



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Network Science

Project 1 - VK friends graph analysis

Gremyachikh Leonid

Network Summary

Friends/ **Giant component**/ **B&A model**:

Nodes number – 225;

Edges number – 1012;

Diameter – 7

Average shortest path length – **3,38**/ **6,33**;

Clustering coef. – 0,52/ **0,58**/ **0,017**;

Average degree – 9/ **10**;

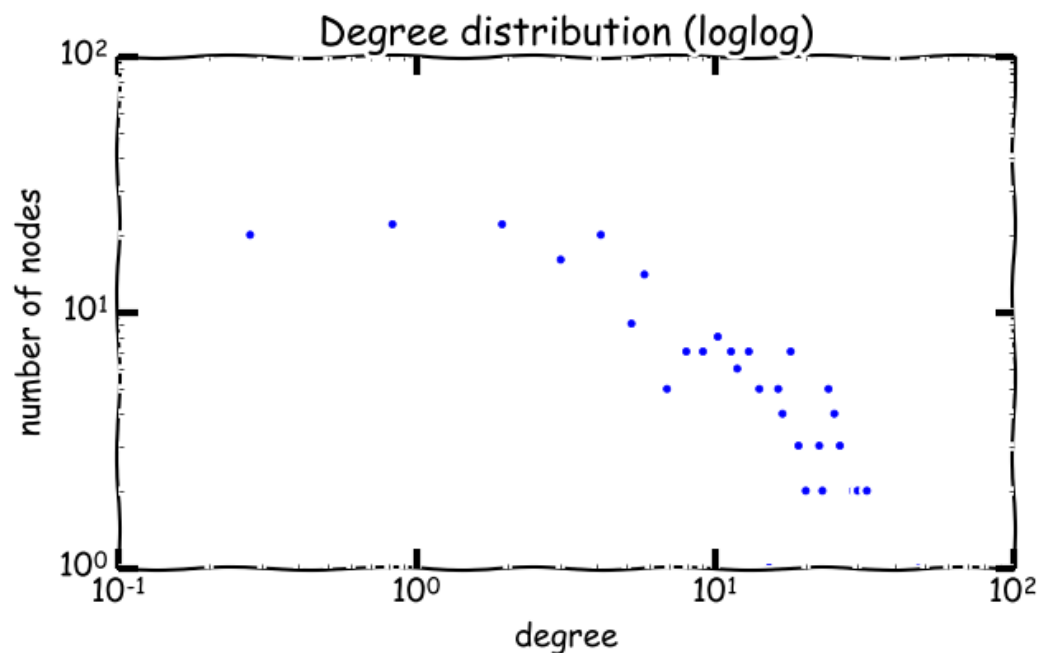
Assortativity coef. – 0,06/ **0,056**.

Remark: ASPL and CC for B&A model:

$$\ell \sim \frac{\ln N}{\ln \ln N} \quad C \sim N^{-0.75}$$

Some observations:

- Clustering coefficient ("all-my-friends-know-each-other" property) is relatively high;
 - Average shortest path length is relatively high;
- => My friends relatively know each other well
- Degree distribution look like Power Law.



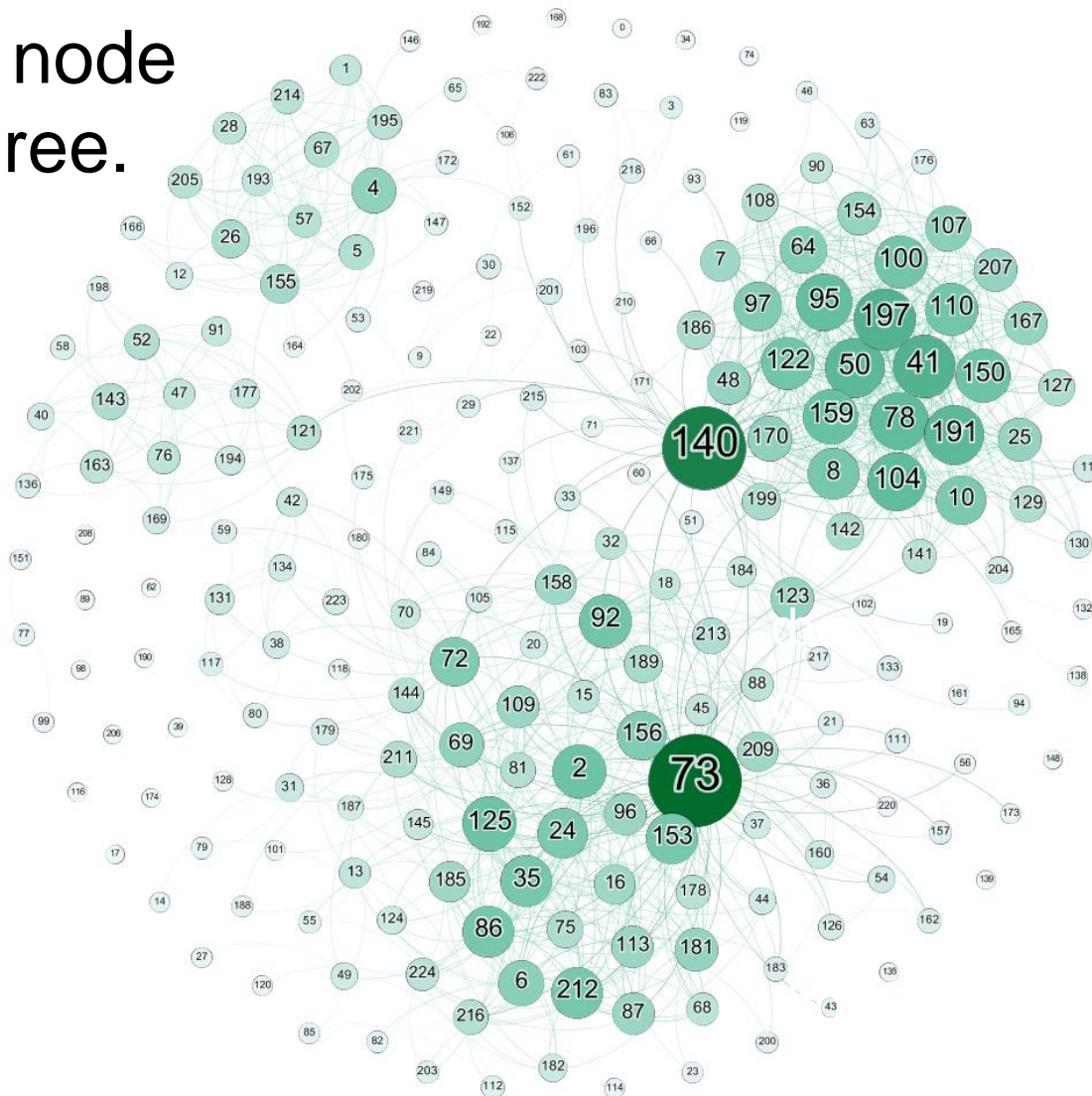
Friends graph

The Friends graph with the node size proportional to the degree.

Some observations:

We can immediately distinguish nodes 73 and 140. Actually, there are my best friend and girlfriend.

They have degree 55 and 48 respectively.



More centralities - Closeness centrality

Closeness centrality - how close an actor to all the other actors in network.

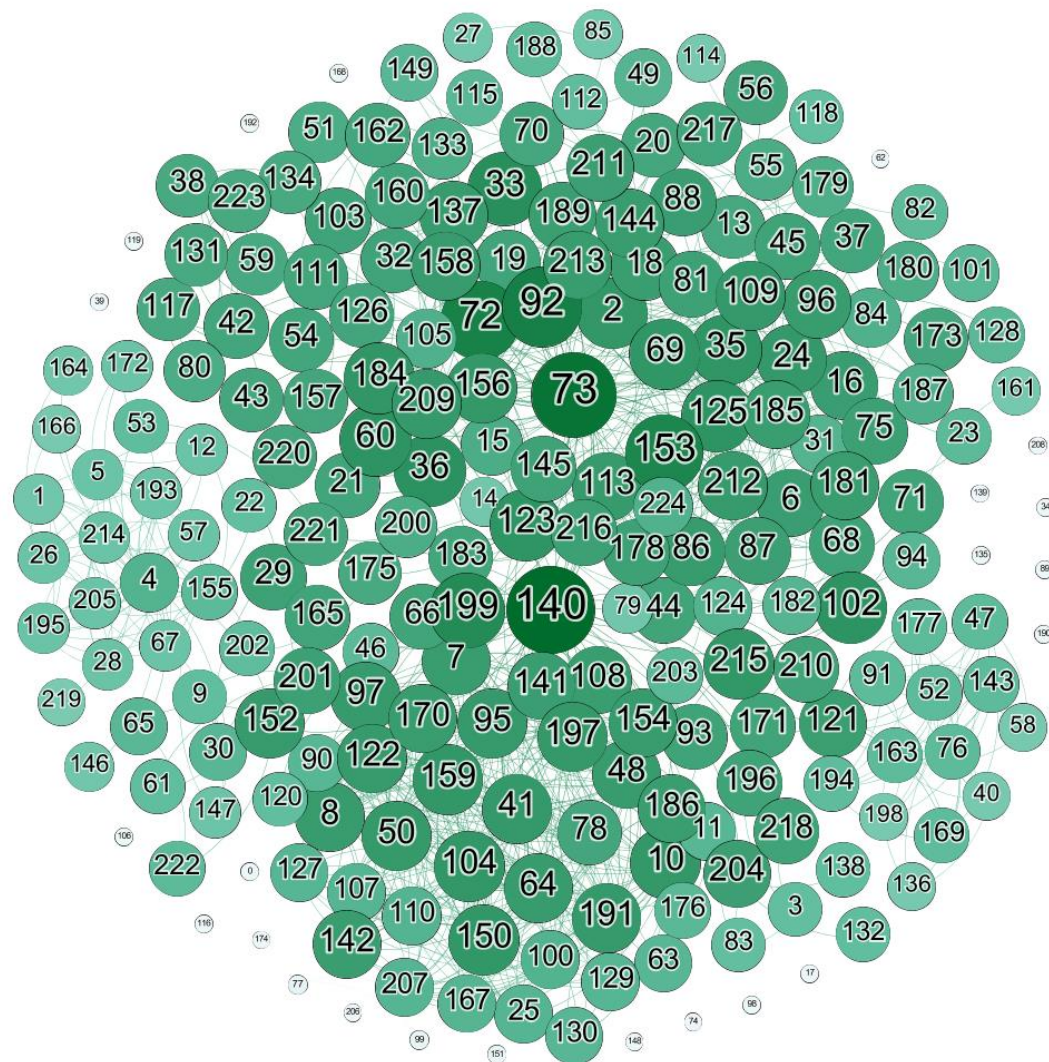
High closeness centrality - short communication path to others, minimal number of steps to reach others.

So, my friends are very close very close to each other.

Top 10:

73, 140, 41, 197, 50, 191, 78, 104, 95, 159

– there are my best friend, girlfriend, guys from theatre and two my fake accounts.



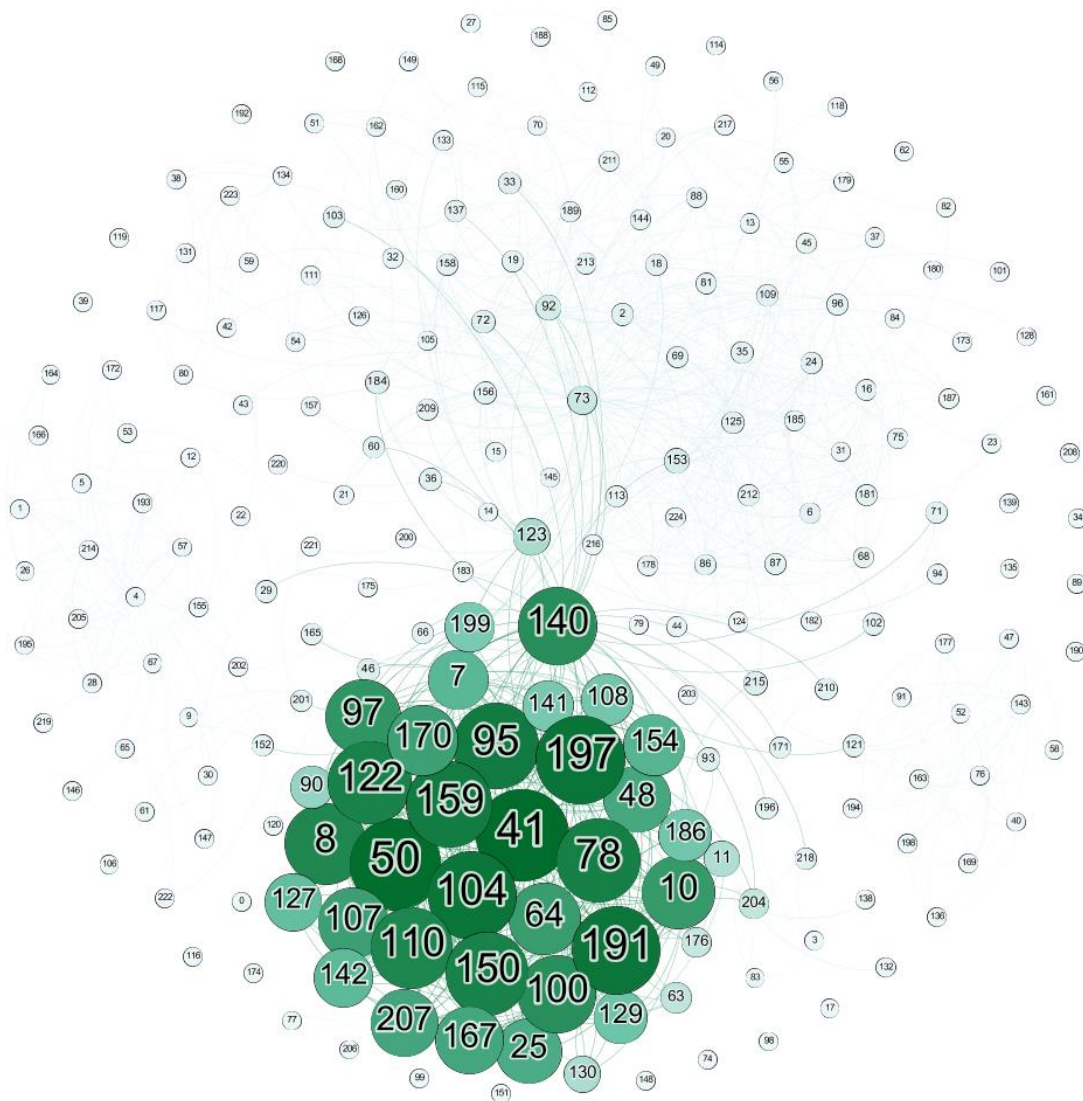
More centralities - Eigenvector

Eigenvector centrality - is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes.

Top 10:

140, 73, 92, 72, 153, 199, 33, 36, 60, 102

– there are my best friend, girlfriend and some mix guys, who know most of my closest friends.



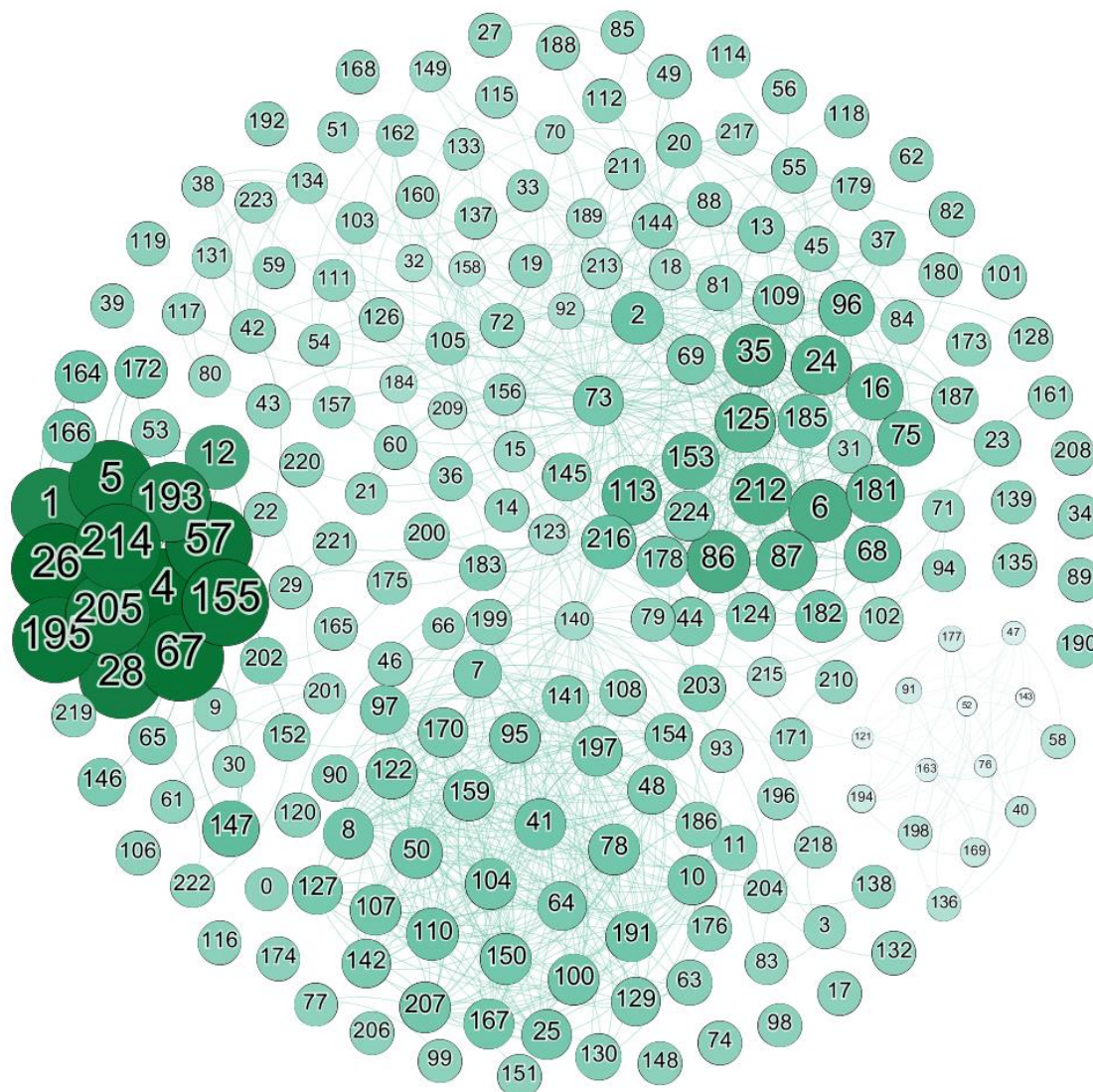
More centralities - Katz centrality

Katz centrality - measures influence by taking into account the total number of walks between a pair of nodes.

Top 10:

4, 26, 67, 155, 57, 214, 195, 5, 205, 28

– there are ALL my classmates from last university.



More centralities - Local load

Load centrality - fraction of all shortest paths that pass through that node.

Top 10:

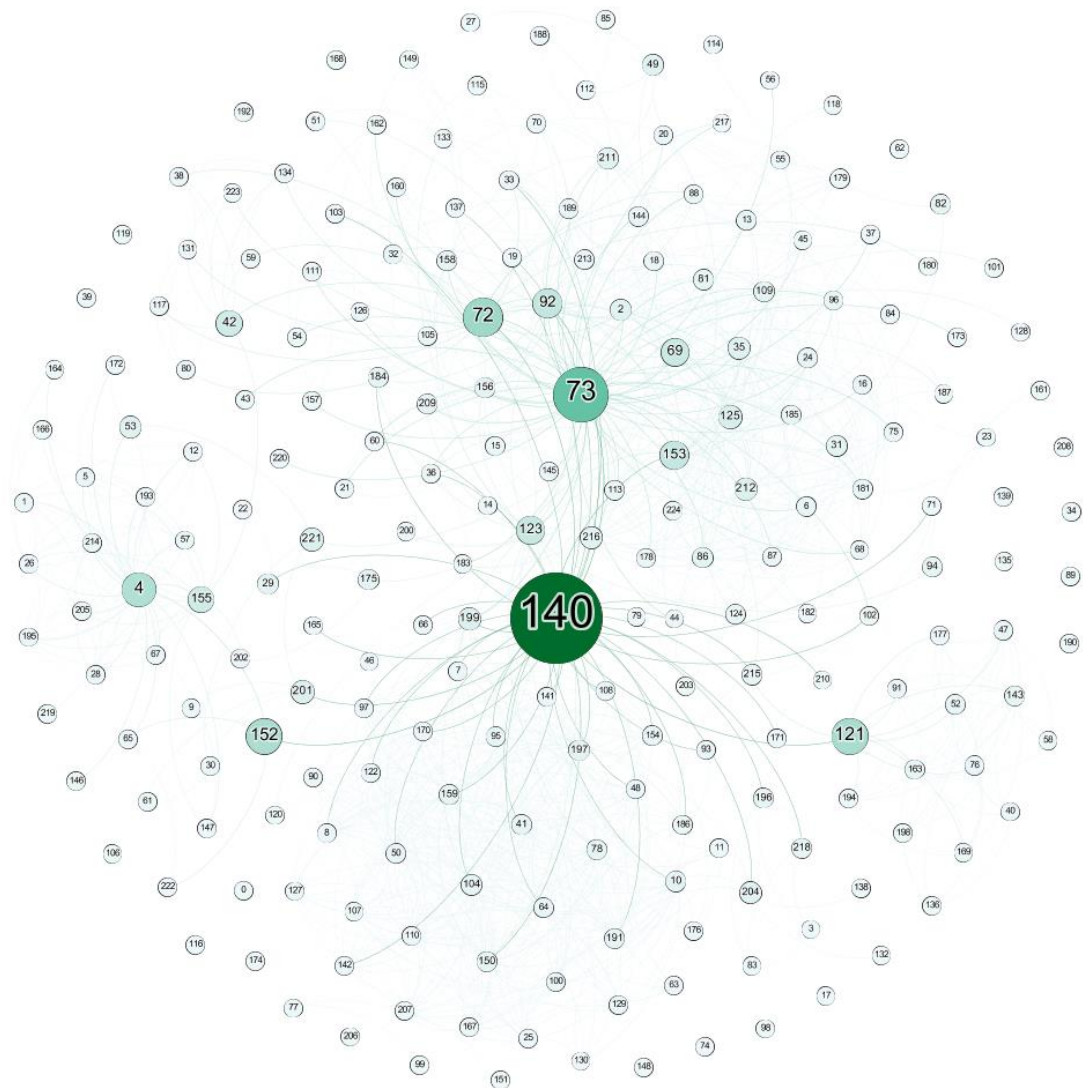
140, 73, 72, 121, 152, 4, 92, 153, 123, 69

– there are my best friend, girlfriend and some quite well-known people from different companies.

Note that the similar result we have by Betweenness centrality:

Top 10:

73, 140, 4, 72, 69, 92, 153, 212, 197, 2



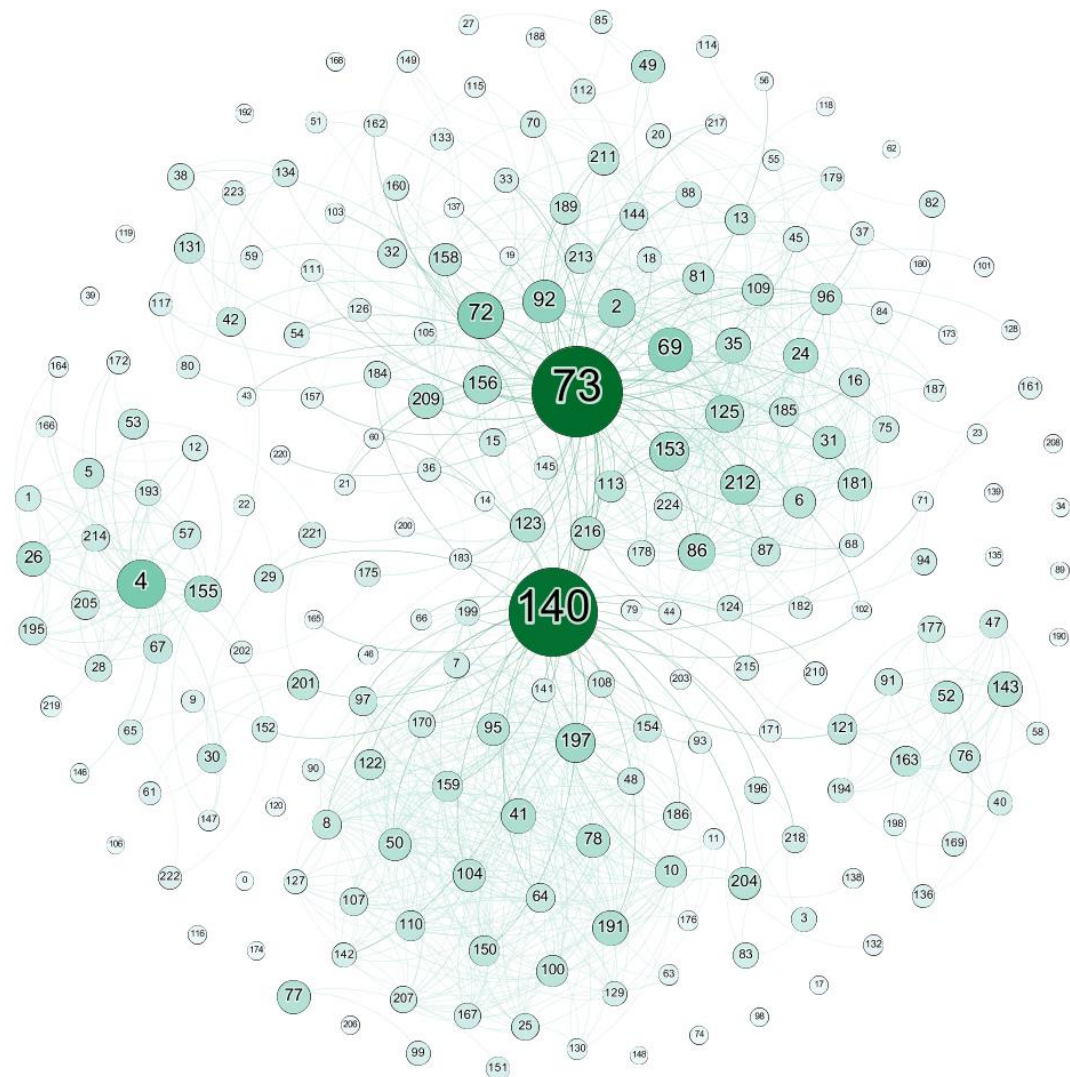
PageRank - is an algorithm used by Google Search to rank websites in their search engine results, is a way of measuring the importance of website pages, is the **probability** that the random surfer visits a page.

Top 10:

73, 140, 4, 72, 69, 92, 153, 212, 197, 2

– there are my best friend, girlfriend and some mix guys, who know a lot of people which don't know each other. They are some kind of connecting links.

One-on-one Betweenness centrality.



Random model

Simple modification of B&A Model give a very similar graph.

Friends graph/ [B&A modification model](#):

Nodes number – 225/ [225](#);

Edges number – 1012/ [1012](#);

Diameter – 7/ [7](#);

Average shortest path length – 3,38/ [2,71](#);

Clustering coef. – 0,52/ [0,05](#);

Average degree – 9/ [9](#);

Assortativity coef. – 0,06/ [0,056](#).

Model parameters:

$m_0 = 10$;

$N = 225$;

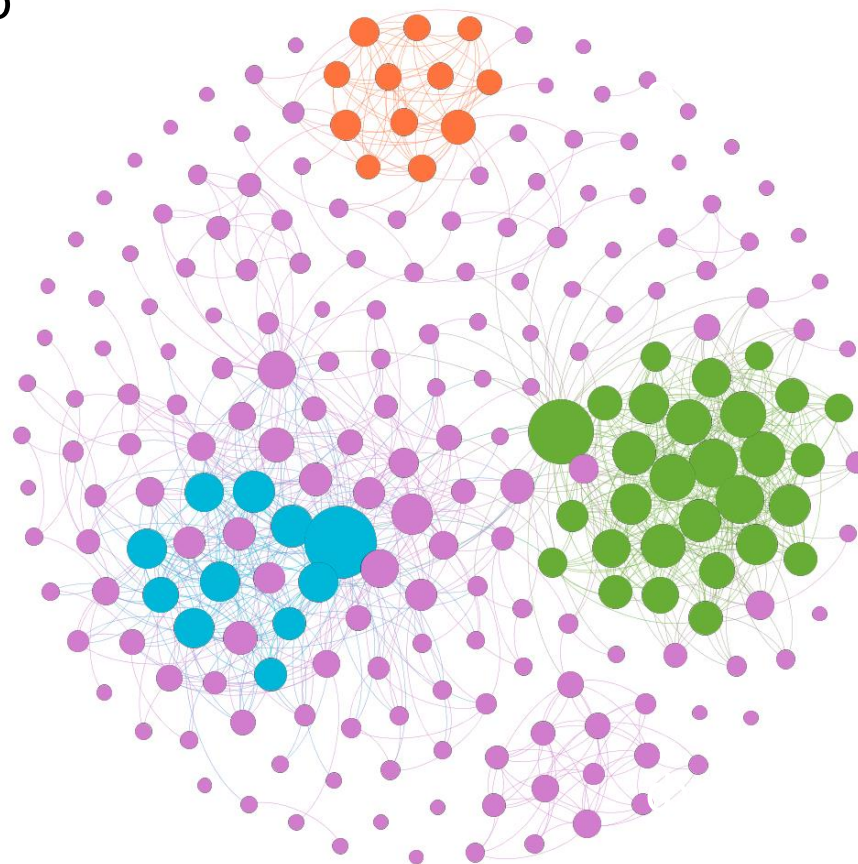
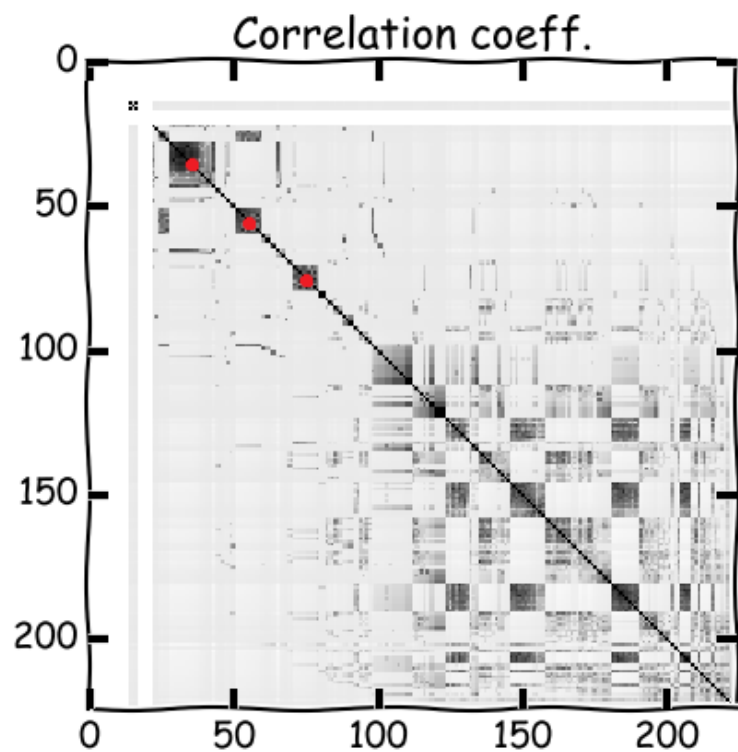
$\text{deg_set} = [4, 5]$.

(deg_set - variants of degree for new node)

Community Detection

Cuthill-McKee and correlation coefficient:

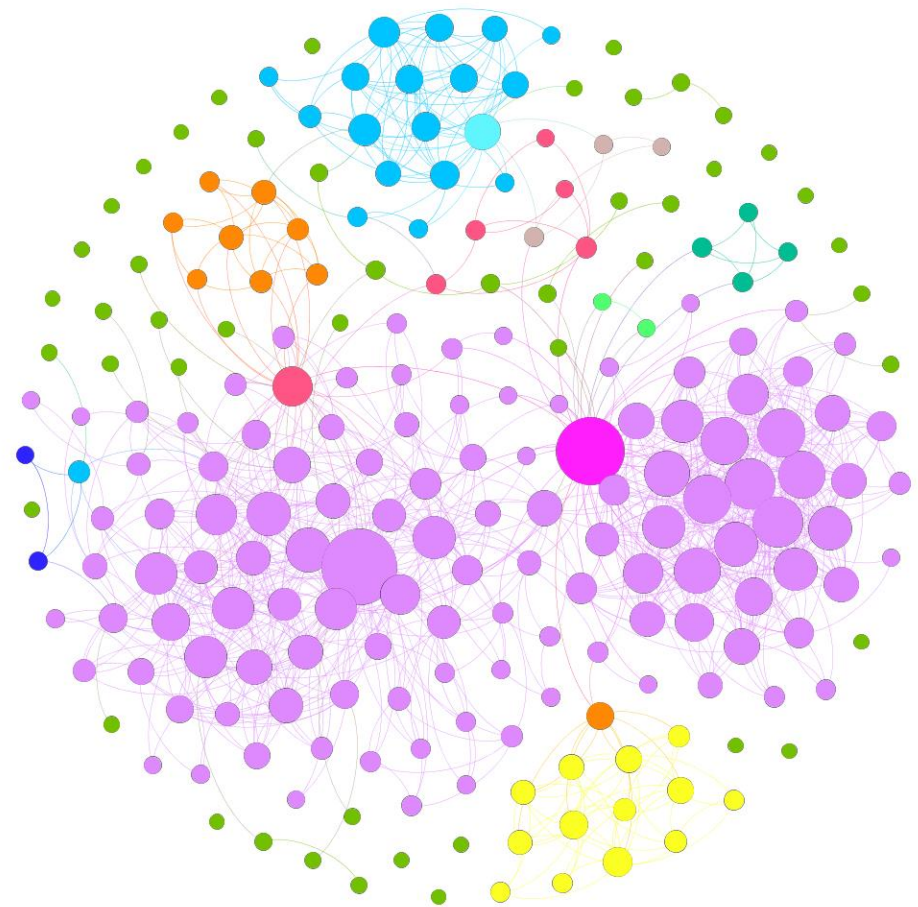
We can recognize 3 groups. Actually there are guys from: last university, current university and climb club



k-clique perlocation:

This way is very good as some measure of closeness. All groups includes mostly different people, but some groups were not identified :

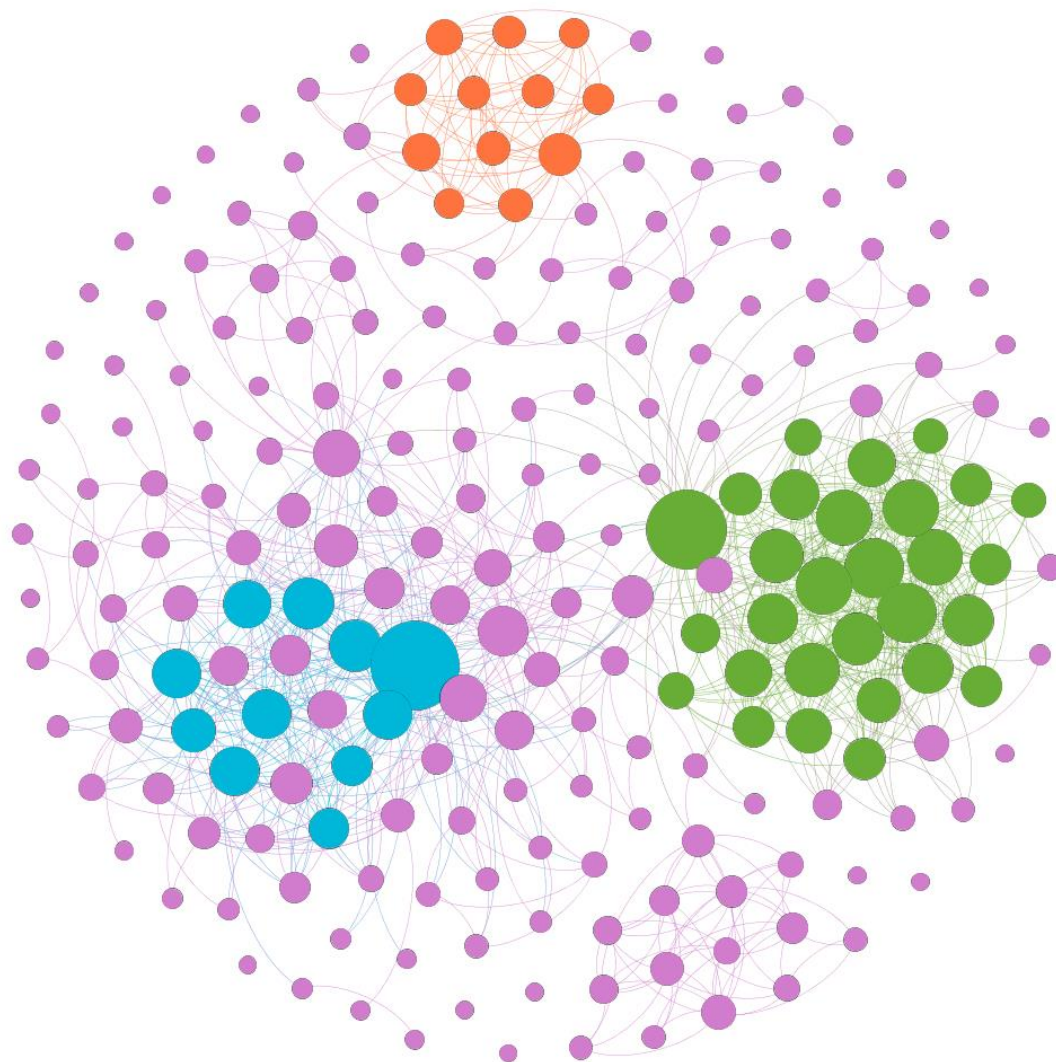
- Random acquaintances
- More close acquaintances (classmates, theatre club, closest friends of my closest friends)
- In fact - it is one guy
- Last university guys
- Current university guys
- ...



Cliques:

This way give:

- Guys from theatre
- Classmates
- Last university guys



Thanks for attention!