

Final Project

For this project, I chose to figure out degrees of separation and Degree distributions, addressing problems such as what is the usual distance between pairs of vertices in your graph. The dataset that I used is social circles since I am very interested in how all people on the internet connect with each other even they don't know each other in real life.

The code that I wrote creates an undirected graph based on the contents of the files, calculates the distance between vertices in the graph using Dijkstra's algorithm, and then calculates the mode (most common distance) between all pairs of vertices.

There are two main functions in this code. The `calculate_usual_distance` function calculates the shortest path between all pairs of vertices in the graph using Dijkstra's algorithm and stores the distances in a hash map. It then finds the mode of the distances and prints it out as the "usual distance" between vertices in the graph. The `degree_of_distribution` function calculates the degree distribution of the graph, which is the distribution of the number of neighbors that each vertex has. It does this by performing a breadth-first search starting from each vertex and counting the number of nodes at each level of the search (up to a depth of 4). It then finds the mode of the counts at each level and prints them out.

The graph produces a total number of nodes of 773030 with usual distance between pairs of vertices of 2.

```
Finished dev [unoptimized + debuginfo] target(s) in 0.01s
Running `target\debug\graph_dataset.exe`
Graph Read Successfully. Total Nodes: 773030
Calculating distance between vertices...
Usual distance between pairs of vertices is: 2

Process finished with exit code 0
```