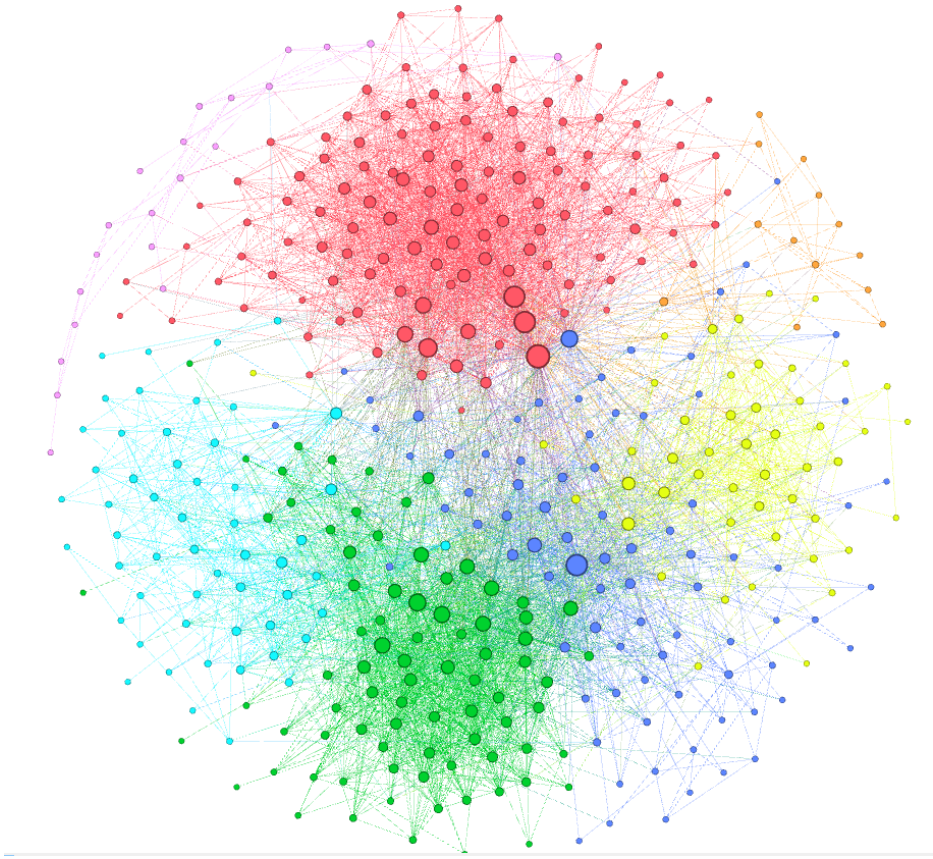
Vk network analysis

Sergey Sterkhov

Higher school of economics

March 27, 2018

Summary

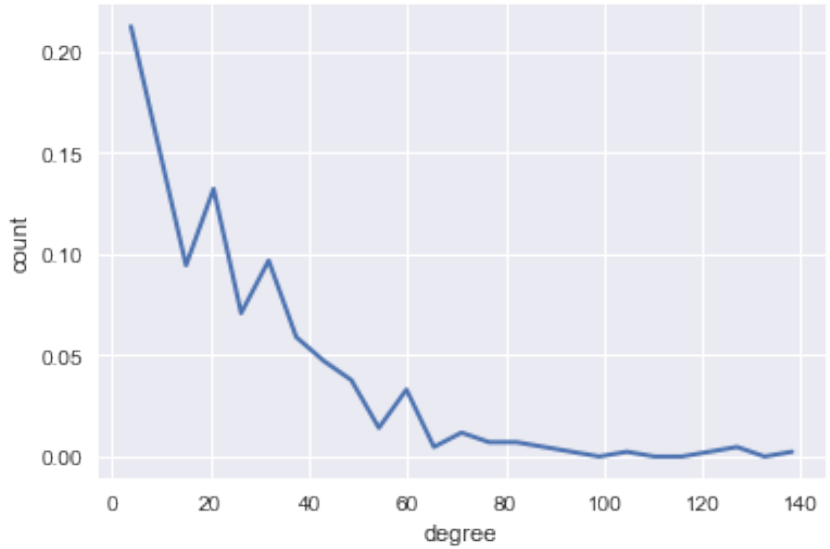


Graph of my friends in VK

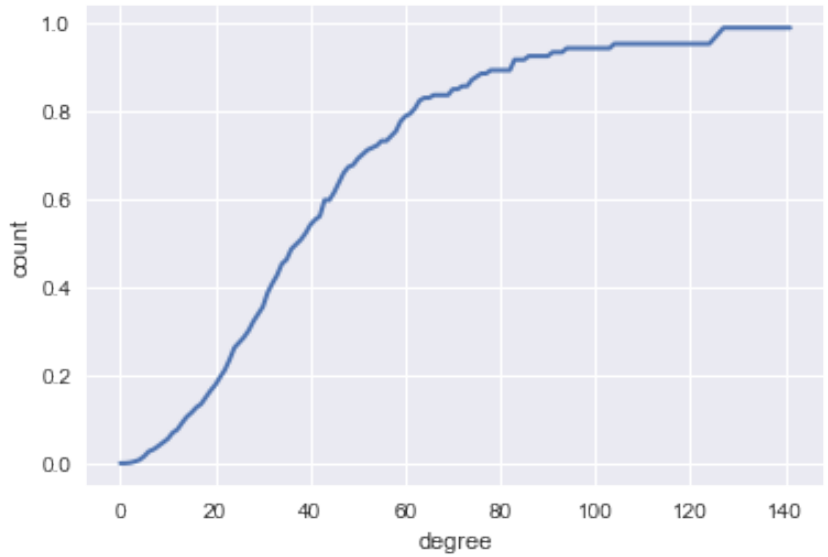
1. 424 nodes in the biggest connected component
2. 5161 edges
3. Diameter = 7
4. Clustering coefficient = 0.509
5. Average shortest length path = 2.594

Degree distribution

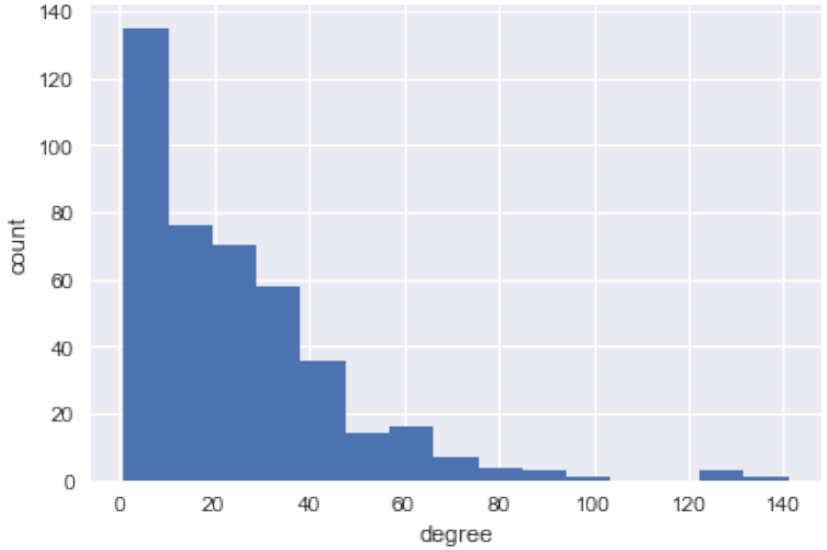
PDF



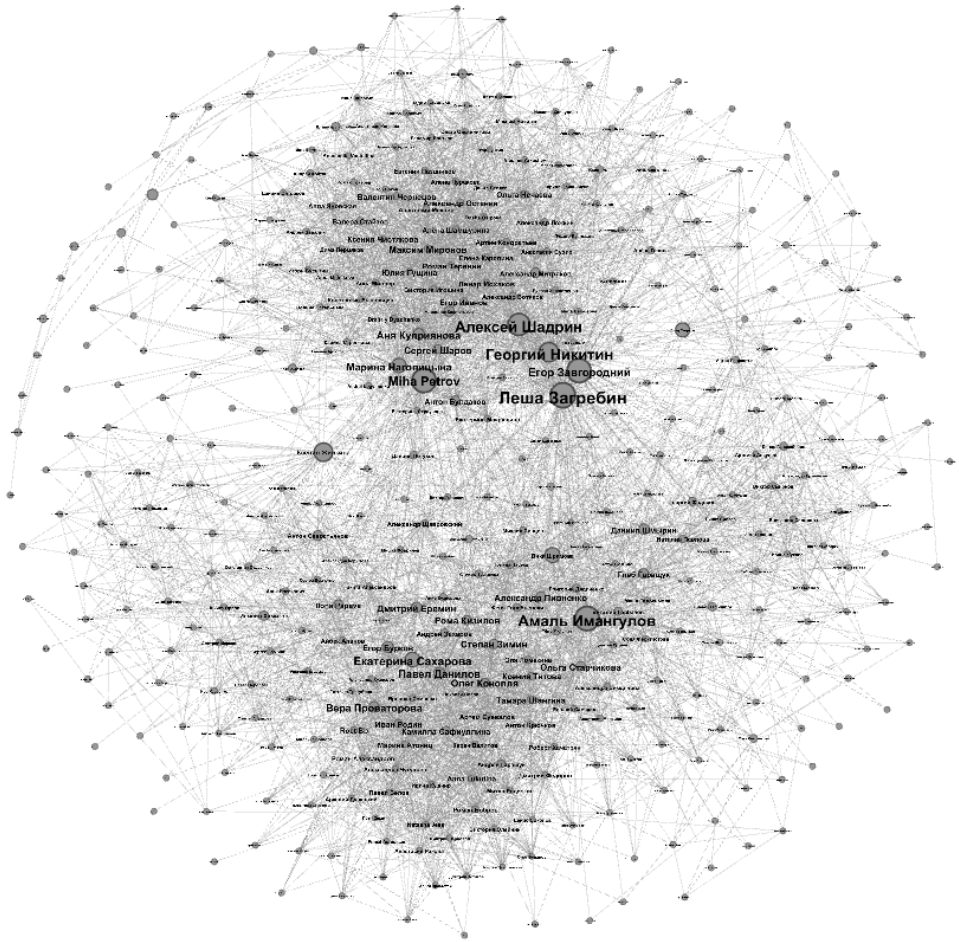
CDF



Degree distribution

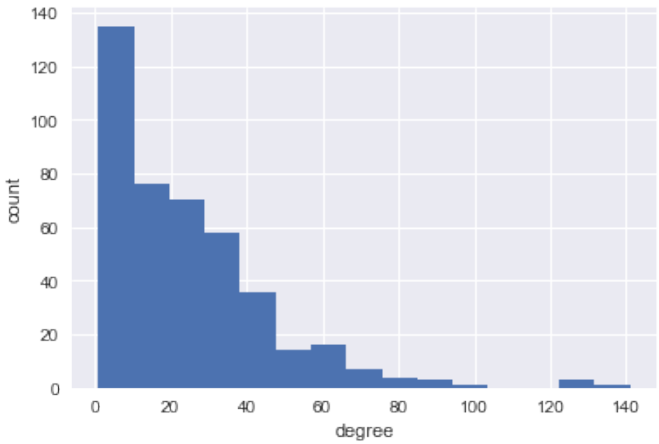


Degree centrality

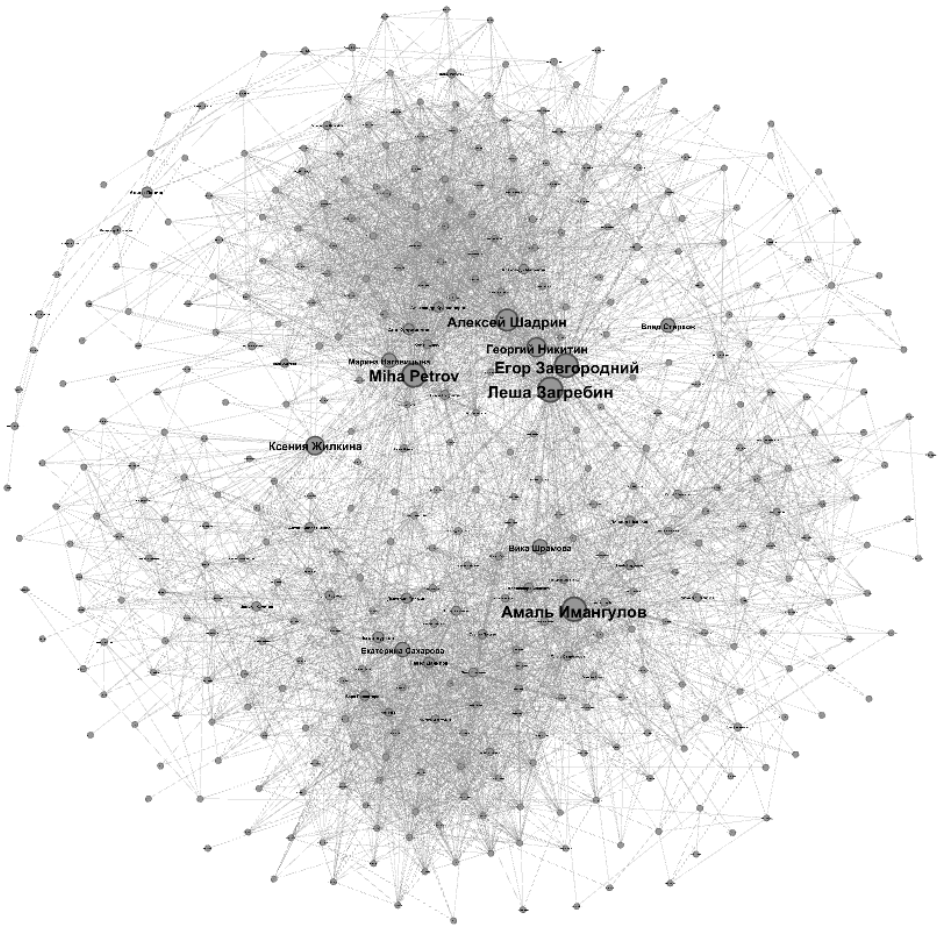


The biggest nodes are:

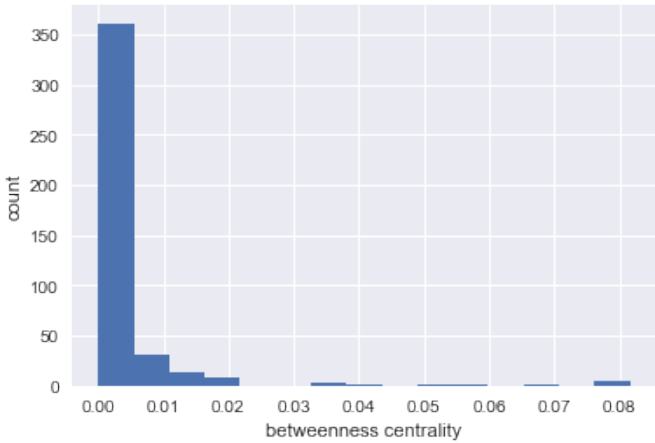
1. persons, who studied both in my school and in HSE
2. very sociable course-mates



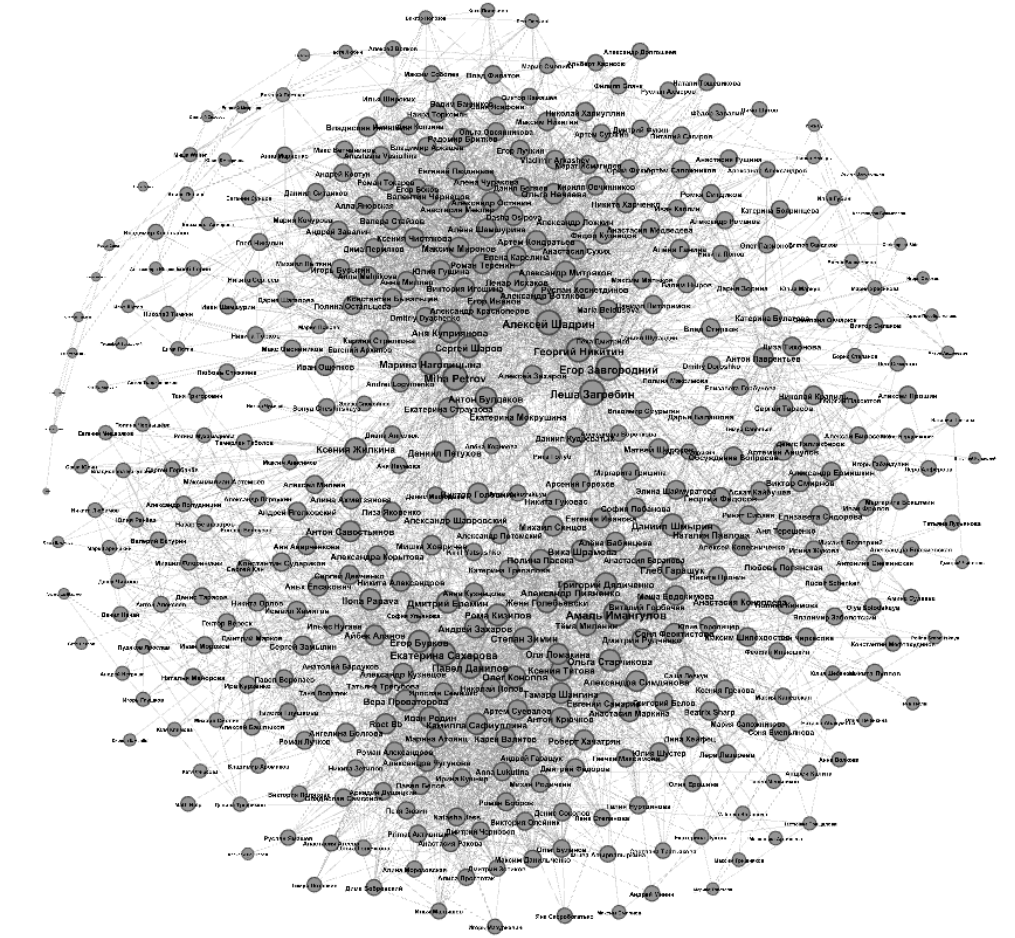
Betweenness centrality



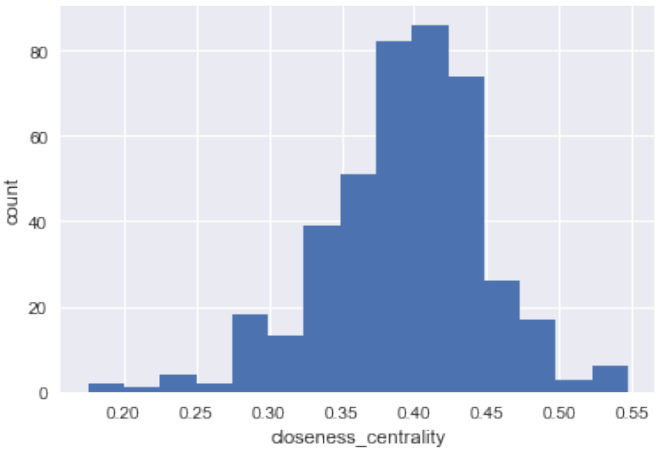
The biggest nodes are almost the same. New two persons (in comparison to the degree centrality), who have edges with almost every subgroup in the graph



Closeness centrality

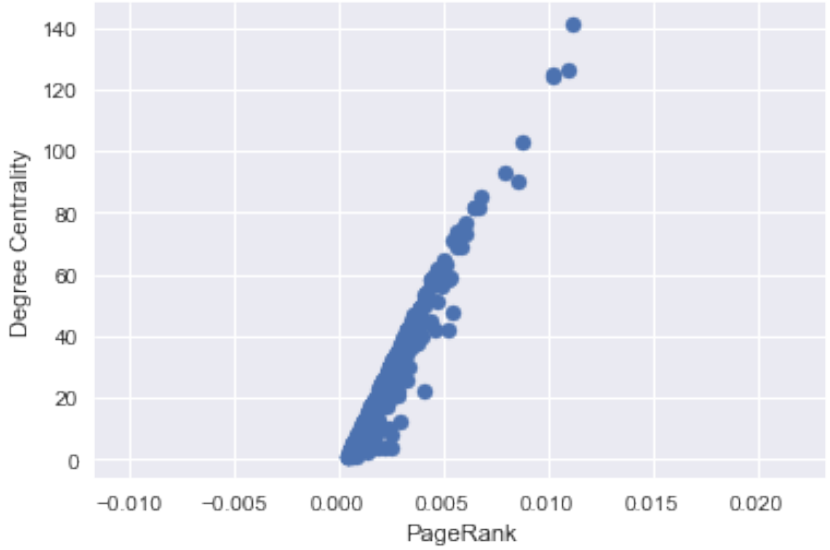


Top nodes are almost similar to the degree centrality

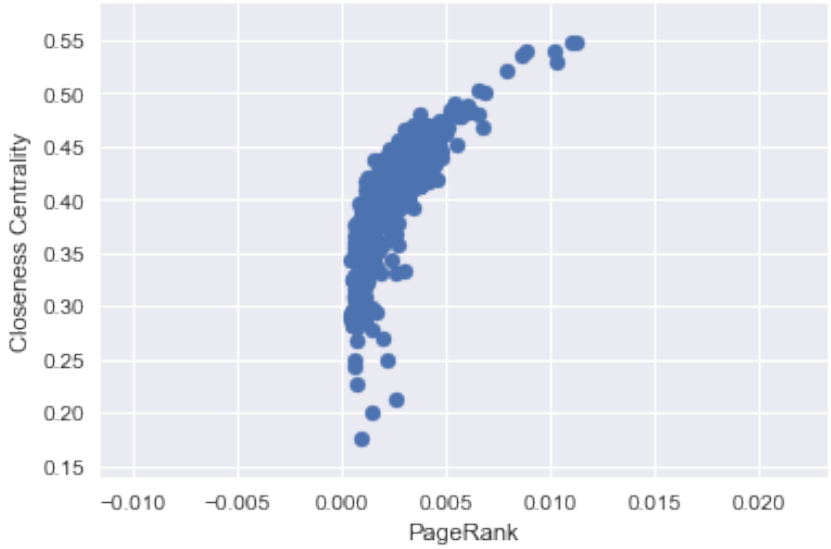


PageRank comparison

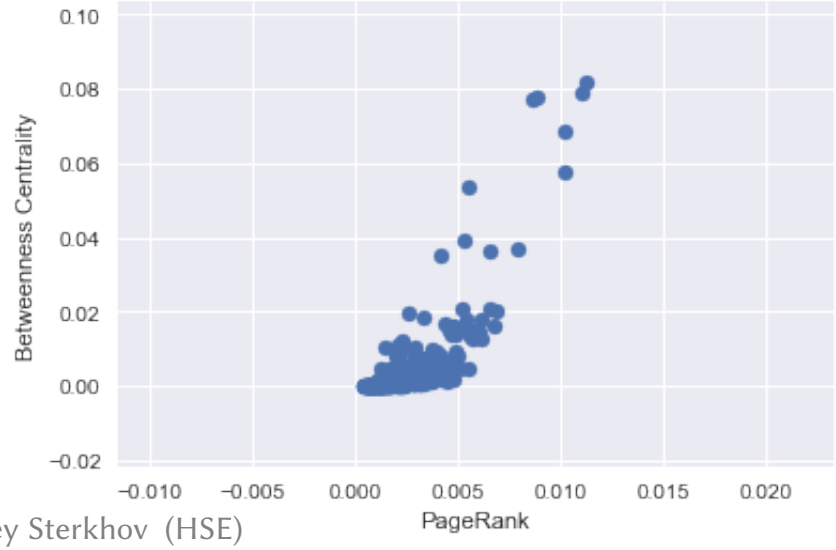
PageRank - Degree Centrality



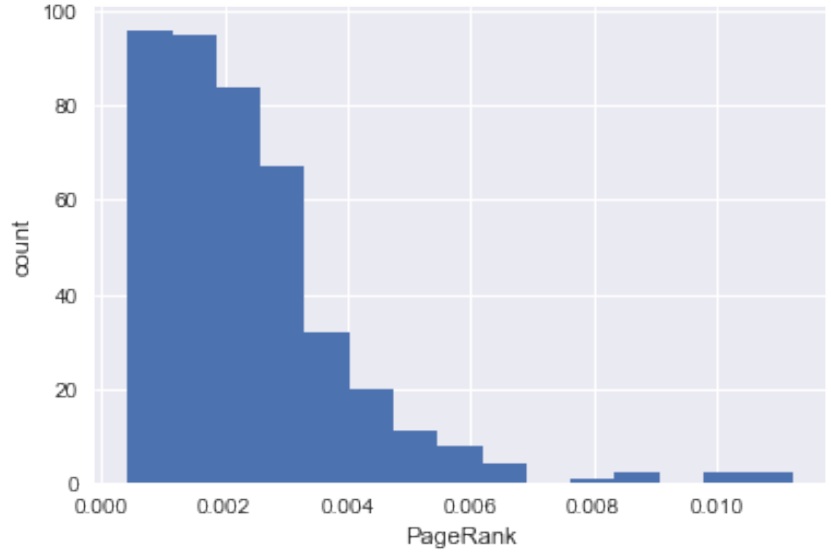
PageRank - Closeness Centrality



PageRank - Betweenness



PageRank - distribution

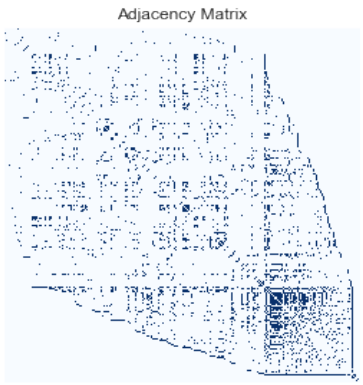


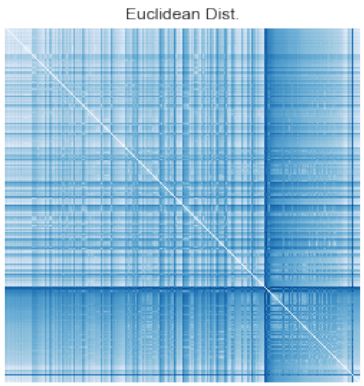
Centralities - table of top nodes

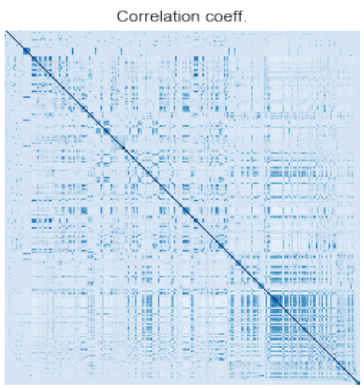
Number of row = place of the person by this centrality

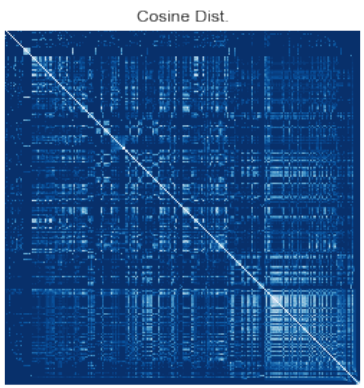
betweenness	closeness	degree	pagerank
Леша Загребин	Леша Загребин	Леша Загребин	Леша Загребин
Амаль Имангулов	Амаль Имангулов	Амаль Имангулов	Амаль Имангулов
Miha Petrov	Алексей Шадрин	Георгий Никитин	Георгий Никитин
Егор Завгородний	Miha Petrov	Алексей Шадрин	Алексей Шадрин
Алексей Шадрин	Егор Завгородний	Miha Petrov	Miha Petrov
Георгий Никитин	Георгий Никитин	Екатерина Сахарова	Егор Завгородний
Ксения Жилкина	Екатерина Сахарова	Егор Завгородний	Екатерина Сахарова
Вика Шрамова	Марина Наговицына	Павел Данилов	Павел Данилов
Екатерина Сахарова	Павел Данилов	Вера Проваторова	Вера Проваторова
Марина Наговицына	Даниил Шмырин	Аня Куприянова	Марина Наговицына

Similarities









Similar nodes by cosine distance

friend_1	friend_2	dist
Машенька Артищева	Максим Гришняков	0.0000
Машенька Артищева	Максим Смирнов	0.0000
Максим Гришняков	Максим Смирнов	0.0000
Ирка Бу	Игорь Стерхов	0.1340
Александр Останин	Ленар Исхаков	0.1379
Елизавета Сидорова	Виктор Смирнов	0.1511
Ирина Жукова	Виктор Смирнов	0.1546

In similarity top:

1. nodes with small degree
2. persons with the same schoolclass and university
3. persons from What?Where?When?

Random Graphs

Generated Erdos-Renyi, Watts-Strogatz and Barabasi-Albert models with different parameters. Best results are presented:

My graph

1. diameter = 7.0
2. clustering coef = 0.509
3. avg path = 2.59
4. edges count = 5161

Watts-Strogatz (p=0.5, k=92)

1. diameter ~ 8.31
2. clustering coef ~ 0.16
3. avg path ~ 4
4. edges count ~ 9964

Erdos-Renyi (p=0.1)

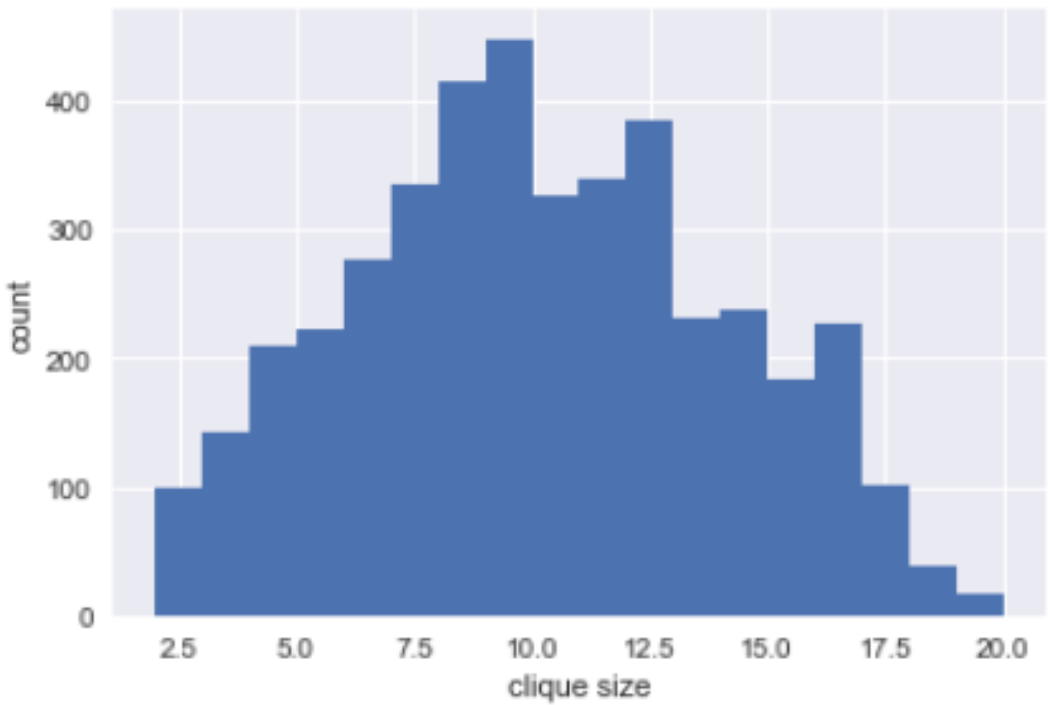
1. diameter ~ 3.0
2. clustering coef ~ 0.1
3. avg path ~ 1.91
4. edges count ~ 8986

Barabasi-Albert (m=10)

1. diameter ~ 3.4
2. clustering coef ~ 0.11
3. avg path ~ 2.3
4. edges count ~ 4140

Lowest value of average deviation is achieved by Barabasi-Albert model

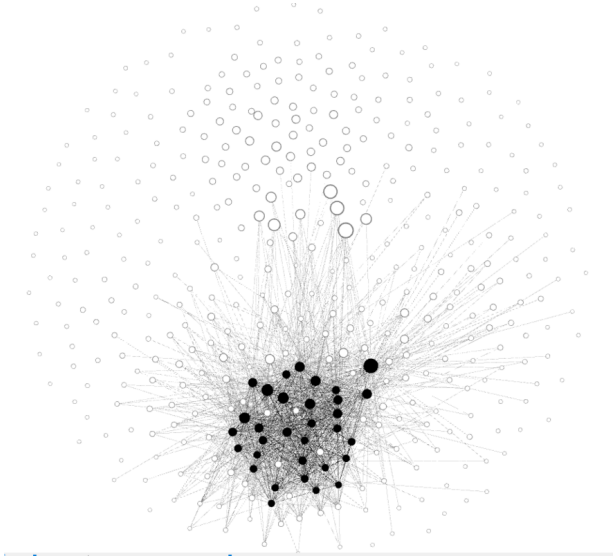
Cliques



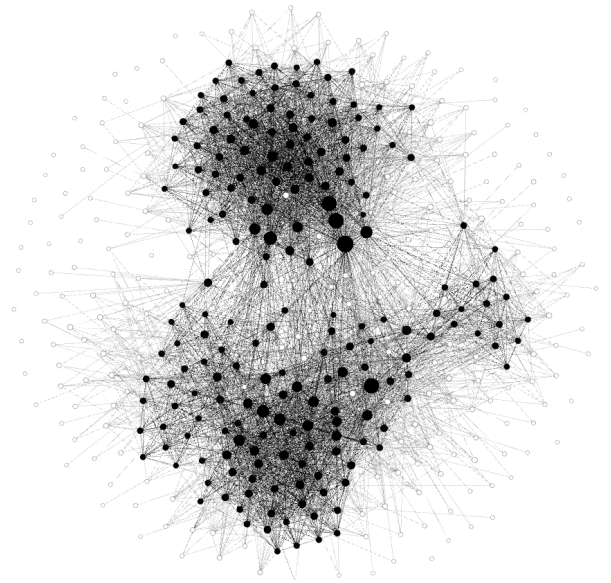
1. maximal cliques number = 4
2. size of maximal clique = 20
3. maximal clique contains course-mates and (sud-denly) one guy from the faculty of law

K-core

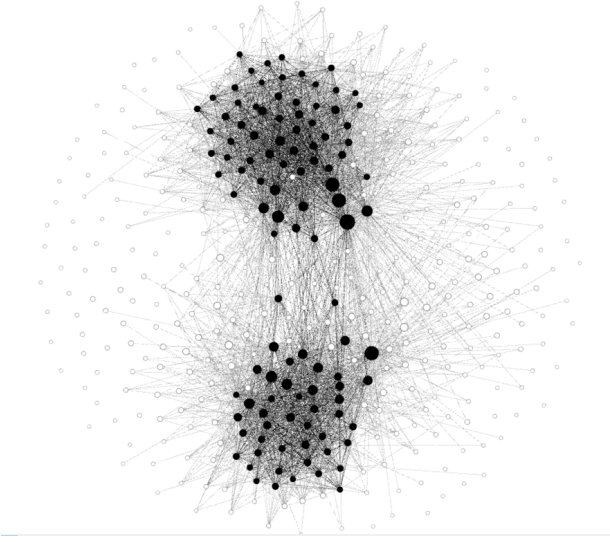
24-core (maximal)



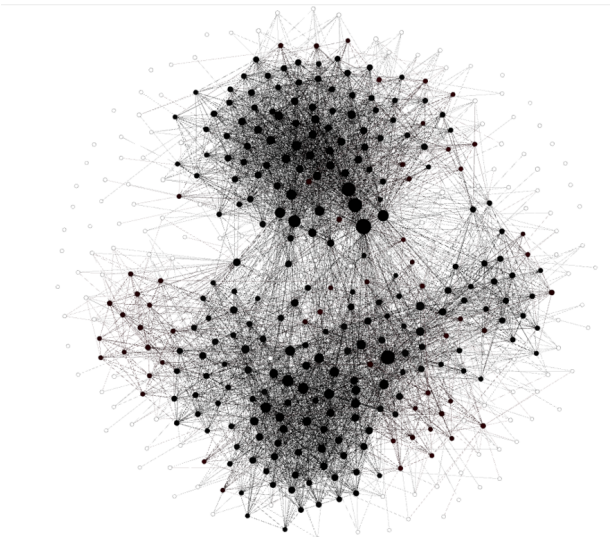
17-core



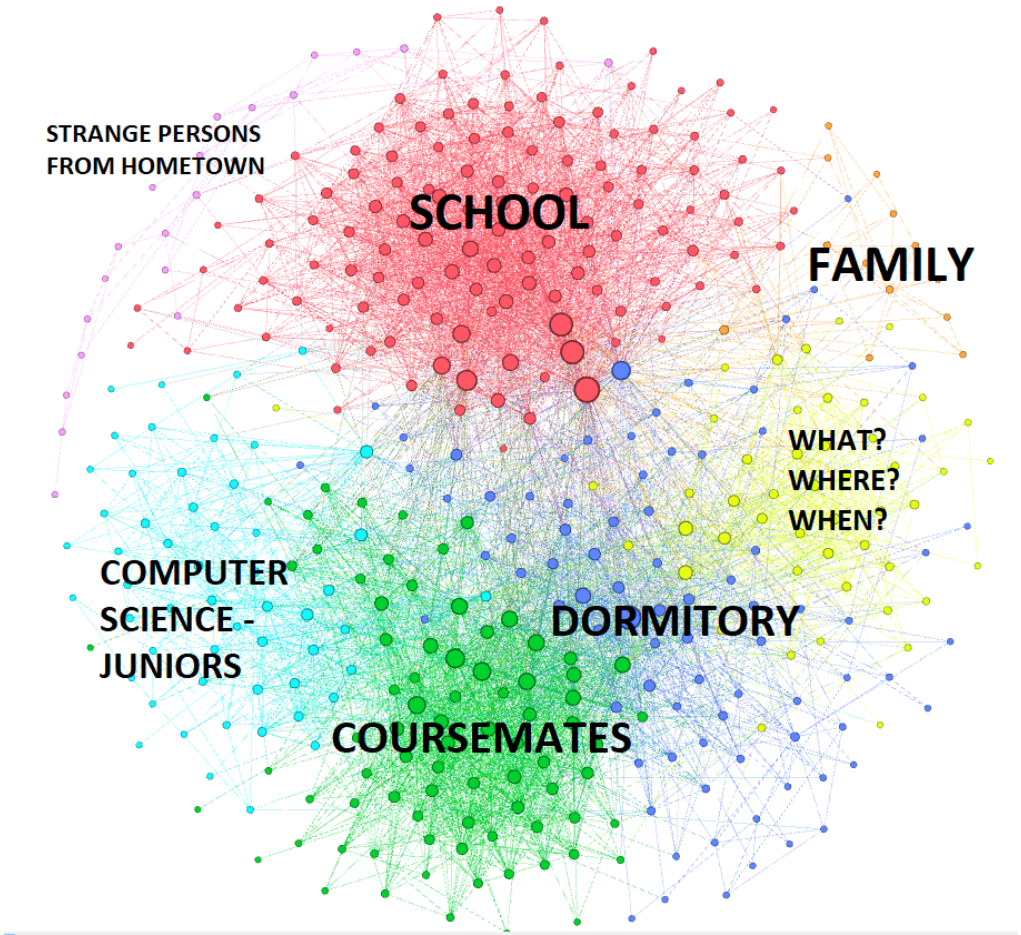
23-core



11-core



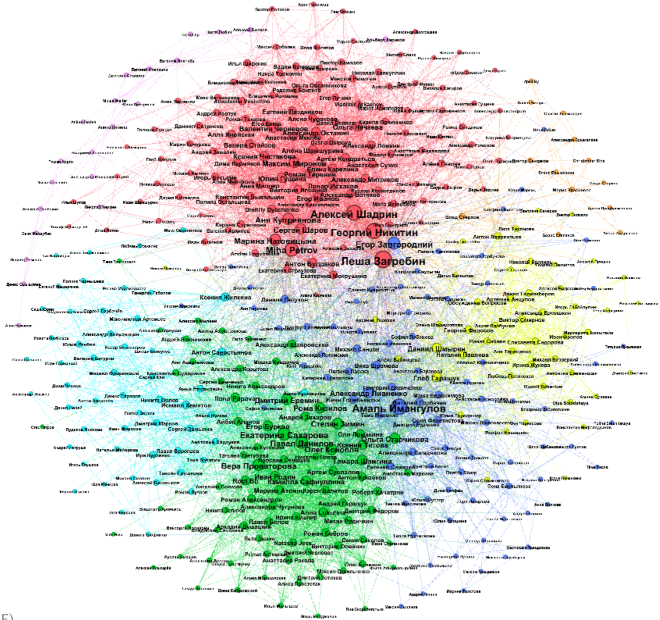
Modularity



7 clusters:

1. Schoolmates
2. Coursemates
3. Persons from the same faculty, but younger
4. Friends from the dormitory and other faculties
5. Persons from What?Where?When?
6. Family
7. Persons from parties and karate in hometown

Modularity - names



Thank you for your attention