CS 225 - Data Structures
Project Goals
11/18/2020

Team Members
Niharika Chintalapati (nchint4)
Jonathan Kang (jk36)
Patrick McNutt (pmcnutt2)
Manny Martinez (mmart330)

Project Goals

Chosen Dataset:

1) **Road Networks -** We have decided to go with one of the provided datasets in the Stanford Large Dataset Collection. More specifically the Pennsylvania road network one. This is a road network of Pennsylvania. Intersections and endpoints are represented by nodes, and the roads connecting these intersections or endpoints are represented by undirected edges. Using this dataset we will implement some graph algorithms specified in the next section.

2) Social Network (backup): In the case we encounter too many difficulties while working with road networks we have chosen our back dataset to be this. This dataset has smaller amounts of nodes and edges and therefore might be easier to work with than the road networks dataset.

Algorithms:

1) DFS Traversal: We have chosen to go with a DFS traversal because it goes hand in hand with the Shortest Path. This traversal might aid in shortest path implementation.

2) Shortest Path (Dijkstra): We have chosen this algorithm because it is very relevant within transportation and we thought it would be interesting to use this to find shortest paths from one node to another.

3) Landmark Path: As above we have chosen this complex algorithm to implement as it is also very relevant and it would be interesting to to compare the two algorithms.

GOAL: Our team will create a small program that will use a given road network graph dataset to implement three algorithms. Our program will implement a DFS traversal and will also implement Dijkstra and Landmark path algorithms. We will also need to create our own test cases to verify our results.

CS 225 - Data Structures
Project Goals
11/18/2020

Mid-Project Modifications:

As we learned more about our data set, we learned that it would be appropriate to implement a BFS traversal, instead of our original plan to implement a DFS traversal. We had many reasons to make this modification, including:

1) Our data does not present in levels with a root node, leaf nodes, or branches, for us to best implement a DFS traversal
2) One of the algorithms we used, Djikstra's, was implemented in BFS fashion using edge weights and a priority queue in order to find the shortest path
3) A BFS traversal is typically faster than a DFS traversal, which is important as we initially wanted to test on a significantly larger number of nodes