COL380 A2 Report

Koduru Suchith (2021CS10572)

1 Optimizations and Performance

1.1 Serial Code (O0):

I collected colors from all nodes in rank-0 and implemented a function to find the degree centrality. It is a serialized code that uses only the rank-0 processor.
Its performance does not depend on the number of processors used, since it uses only a single processor.

1.2 Parallelized Code (O1):

Optimization:

1. At start each core gets a part of graph and a part of colors of nodes. To calculate degree centrality in each core independently we need colors of all nodes present in that core.
2. I used MPI_Allgather, so that each core shares its node colors with all other nodes, then all the cores will have colors of all nodes.
3. After each core calculates the degree centrality of the nodes present in them, all cores share their result with rank-0.
4. At rank 0, we receive the cumulative result of degree centrality of all nodes, then we sort the nodes to return the top k nodes for each color.

Performance: From the times measured in data.csv, we can observe that the time taken to compute the degree centrality decreases with increasing nodes.