Analysis and Modeling of Social and Information Networks
CIS 4524/5524, Spring 2019

Assignment 1, due January 31th in class

*Please write your name and TUID at the top of your hard copy submission and staple all pages together.

Homework Policies (applicable for all assignments):

1. You are required to do the homework problems in order to pass.
2. Understandability of the solution is as desired as correctness.
3. Penalty for late homework assignments submissions is 20% per day. So, do it on time.
4. Solutions are expected to be your own work. Group work is not allowed unless explicitly approved for a particular problem. If you obtained a hint with help (e.g., through library work, etc.) acknowledge your source, and write up the solution on your own. Plagiarism and other anti-intellectual behavior will be dealt with severely.

The Course Network
A bipartite network attached to this homework is generated according to the background survey. Each of you is representing a node of this network (labeled as a number). In this network there exists an edge between a student node and a topic node if and only if this student has taken a course on that topic. In this exercise your job is to project the attached bipartite network to three student networks. In each projection network a link between two students is obtained by connecting corresponding students if they both took a course on a certain topic. The difference in three projection is which courses are considered to construct a projection. Construct Students1 projection network by considering students background related only to the Social Science, Economics, Optimization, Marketing and Social Media topics. Students2 projection network is more inclusive and should in addition include background on Information retrieval, Vision, Networks, Database and Artificial Intelligence. Students3 projection network should include all topics considered in Projection2 and also Artificial Intelligence, Data Mining and Machine Learning. Exclude students 10,11,12 and 23 from the analysis (missing data).

Problem 1. [Projecting and Visualizing Networks] Visualize 3 constructed projections using a graph visualization software. You can use Gephi, a platform-free software for graph visualization and analysis, which you can download from https://gephi.org/ (follow the quick start guide to import the student network and visualize it; before importing the student network, you need to transform the network into a supported graph formats). The the CSV format is recommended. Either edge list or matrix under CSV format should be fine. You can also use Python library NetworkX. Check sections 'Creating a graph' and 'Drawing graphs' to see how to visualize a graph. Make a screenshot of your visualized networks, print them out.

Problem 2. [Computing Global Network Properties] For each of 3 constructed projections compute:
   a) the size and diameter of the network largest connected component,
   b) degree distribution,
   c) average path length and
   d) clustering coefficient.
   All metrics can be calculated using Python library NetworkX, but you can use any freely available packages to compute these properties or you develop your own code.

Problem 3. [Topics Network Visualization and Analysis] Construct, visualize and compute global network properties for a topics network obtained by projecting the attached bipartite network to a network of topics where two topics are linked by a weighted edge representing number of students who took a course on both topics. Exclude Programming, Statistics, Graphs, Algebra and Algorithms (pre-requisites for the CIS 4524/5524).