

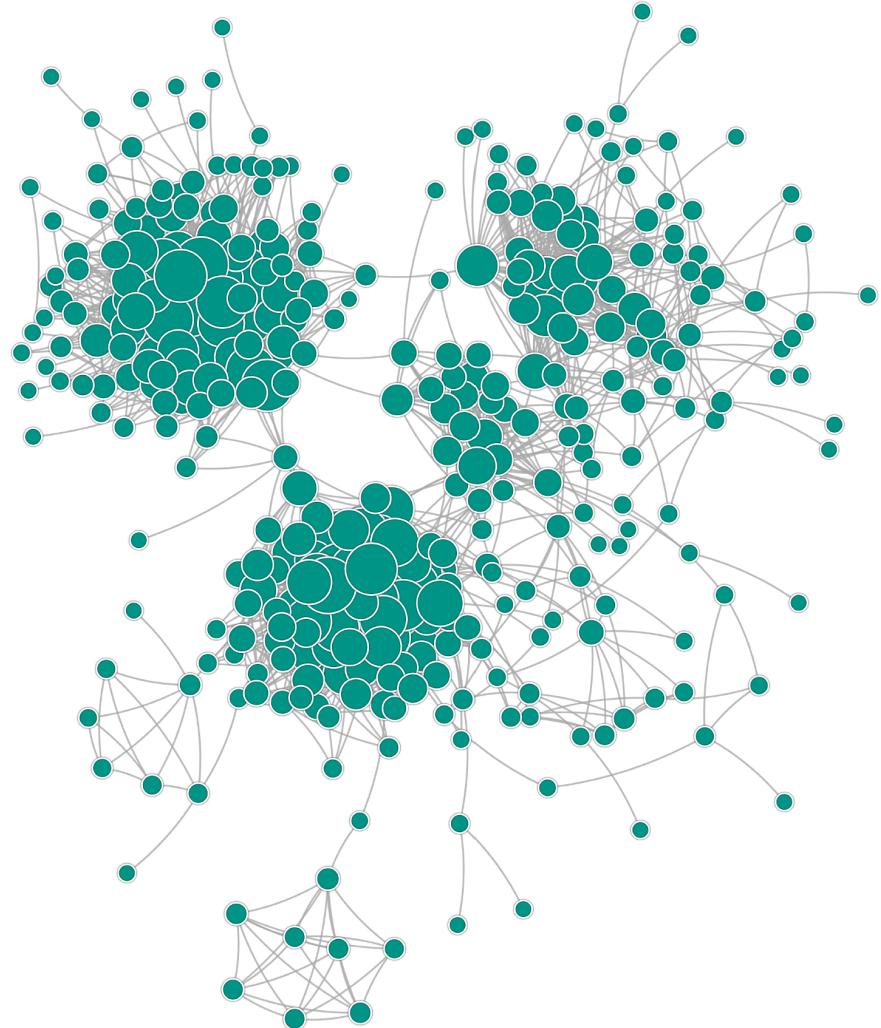
# Network Science Project

VK friends network analysis

Prepared by  
Pugachev Alexander

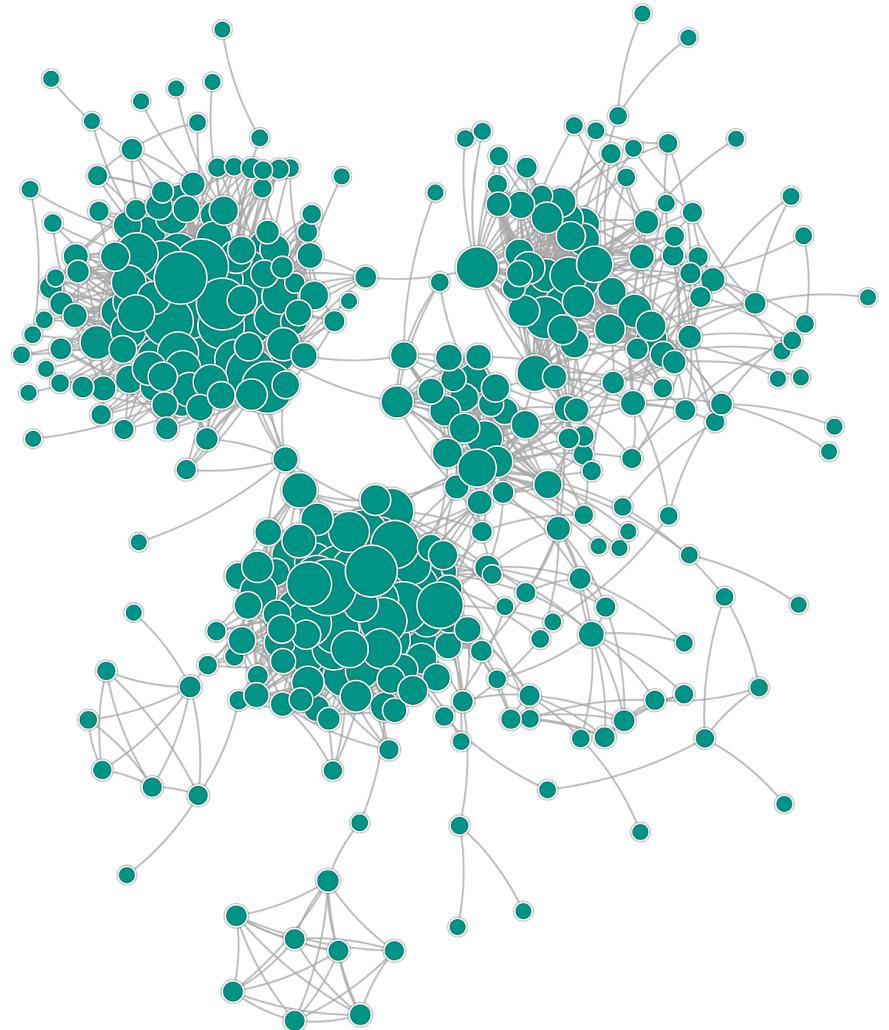
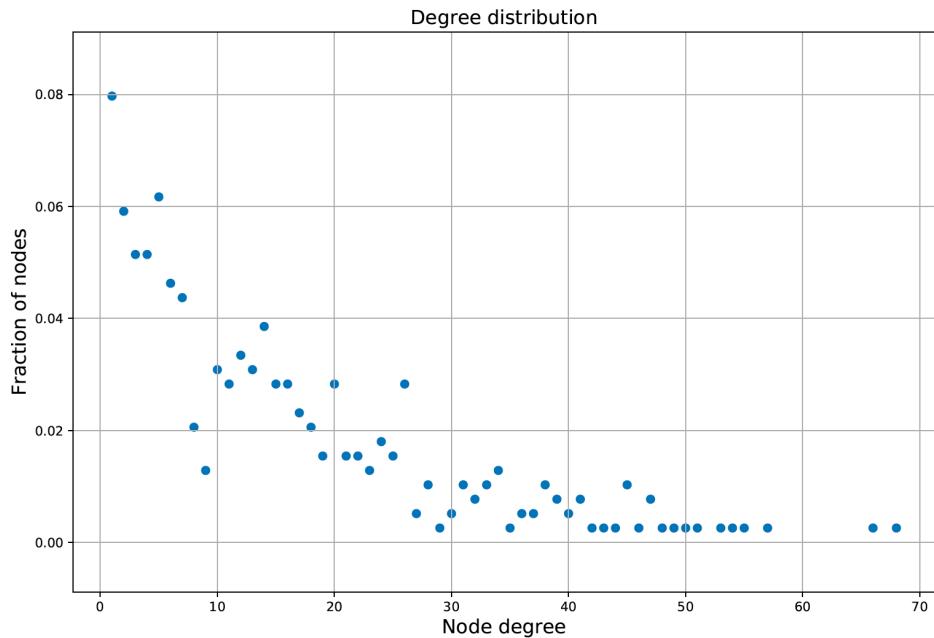
# Network summary

- Number of nodes: 388
- Number of edges: 2885
- Node attributes:
  - Name
  - Sex
  - City
  - Studying at HSE
  - Participation in Wylsacom VK group



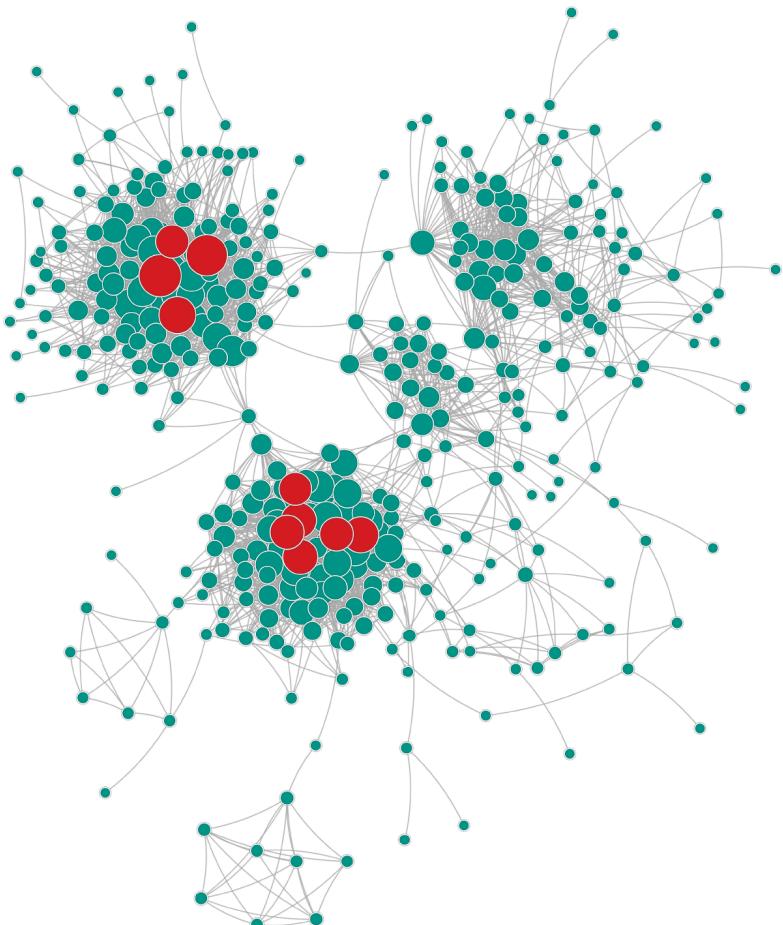
# Network summary

- Diameter: 10
- Radius: 6
- Clustering coefficient: 0.449
- Average shortest path: 4.103

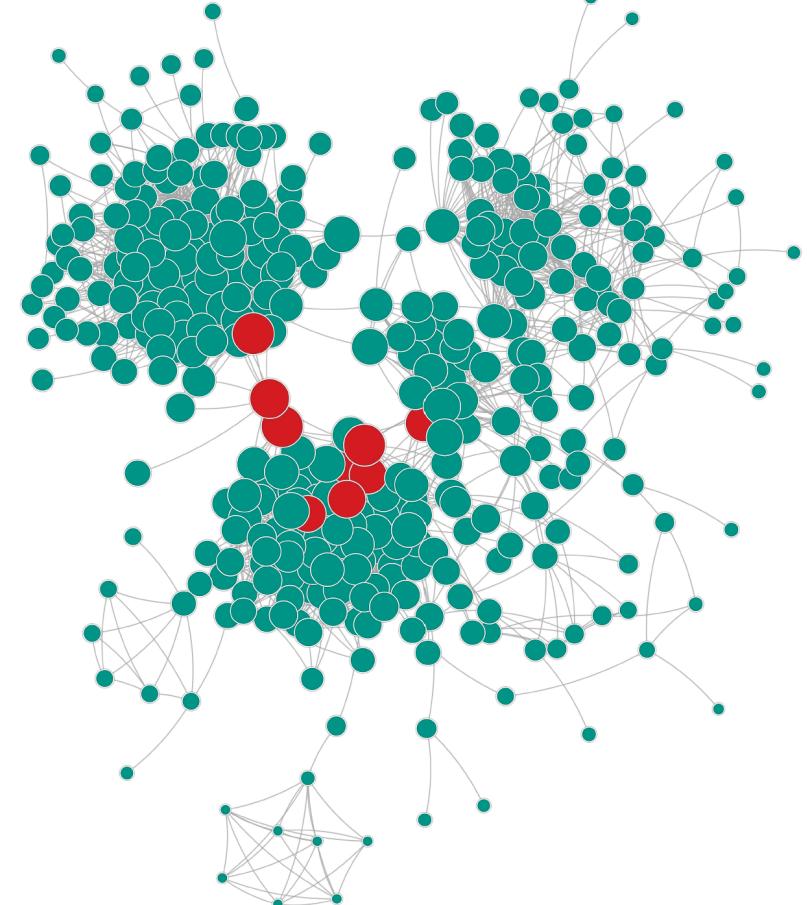


# Centralities

Degree centrality



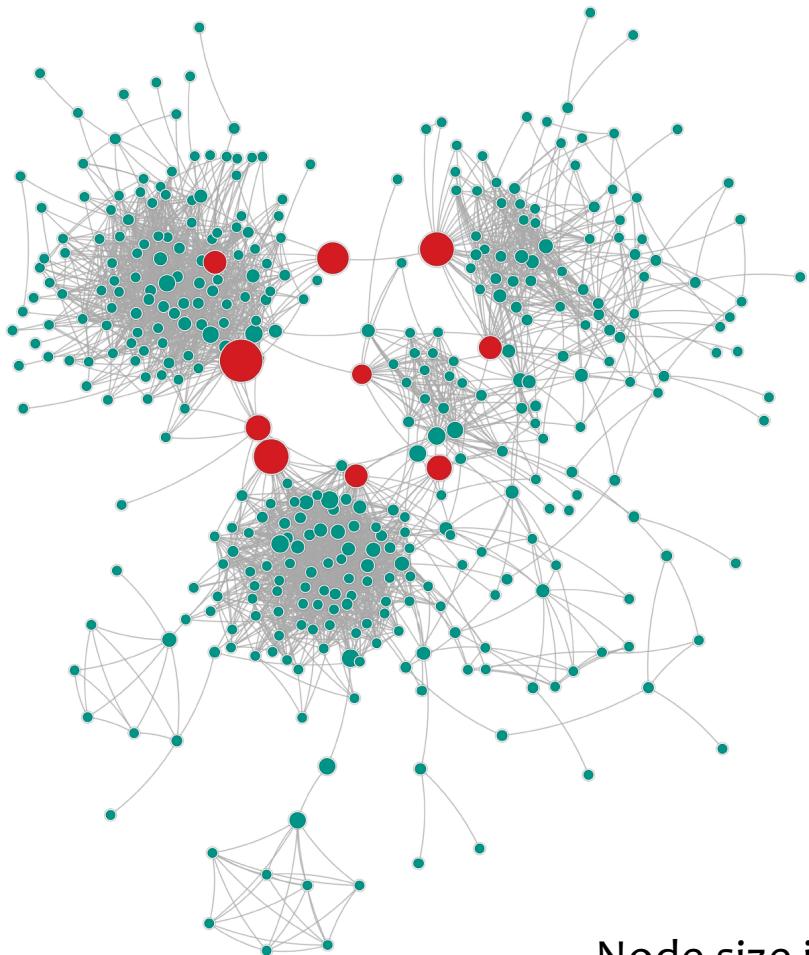
Closeness centrality



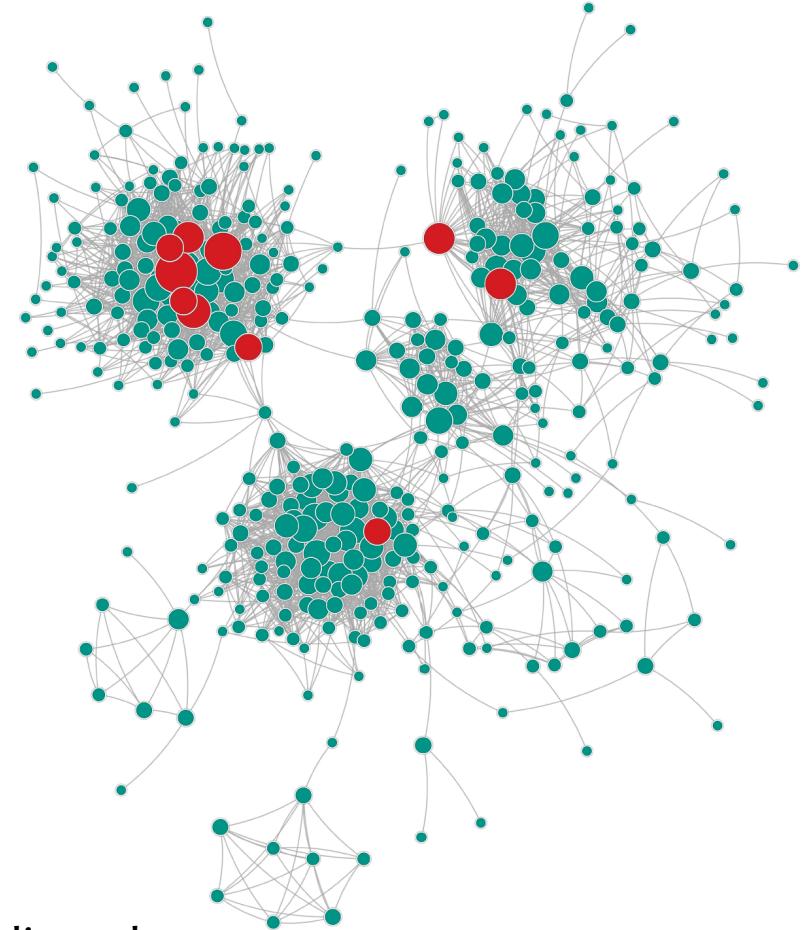
Node size is shown according to centrality value  
Top 10 nodes are colored red

# Centralities

Betweenness centrality



Pagerank centrality



Node size is shown according to centrality value  
Top 10 nodes are colored red

# Top nodes by centralities

	Degree Centrality	Closeness centrality	Betweenness centrality	Pagerank
1	Александр Ермачков	Христофор Парфиров	Христофор Парфиров	Александр Ермачков
2	Pavel )))	Александр Проскуряков	Стас Рыбин	Pavel )))
3	Georgiy Demenchuk	Стас Рыбин	Степан Куртев	Georgiy Demenchuk
4	Валерий Батурин	Ксения Жилкина	Cjb Bot	Степан Куртев
5	Денис Тарасов	Антон Таныгин	Антон Таныгин	Андрей Быстрицкий

Blue color: people from Wylsacom group

Orange color: people who studied in HSE

Green color: people from Kishinev

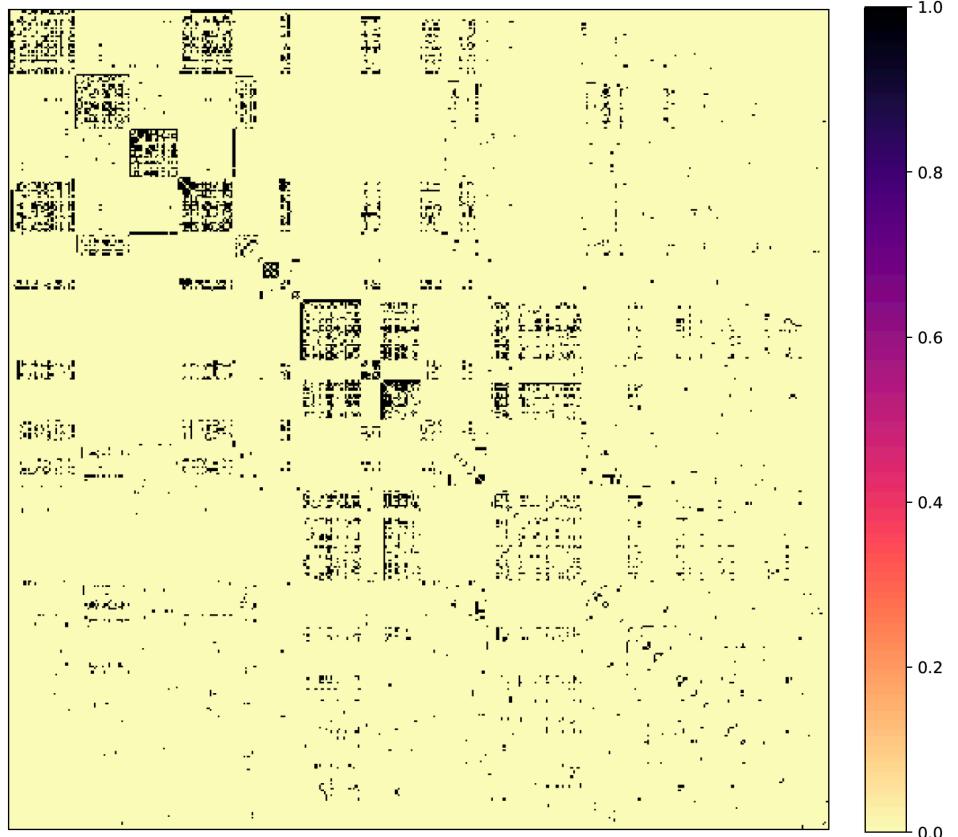
# Assortative mixing

Attribute	Sex	City	HSE	Wylsacom
Assortativity coefficient	0.245	0.107	0.394	0.199

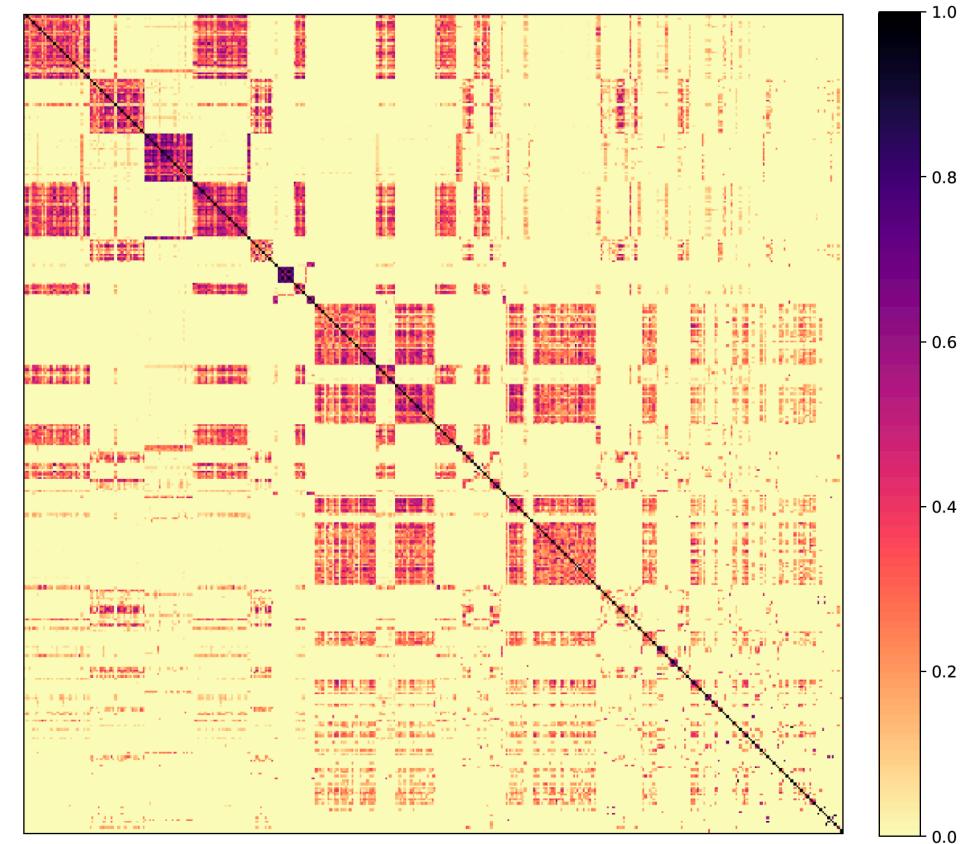
- Two people from HSE are more likely friends than two people with the same gender or from the same city
- Two people with the same gender are more likely friends than two people with different gender

# Node structural similarity

Adjacency matrix

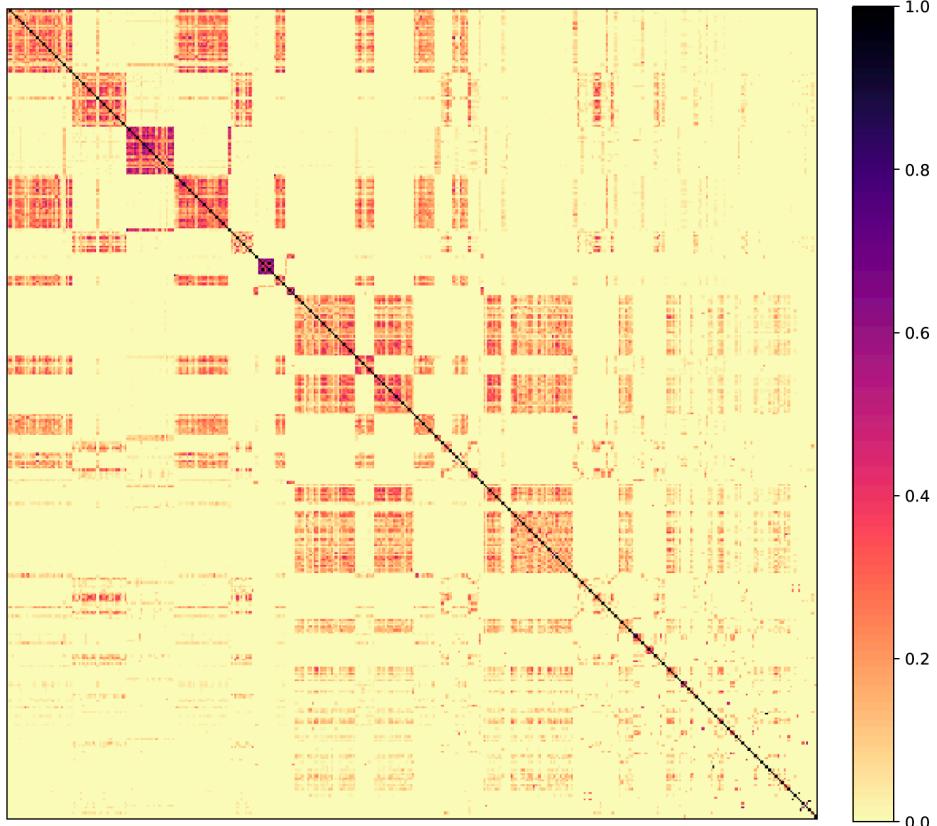


Cosine similarity matrix

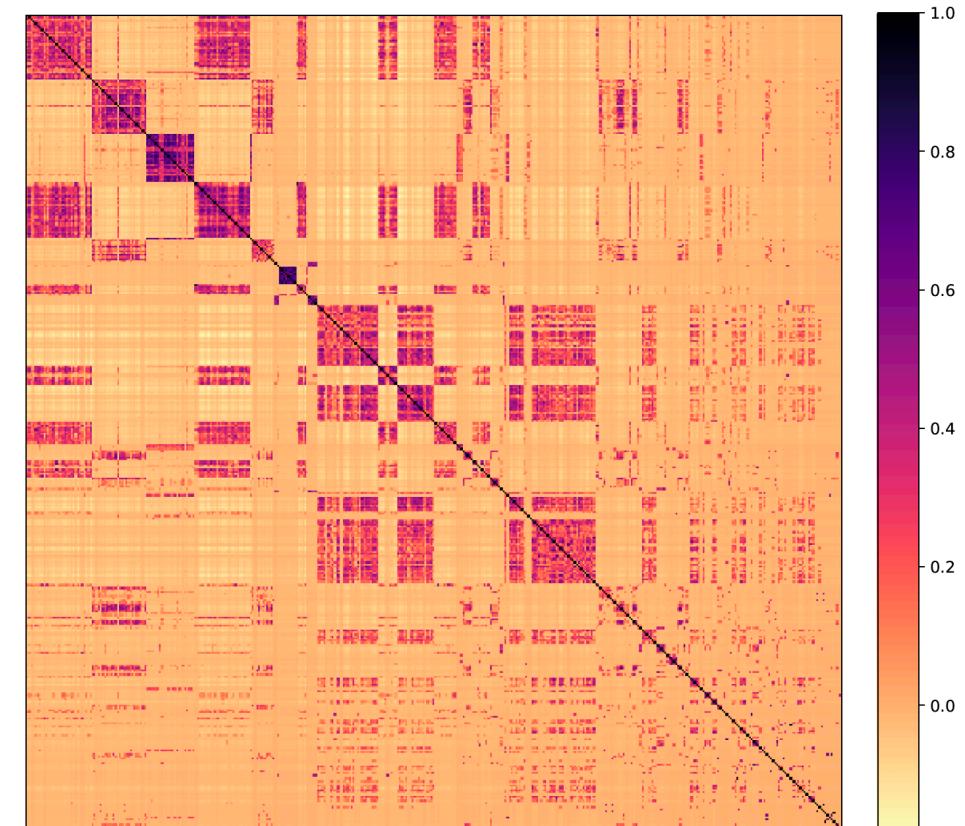


# Node structural similarity

Jaccard similarity matrix



Pearson correlation matrix



# Top by cosine similarity

#	Node name 1	Node name 2	Cosine similarity
1	Лена Мимоглядова	Наталия Селюто	1.0
2	Ирина Белоусова	Владислав Бадеха	1.0
...			
7	Настя Лебедева	Николай Цалко	1.0
8	Дмитрий Карфидов	Николай Герасименко	0.912
9	Дина Нагибина	Дмитрий Карфидов	0.912
10	Георгий Плотников	Андрей Поляков	0.894

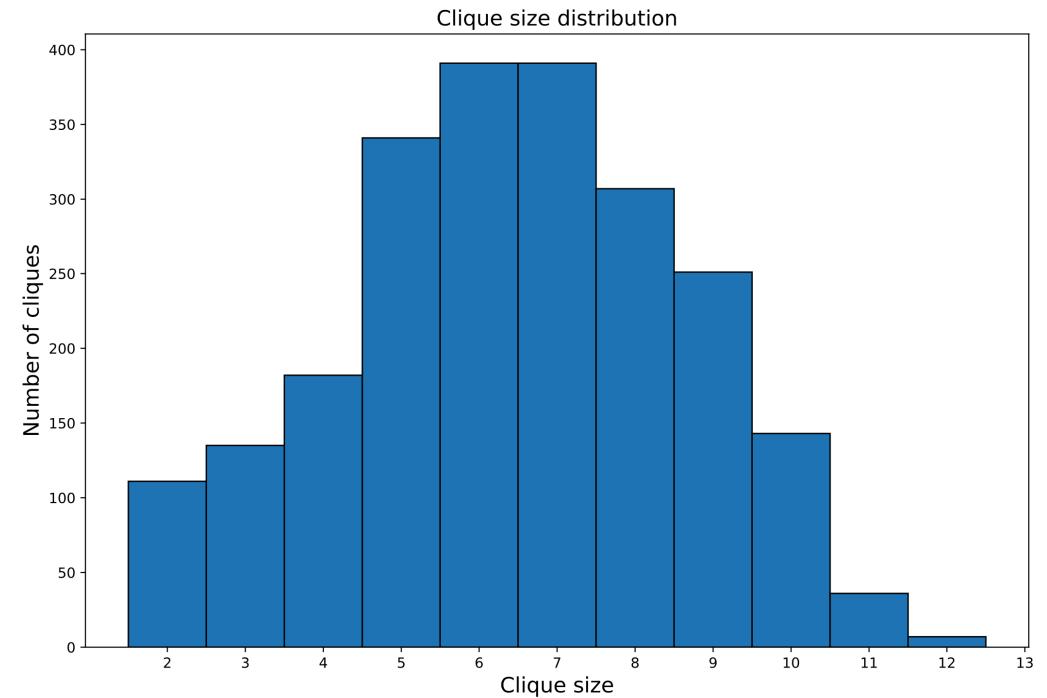
- Node pairs with cosine similarity equal to 1.0 have one and only common friend
- For Jaccard similarity and Pearson correlation Top – 10 is the same

# Closest random graph model

	<b>Friends Network</b>	<b>Edros – Renyi</b>	<b>Watts – Strogatz</b>	<b>Barabasi - Albert</b>
Parameters		$N = 388$ $P = 0.04$	$N = 388$ $K = 15$ $P = 0.05$	$N = 388$ $M = 8$
Number of nodes	388	388	388	388
Number of edges	2931	2903	2716	3048
Average node degree	15.069	14.963	14.0	15.670
Average shortest path	4.091	2.500	3.109	2.400
Clustering coefficient	0.449	0.038	0.489	0.104
Diameter	10	4	5	4

# Clique search

- Total number of cliques: 2218
- Size of the biggest clique: 12
- Number of biggest cliques: 7
- Number of unique nodes from biggest cliques: 20



# Nodes which are in biggest cliques

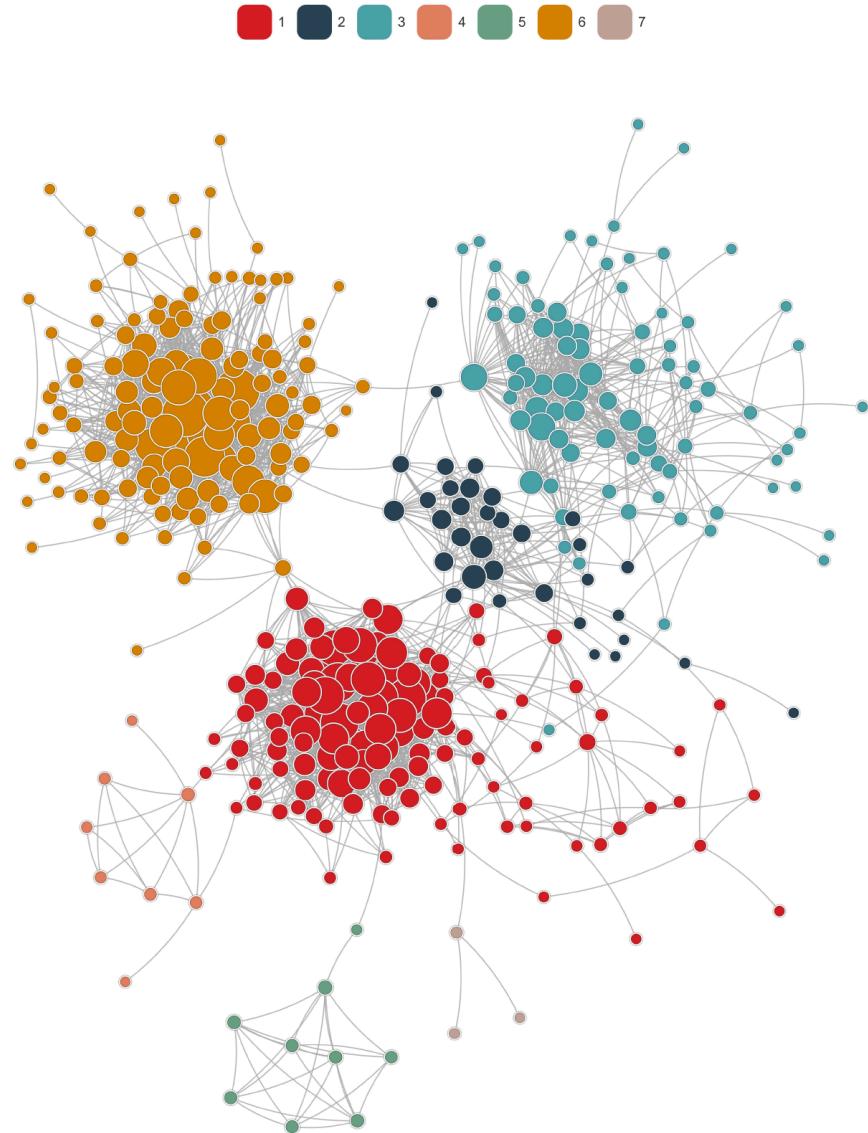
Андрей Соколов	Иван Толкушкин	Дмитрий Кириллов	Валерий Батурин	Илья Шубаев
Ксения Жилкина	Даша Чижова	Александр Цой	Сергей Павлов	Гектор Вереск
Денис Тарасов	Дарья Корепанова	Екатерина Докучаева	Астра Никитина	Антон Наумов
Денис Шаклеин	Юлия Ренёва	Оля Антонова	Pavel Хрушков	Виктория Тороп

- All these people are those with whom I studied at HSE
- Green color : my former group mates

# Community detection

## Edge Betweenness method

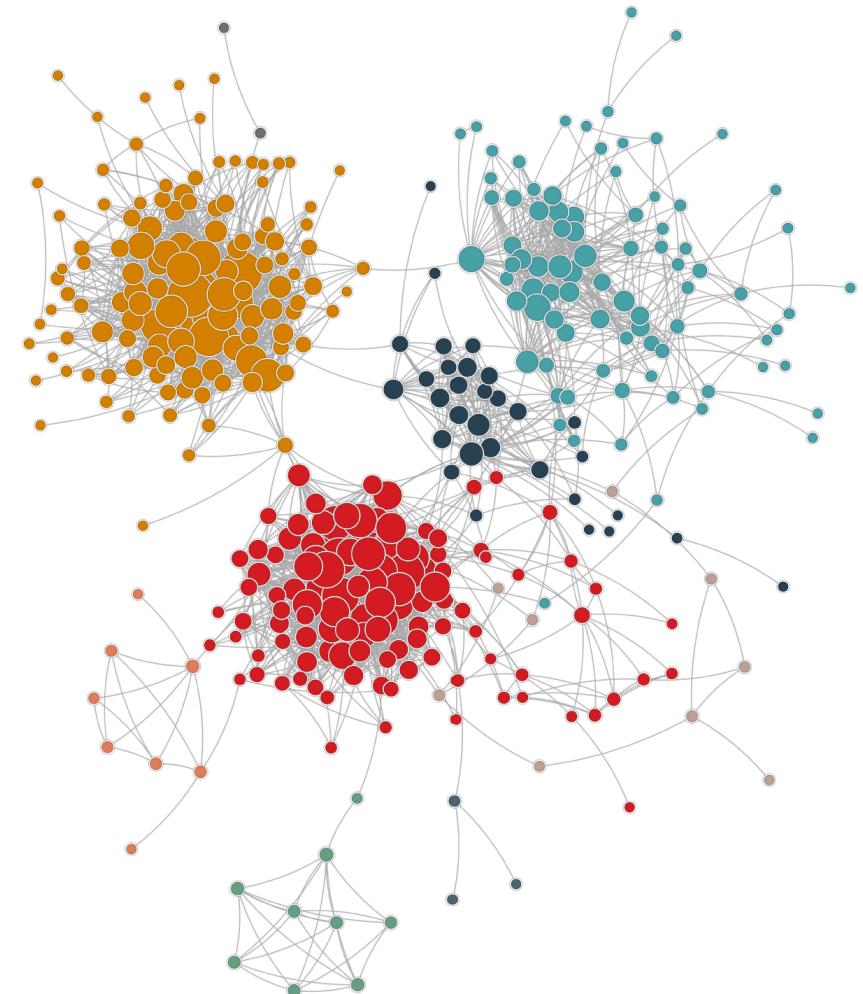
- Detected 6 prospective communities:
  1. HSE students
  2. People from HSE summer camp
  3. People from Kishinev
  4. People from HSE winter school
  5. Former Sberbank Internship colleagues
  6. People from Wylsacom VK group
- And one more community:
  7. Clash of Clans clanmates



# Community detection

## FastGreedy method

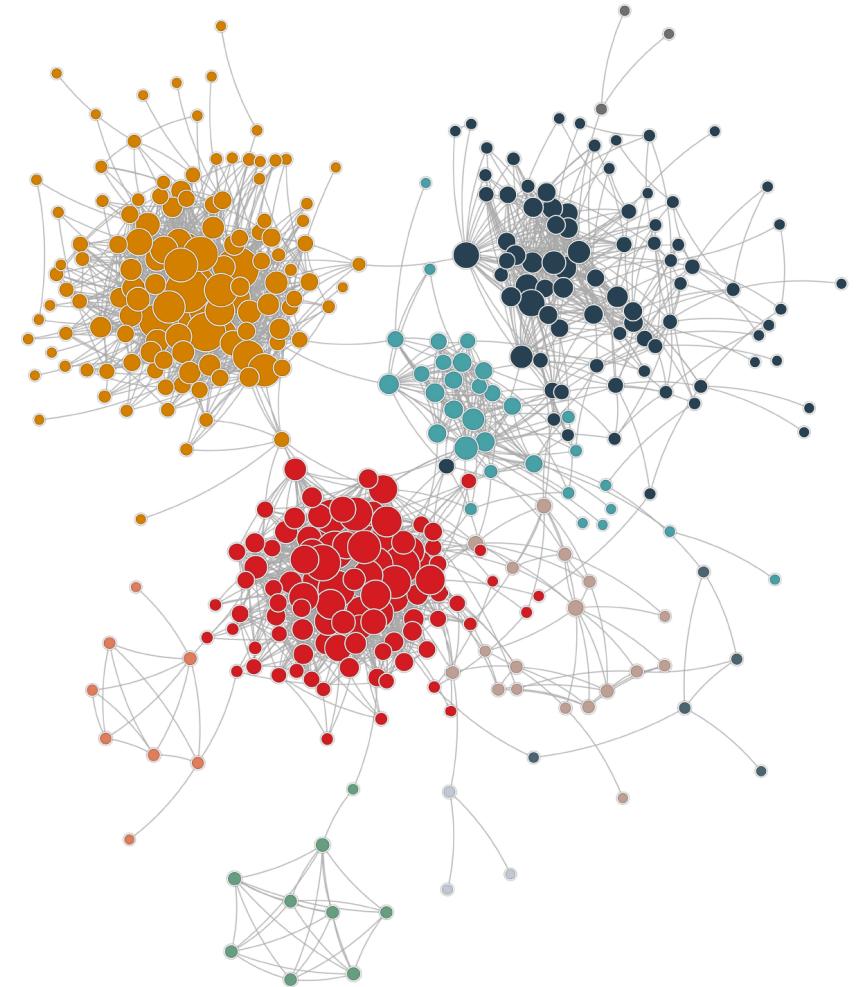
- Detected 7 above mentioned communities
- Distinguished a separate community from HSE students
- Distinguished a two-node community from Wylsacom cluster



# Community detection

## WalkTrap method

- Detected 10 communities
- 7 of them are above mentioned
- Divided HSE cluster into 3 clusters
- Distinguished a cluster which consists of my mom and 2 her friends



# Community detection

Modularity score

Method	Edge Betweennes	FastGreedy	WalkTrap
Modularity	0.654	0.655	0.662

Full code and interactive visualizations are available here:

<https://github.com/apugachev/network-science-project>