6°

Group Members

- Anastasia Vela
- Joshua Asuncion
- Kaiwen(Kevin) Pang
- Steve Hwang

Date:
2017/10/16

1. Project Description and Background

Six degrees of separation is a concept that was created in 1929. It stated that a person could be connected to another person in the world within 6 connections. However, with the advent of social media, there may be a change in how connected we are as a society, and result in lower degrees of separations between 2 people. We want to test if facebook has achieved their mission statement of bringing the world closer together by seeing if people are generally less separated than they were before facebook was founded. In addition, we want to create a model that could visualize how strong the connections are between two people, not just the number of people it takes to connect. We will do this by weighing each connection based on the number of mutual friends.

2. Motivation

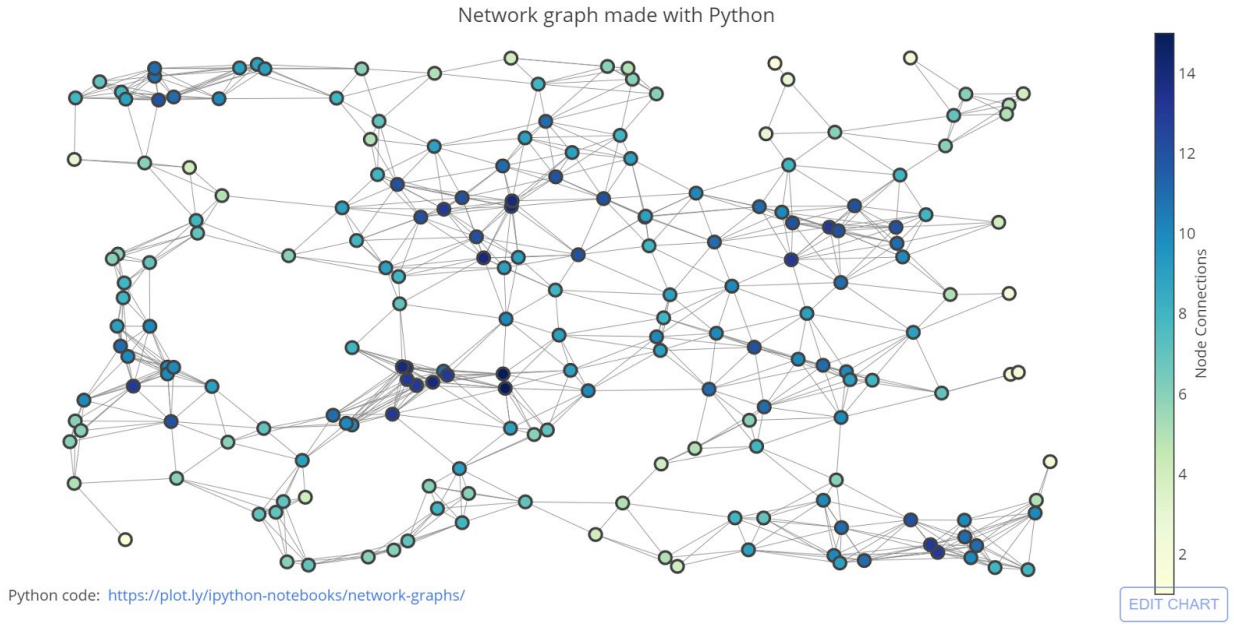
We decided to focus specifically on the UC Berkeley student population. We will single out those on Facebook that mark themselves as attending UC Berkeley, and from that dataset we will analyze numbers of mutual friends. Our motivation for this is that we believe focusing on this population presents unique questions. Does a smaller population mean that people are more separated, because there are less paths to get to everyone? Or will we find that people are less separated, because there are less people to get to overall?

In addition, we were motivated to look at the number of mutual friends two people have because we reasoned the number of mutual friends generally shows how close of a relationship they have. If two people are close, then most likely they also share many of the same friends (and this is true especially for a population of college students such as the one we are observing). With this reasoning, we will check the strength of the connection between any two given Berkeley students through their number of mutual friends.

There are, of course, assumptions that we are making, such as the assumption that the Facebook UC Berkeley population is representative of the actual student population. We understand that not everyone uses Facebook, and we also understand that there might be cases where there are few mutual friends in a strong relationship. However, we believe that, because the use of Facebook is so widespread, and also because of the very fact that most people use Facebook to get connected with friends, those cases are few and we can make such generalizations.

3. Plan of Action

Using Facebook's Developer API, we will record a list of a person's friends who claims to go to UC Berkeley. Each person in that list will also have a list of friends who attend UC Berkeley. The weight of each branch from one person to their friend is determined by how many mutual friends they have with each other; the more mutual friends, the easier it is to pass to the other node. The graph will virtually look as followed:



With the graph developed, we will use an algorithmic search called Dijkstra's Algorithm to find the shortest, heaviest path from any arbitrary node. After finding how many steps it takes to get to any person, we will average the paths to find the average number of steps to get from one person to any other person.

4. An initial task division

- Facebook API and Coding - Anastasia
- Tableau and R Visualization - Josh
- PPT Slidedeck and Demos - Kevin
- Final Report - Steve

The team has decided to meet every Friday to work the project. Preliminary timeline is shown as below:

Task	Completion Date
Project proposal	**October 17th**
Test Facebook API to extract mutual friend data and search for libraries that help to visualize	October 20th
Continue to extract data	October 27th
Work on status update on the project	November 3rd
Status update on the project and design for the system	**November 7th**
Visualize using Python libraries, Tableau and/or R	November 10th
Compile trends and visualizations; start working on PPT	November 17th
Finalize and practice presentation and demos	November 24th
Project presentations and demos	**November 28th**
Work and finalize report	November 30th
Final report	**December 5th**