Project 1 Proposal

DATA 620: Web Analytics

Igor Balagula

Data

For this project I created a dataset that lists all of the college friends of myself and my wife and our friend's children.

Nodes

Nodes are college friends of myself and my wife and our friend's children.

Edges

Edges indicate close relationship between people. Any pair of people who call themselves friends are connected.

Categorical Variable

Gender: male or female

Method

I asked our friends to fill out a quick online survey where I listed all of our college friends and their children and asked everyone to check if they have a close relationship with all of the other people and consider them close friends. Data were loaded into a file in Pajek format and analyzed using NetworkX package.

Questions

Who are the people with the highest degree centrality?

Is there a statistically significant difference in degree centrality between males and females? I.e. do females have more friends?

How much different is centrality measure of our friends and our friend's children?