefficient has been assumed to be 0.

## Methodology

* Node data was mapped to 0 base node indexing.
* Adjacency Matrix was created for given Dataset.
* The Dataset contains 5242 nodes and 14496 edges.
* For closeness, Shortest Path was calculated using BFS.

## Observations and Results



* From the given graph, it can be inferred that the given data results are mostly a sparse Graph, i.e. most the nodes are not connected to a large number of nodes.
* A large number of nodes are either not connected to the Graph or re connected to Small number of nodes in degree [1,5].
* The Average Cluster Coefficient for the graph is approximately 0.53.
* Average Closeness for the Graph is approximately 1942.8. The number is coming out to be high as there are large number of nodes that have low degree.
* Average Betweenness for the Graph is approximately 0.006.

## NOTE

* Closeness of each node can be found at “closeness.json”.
* Betweenness of each node can be found in “betweenness.json”.
* Cluster Coefficient for each node can be found at “clustering\_c.json”.