Assignment 05 - Graphs

CS 581: Online Social Networks
Spring 2021

Assignment Description:

This project has two parts. Write a one-page report describing your program. The report must include the following labeled sections:

- PURPOSE;
- INPUT
- OUTPUT
- WHAT THE PROGRAM DOES
- ADDITIONAL INFORMATION

(Anything you want to share--students often talk about special issues that came up and how they addressed them, for instance. You may not have anything you want to include, and that is fine.)

Report

1.1 Purpose:

The purpose of this assignment is to process the input file and identify triads. Your program should successfully run on file epinions1.csv within 15 minutes.

1.2 Input:

The inputs that this program takes from the user is the name of the csv file that it is processing i.e. epinions0.csv or epinions1.csv. I was not able to process epinions2.csv the way I needed to.

1.3 Output:

The outputs of this program include number of edges used to identify triads, number of positive (trust) edges, number of negative (distrust) edges, probability p that an edge will be positive: (number of positive edges) / (total edges), probability that an edge will be negative: 1 – p, number of triangles, expected distribution of triad types and an actual distribution of triad types.

1.4 What the program does:

Once the program takes in the needed and correct csv file, it will process the file through graphs using the networkx package and processes it in a way that it can derive the outputs as were listed in the previous point. It uses the “parse_edgelist” method in order to get the graph from the csv file and uses methods such as “enummerate_all_cliques” and “edges” to get all of the needed information.

1.5 Additional Information:

This assignment was pretty challenging when it came to the actual vs expected distribution part of the assignment. Took a bit of python knowledge and learning the networkx package but in the end it was able to get done.