Christopher Rudel

Professor Dugas

CS 581 WS

Assignment 5

The purpose of this assignment is to take the input of graph data stored in a csv and to identify all the triads and the number of types of triads that are found. The program prompts the user to name the file they would like to analyze, and takes in the data from the csv. The program assumes that the data in the csv is formatted “from,to,weight” e.g. “4,9,-1” where 4 is the starting node, 9 is the ending node, and -1 is the weight. All weights are assumed to be 1 or -1, symbolizing a trusting relationship or a distrusting relationship respectively. I then took the probability that an edge would be trusting or distrusting from the data and predicted the amount of triads that were TTT, TTD, TDD, or DDD. TTT represents full trust, TTD represents one distrusted edge, TDD represents only one trusted edge, and DDD represents a fully distrusting triad. Finding triangles within the nodes was more time consuming than I anticipated. I originally tried to implement my own algorithm, however after some analysis I realized that the time complexity of it was $O(n^3)$, so I decided to rely on NetworkX’s algorithm. In the end I was not able to run the epinions.csv file because of the size, and epinions_small.csv still takes a bit of time to complete running (about 5 minutes). In the epinions_small.csv file, I predicted that the triad's probability would be 64%, 30%, 5%, and .2% for TTT, TTD, TDD and DDD respectively based on the total number of trusting and distrusting edges. However I found that the respective probabilities were actually 73%, 15%, 10%, and 2% based on the actual triads. I find it reasonable that the TTT probability was higher than predicted because generally I would expect

that people respect and like one another. It also makes sense to me that TDD is less in order to compensate for the rise of TTT triads. I am not too sure of why TDD and DDD are higher, maybe if one person notices untrustworthy attitudes from another person, they in turn distrust that person as well creating a TDD or DDD triad more often than expected.