Rating Profiles

Network Analysis, University of Helsinki

Team Members

- Tomáš Vopat
  - 015148847
  - tomas.vopat@helsinki.fi
- Teemu Koivisto
  - 014211393
  - teemu.koivisto@helsinki.fi

Project Description

We are using dating site dataset to understand how user ratings of other users' profiles work. The dataset contains 220 970 users, their gender and 17 359 346 edges representing the ratings between different users. Our goal is to find communities in the network, see if men/women use/receive ratings differently. And lastly, how the ratings and the connectedness of nodes (popularity) in the network correlate. Dataset: http://www.occamslab.com/petricek/data/

Effort Distribution and Timeline

We will use Python with Jupyter notebook and NetworkX library for graph representation, the code in Github. In the first step, we try to visualize the networks using some graph visualization tool and do fundamental analysis such as the number of nodes/edges or degree distribution. Next, we try to find the different groups and do some preliminary analysis of how the ratings and gender affect the network. Also, to try to predict the unobserved ratings between given profiles. Lastly, we will try to find the most exciting properties of the network and figure out how to present them. All of this is done in the course of the next 3-4 weeks. We will be splitting the work so that both of us have a chance to try all the models by themselves, and see if they come up with the same results as the other person.