Project 6 – Community Detection in a Youtube Network

Aim:

The aim of this project is to analyze the social behavior in a Youtube friendship network and identify communities and influencers. The tool should be able to:

- Load a social graph
- Run community detection and network analysis methods
- Visualize the network

Tasks:

1- Dataset

- Load the Youtube dataset: http://snap.stanford.edu/data/com-Youtube.html
- Explore the properties of the graph. Consider reducing the network to a smaller connected subgraph.

2- Implementation

- Implement the Girvan-Newman clustering algorithm.
- Implement the Betweenness centrality.

3- Analysis

- Use Girvan-Newman algorithm to find clusters of users on the generated network. Evaluate with different values of the iteration level.
- Identify the top *k* users with highest Betweenness centrality. Choose an appropriate value of *k*.

4- Visualization

- Visualize the output of Girvan-Newman algorithm, by coloring nodes according to their assigned groups.
- Visualize the generated network, and highlight the top *k* authors with highest Betweenness centrality.

Contact person: Rana Hussein (rana.hussein@unifr.ch)