Austin Thiel

Dr. Linstead

CPSC-350

11 December 2014

Graph Research

Dataset Used: “Social Circles: Facebook”

Source: snap.stanford.edu

Nodes: 4039

Edges: 88234

The data used was of ‘circles’ or friends lists from Facebook collected from an anonymous survey. Information such as political party and religion were given random key values so as to obtain affiliations between users, but keep the information anonymous. This dataset is of all egonets combined.

I attempted to answer a few questions involving depth vs. breadth first search as well as analyze the results of using Prim’s algorithm. I’m not sure if I’m lacking some understanding of the algorithms, or if my data is clashing and my code is incorrect, but there were some inconsistencies somewhere in my project. The results are within a range of reason, however. Implementing the boost library proved to be more of a challenge than it was worth for this assignment, so I used a templated Graph class in order to analyze the dataset. I greatly regret not working with others on this assignment, as I found myself stuck in a hole a few times that I feel would have been much easier to get out of with a little help from my friends.

Running a depth first search from any position in the graph returned anywhere from 4006 to 4010 visited nodes, which means some people share no friend relations with others and are therefore in a sub-graph on their own. A breadth first search returned nearly the same results (3999 to 4010), with a slight amount of less visited nodes in some areas when run from any given position.

Prim's minimum spanning tree algorithm produced interesting results. Whenever it was run from any given starting vertex, it always returned 4039 connections, which implies that all nodes are, in some way, connected to one another and that there are no sub-graphs in this dataset.

I was also able to find out how many friends were connected to a specific person, like how person 10 only has 9 friends, or how person 0 has 347 friends. This is useful for finding out general relation statistics like the average number of friends per person, or who shares the most amount of friends with another person. Although, I could not find an algorithm to perform the latter.

All in all, the results are a little plausible but by no means worthless. Finding relational information between anything is incredibly interesting and this assignment made me realize the bounty of things it can be used for, and how relatively simple it is to implement on a small scale.