Gopi Miyani 10437266

Assignment 06 - Report

I have implemented a program in python named GopiMiyani_Assignment06.py for the purpose of processing graph data from csv file and do some analysis on the results. User can run this program with command: "*python3* GopiMiyani_Assignment06.py -filename myfile, where in place of myfile, user has to specify file name that he/she wants to process. The output of program will be the results of graph data processing such as number of edges in the network, number of self-loops, number of triangles, number of trust edges, distrust edges, probability of trust and distrust edges, expected and actual distribution of triad types with number and percent etc.

I have used networkx library of python, which is a library for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. I used different functions available in network library to get various results on graph data. As illustration, I have used nodes_with_selfloops() function to count self-loops, triangles() function to get all triads of graph, enumerate_all_cliques() to get list of triangles of clique size of 3 etc.

I faced problems in finding out expected distribution and actual distribution of triad types as I was not able to find out any specific documentation that I can refer to. But, finally after doing much effort, I can execute and get the results. From the results, I found that the major difference between expected and actual distribution of triad types is in the percentage or number of TTT and TTD triad types because in expected distribution, we are calculating based on probability and random distribution whereas in actual distribution, it is based on actual weight edge and calculating triad types using actual distribution.