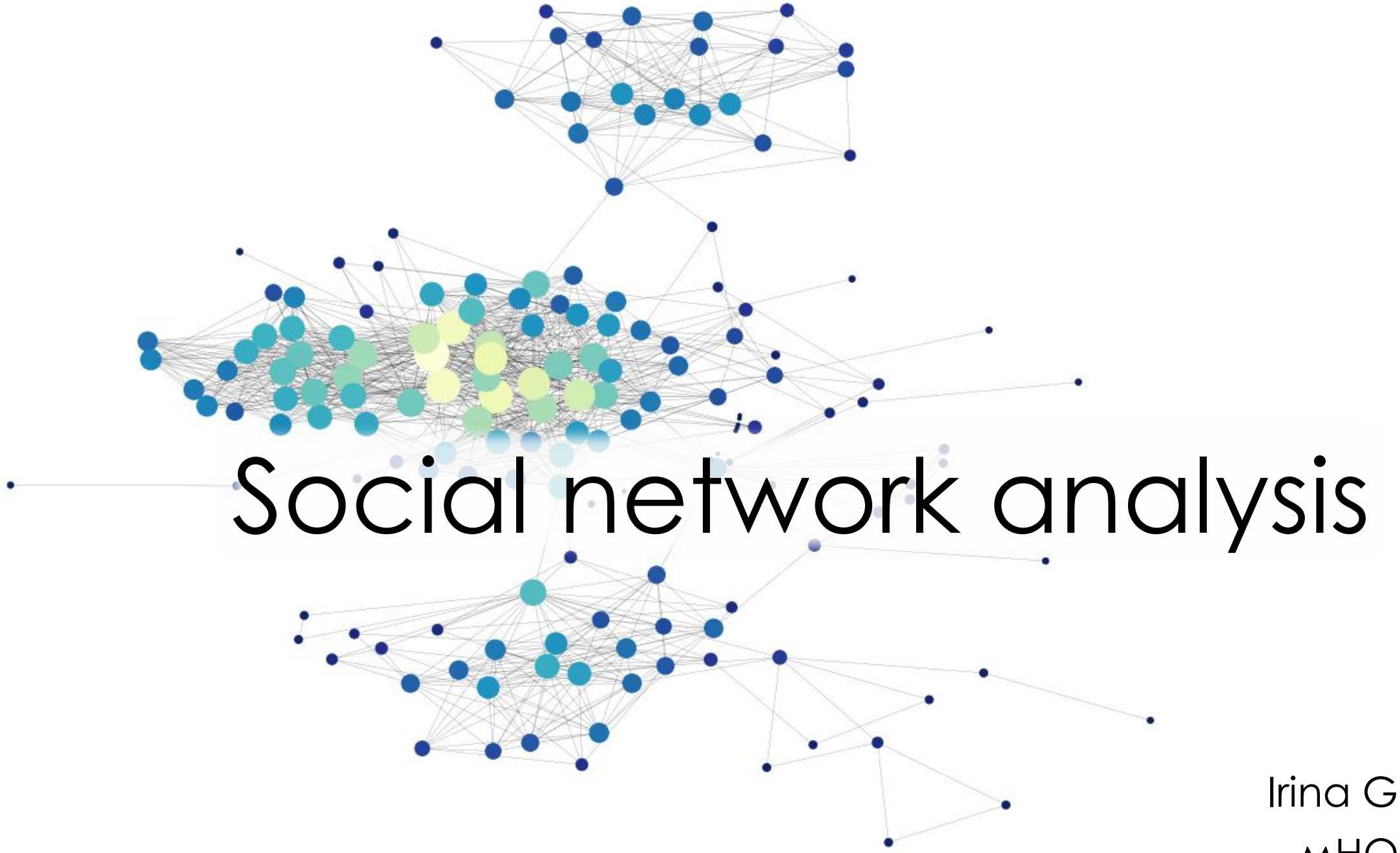


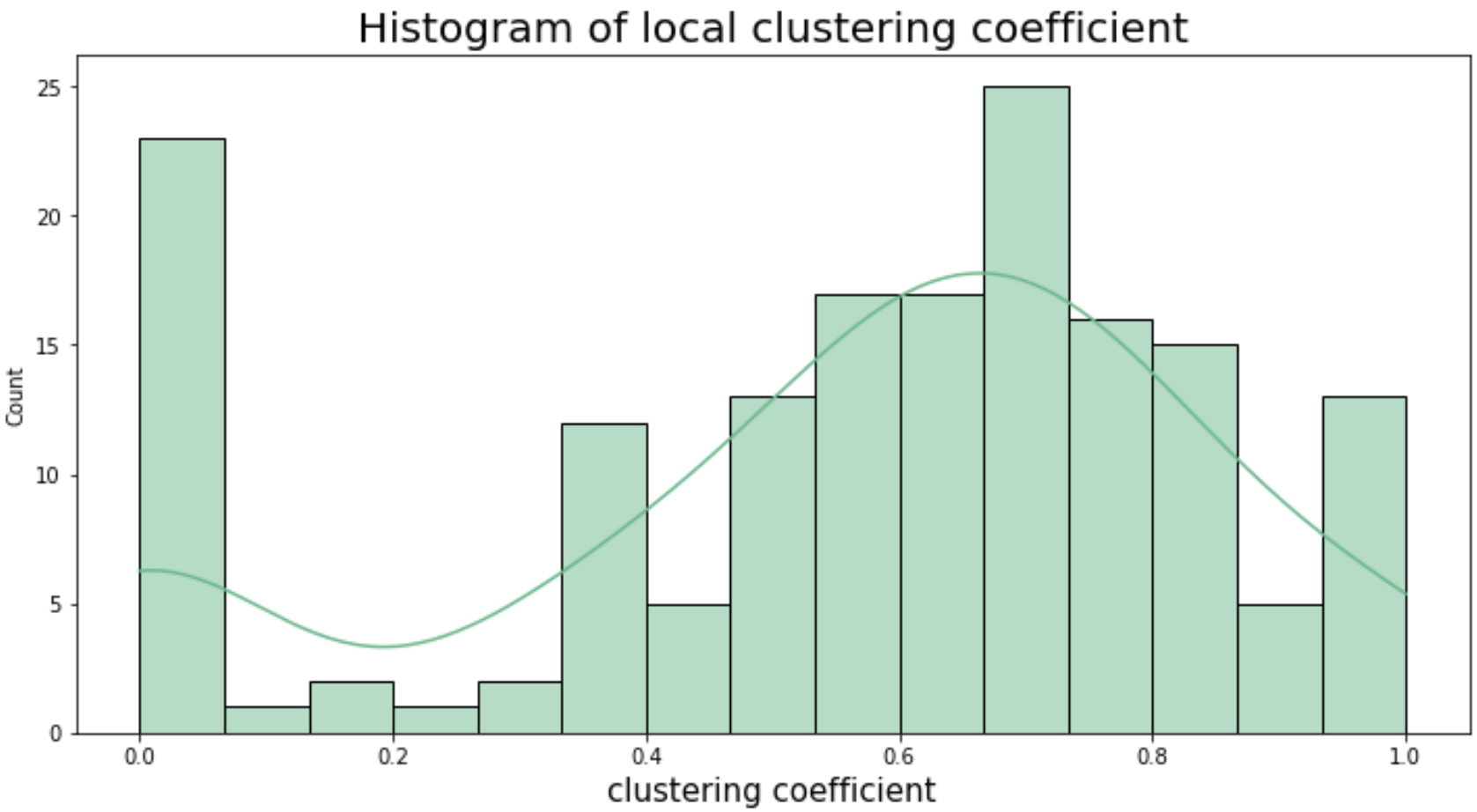
Social network analysis

Irina Govorova
мНОД-2020

Network summary

- Number of nodes: 167
- Number of edges: 1100
- Number of connected components: 10
- Radius: 5
- Diameter: 9
- Average clustering coefficient: 0.56
- Average path length (largest cc): 3.49

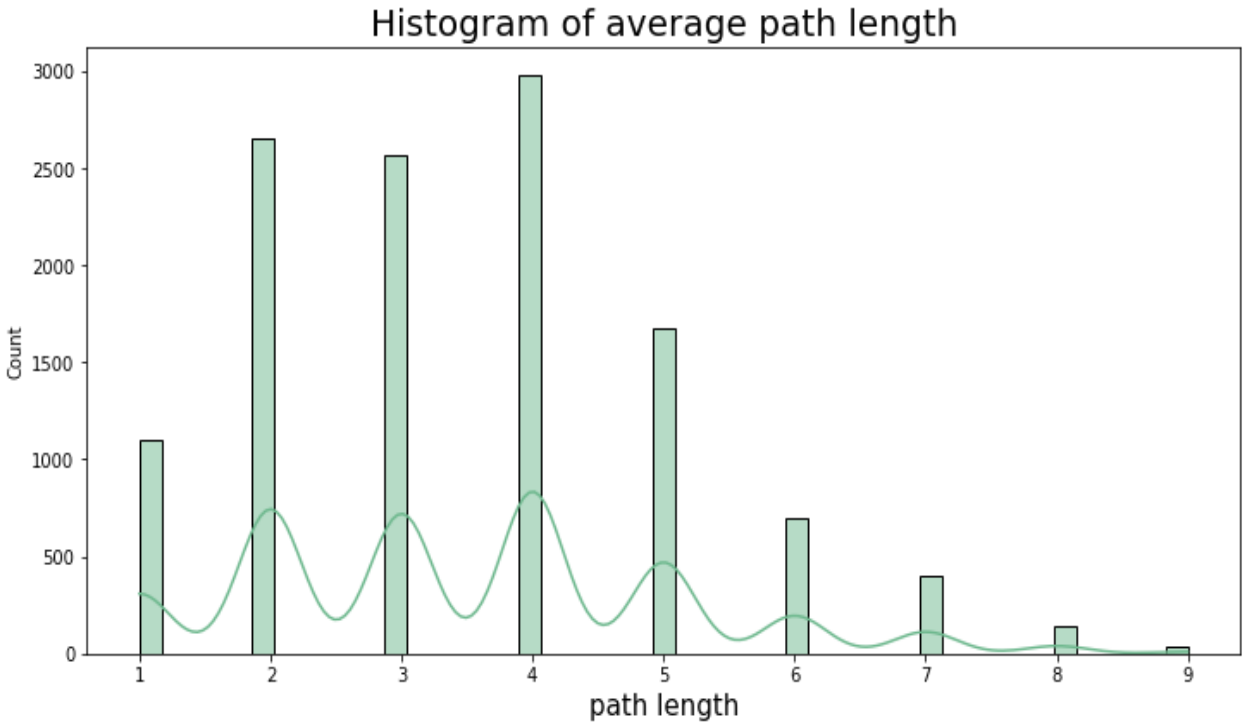
Histograms

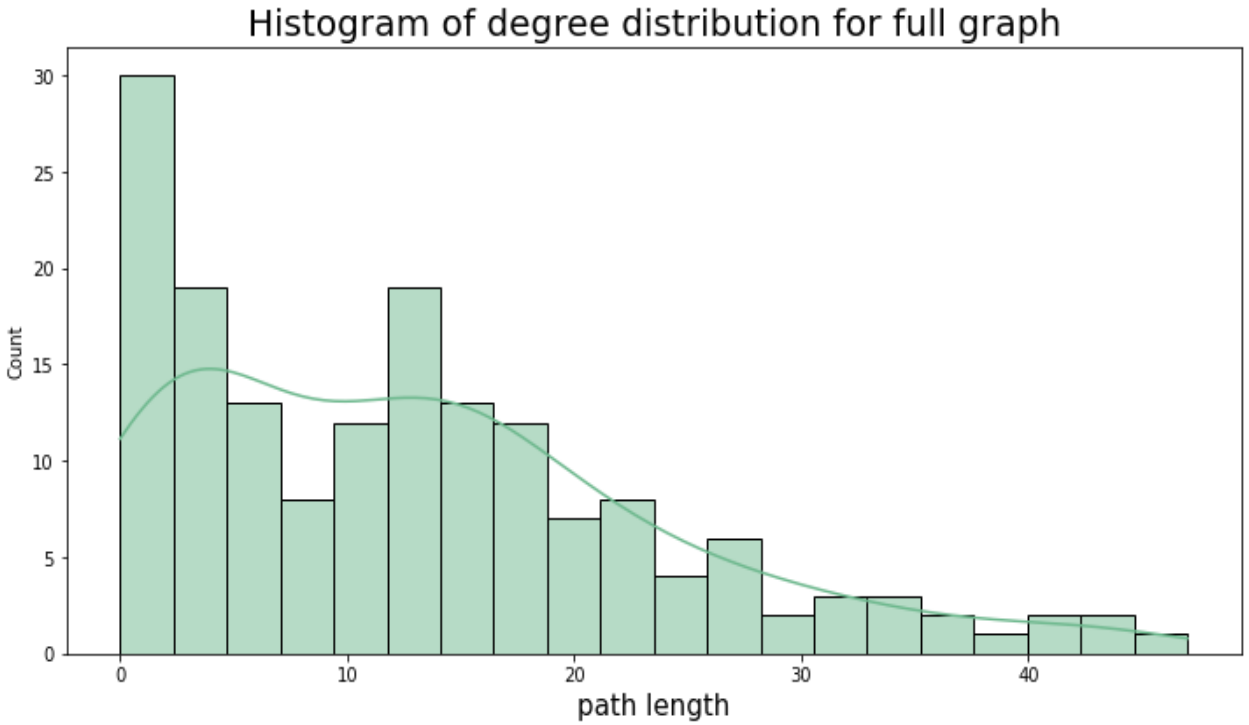


Majority of nodes have high clustering coefficients, so they to tends to cluster.

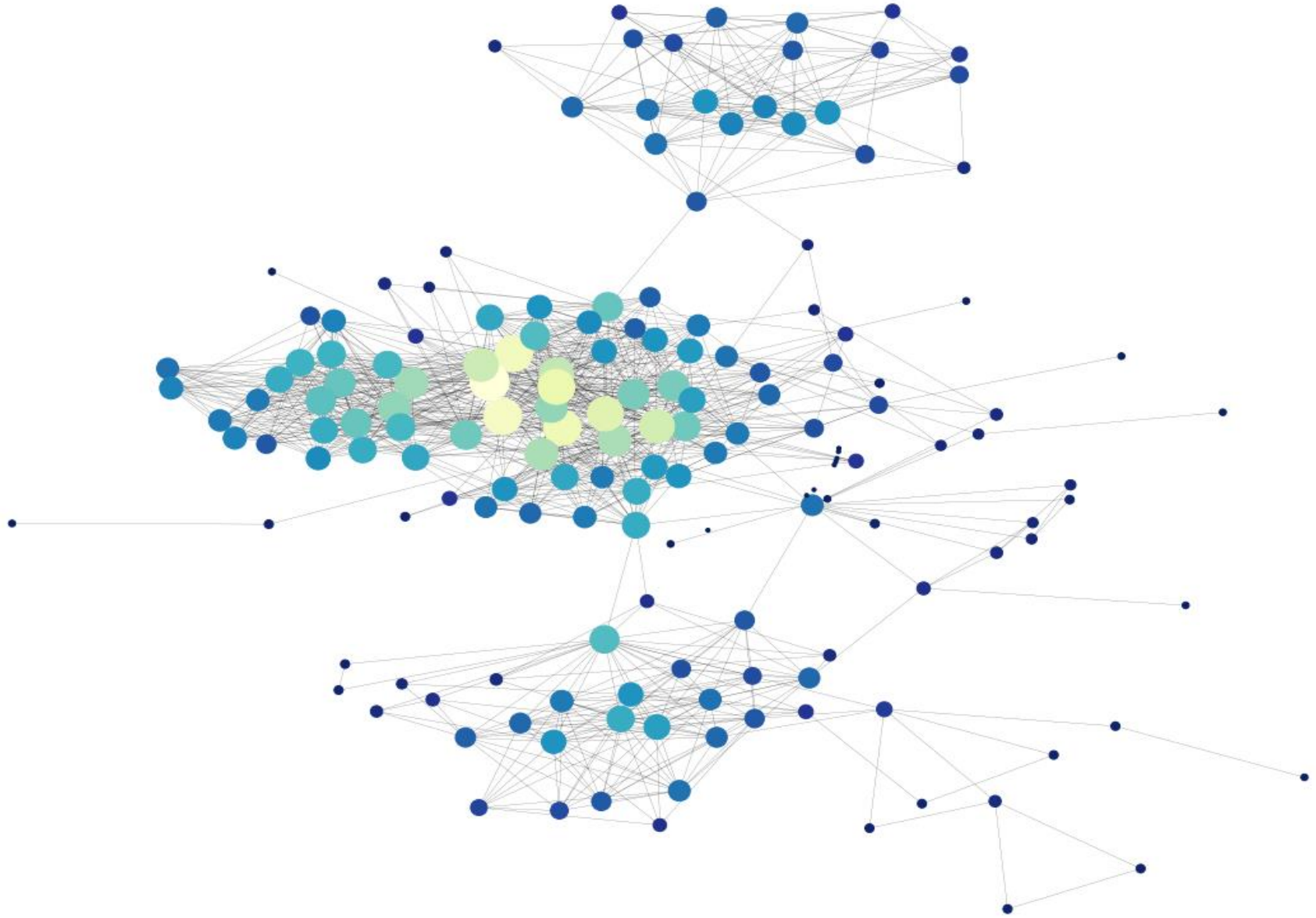
But there are many nodes with 0 clustering coefficient and that reduce average value to 0.56

Histograms


Histogram of average path length
Count
path length
0
500
1000
1500
2000
2500
3000
1
2
3
4
5
6
7
8
9


Histogram of degree distribution for full graph
Count
path length
0
5
10
15
20
25
30
0
10
20
30
40

Network layout



Random network models

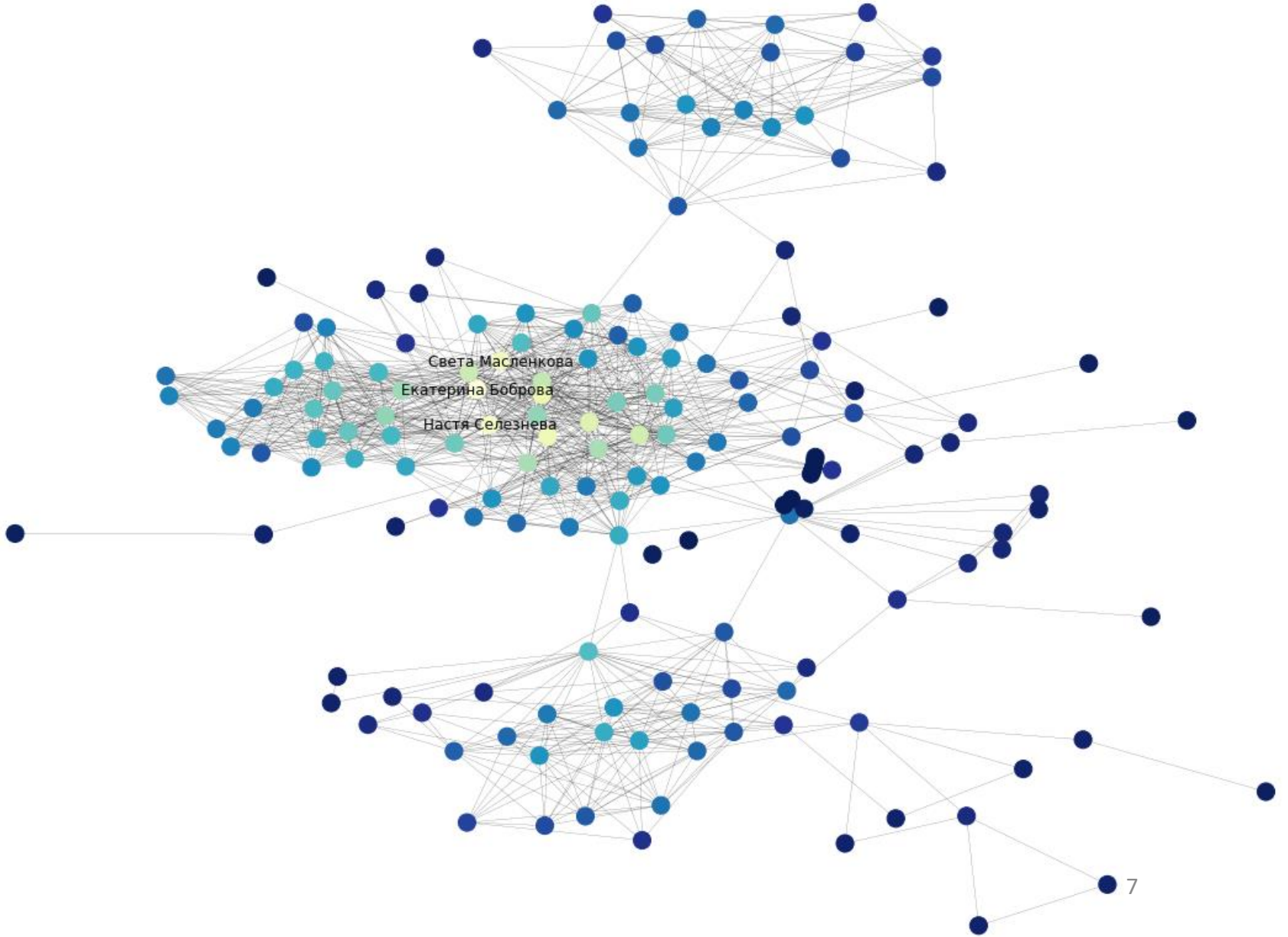
	Number of nodes	Number og edges	Clustering coeff	Path length	Diameter
My network	157	1099	0.594548	3.494692	9
Erdos-Renyi model	157	12246	1.000000	1.000000	1
Barabasi-Albert model	157	1050	0.155901	2.184305	3
Watts-Strogatz model	157	1099	0.148927	2.216724	3

The closet random model is Watts-Strogatz model. It has the same number of edges, closest average path length and diameter.

Degree centrality

Top-10 nodes

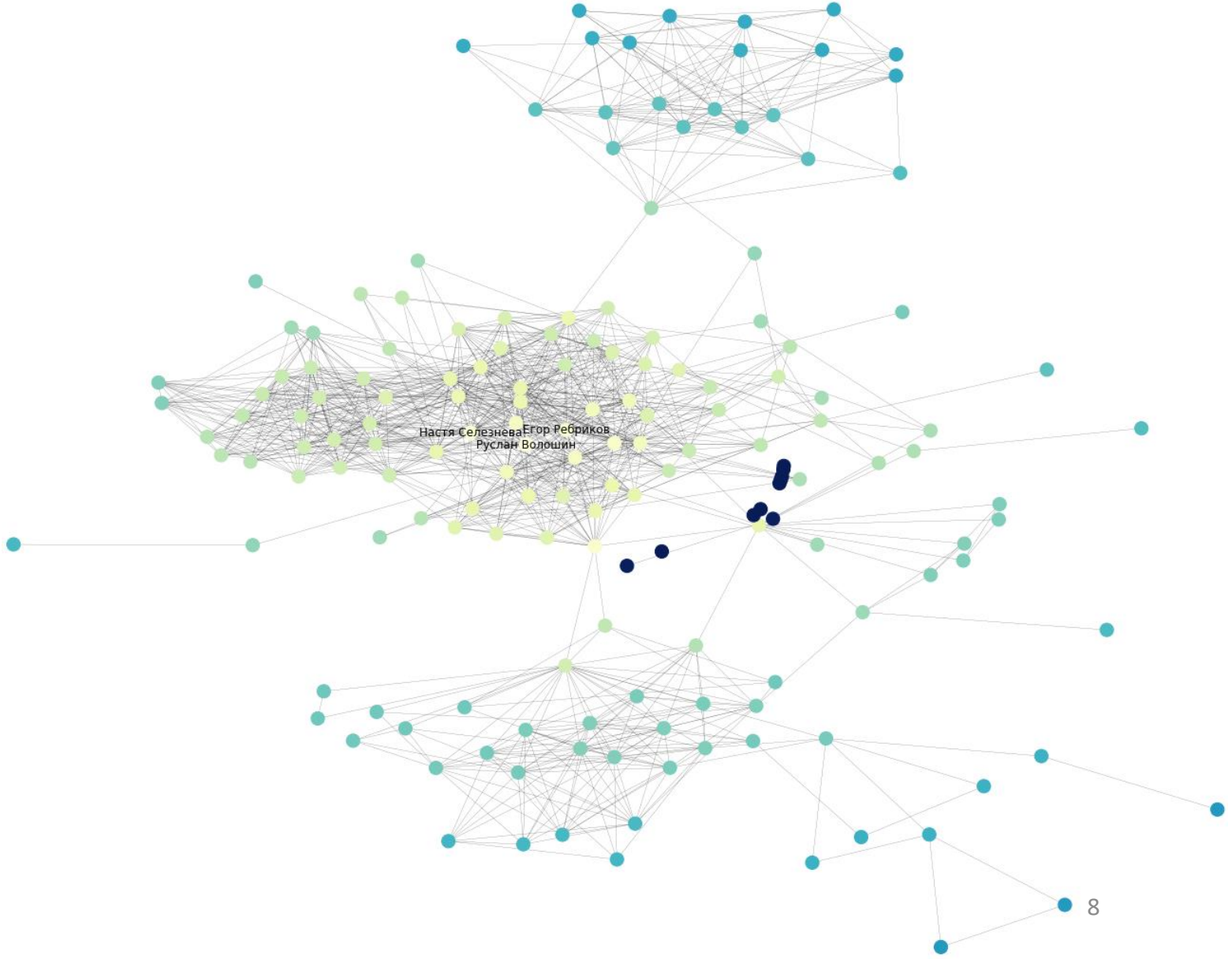
	id	Name	Degree centrality
0	152573209	Екатерина Боброва	0.283133
1	124330617	Настя Селезнева	0.26506
2	117958697	Света Масленкова	0.259036
3	156099664	Руслан Волошин	0.253012
4	16546514	Гриша Моргунов	0.246988
5	197625757	Егор Ребриков	0.23494
6	66917770	Полина Татарская	0.222892
7	161085240	Вячеслав Саушкин	0.216867
8	143103666	Настя Бердникова	0.210843
9	23506333	Елизавета Ильичёва	0.198795



Closeness centrality

Top-10 nodes

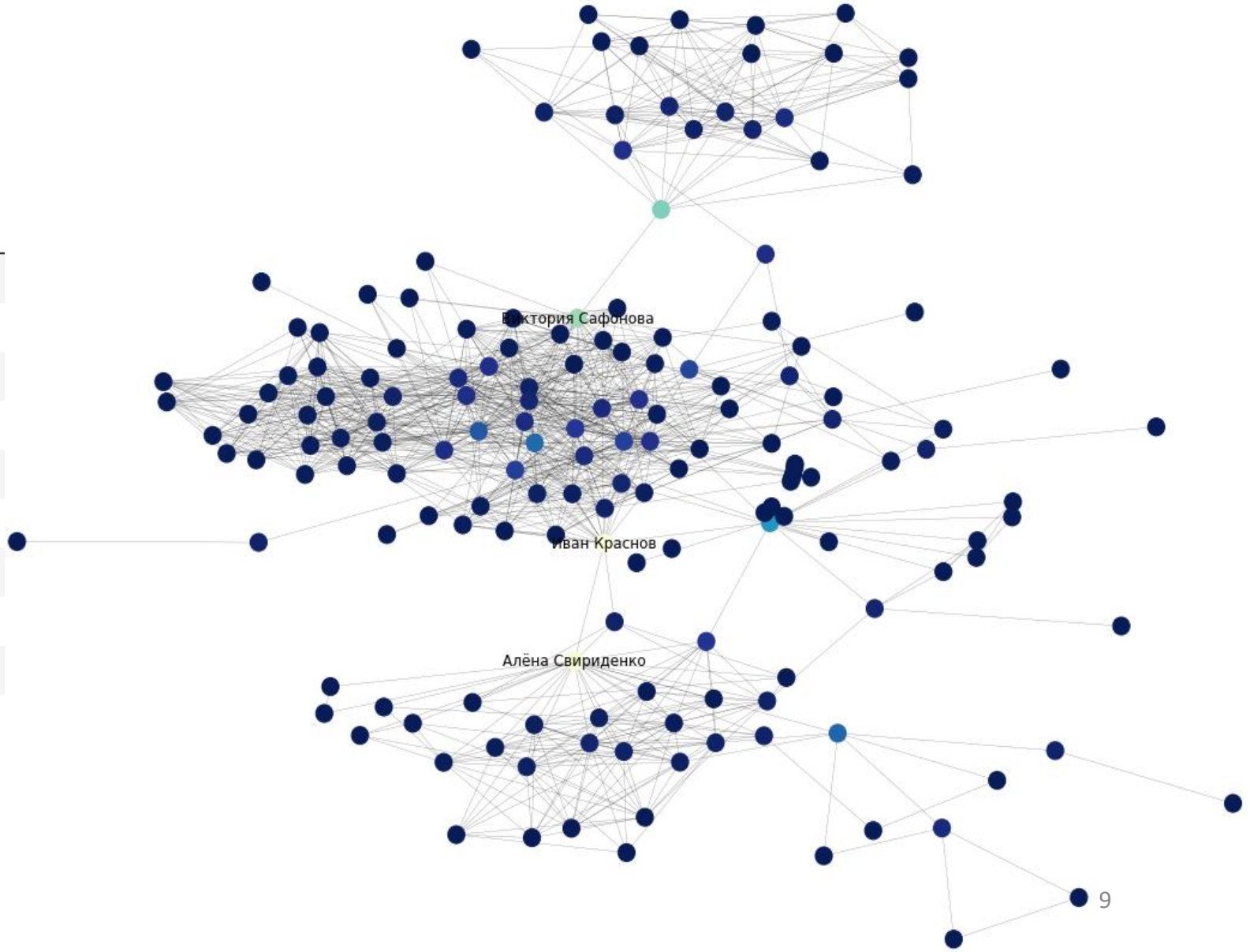
	id	Name	Closeness centrality
0	156099664	Руслан Волошин	0.408363
1	124330617	Настя Селезнева	0.399462
2	197625757	Егор Ребриков	0.398376
3	86644662	Иван Краснов	0.393036
4	66917770	Полина Татарская	0.391985
5	23506333	Елизавета Ильичёва	0.383776
6	217299615	Вика Троянова	0.382774
7	138303505	Елизавета Лошак	0.377841
8	33961841	Екатерина Каплун	0.37687
9	55447526	Роман Паршин	0.373034



Betweenness centrality

Top-10 nodes

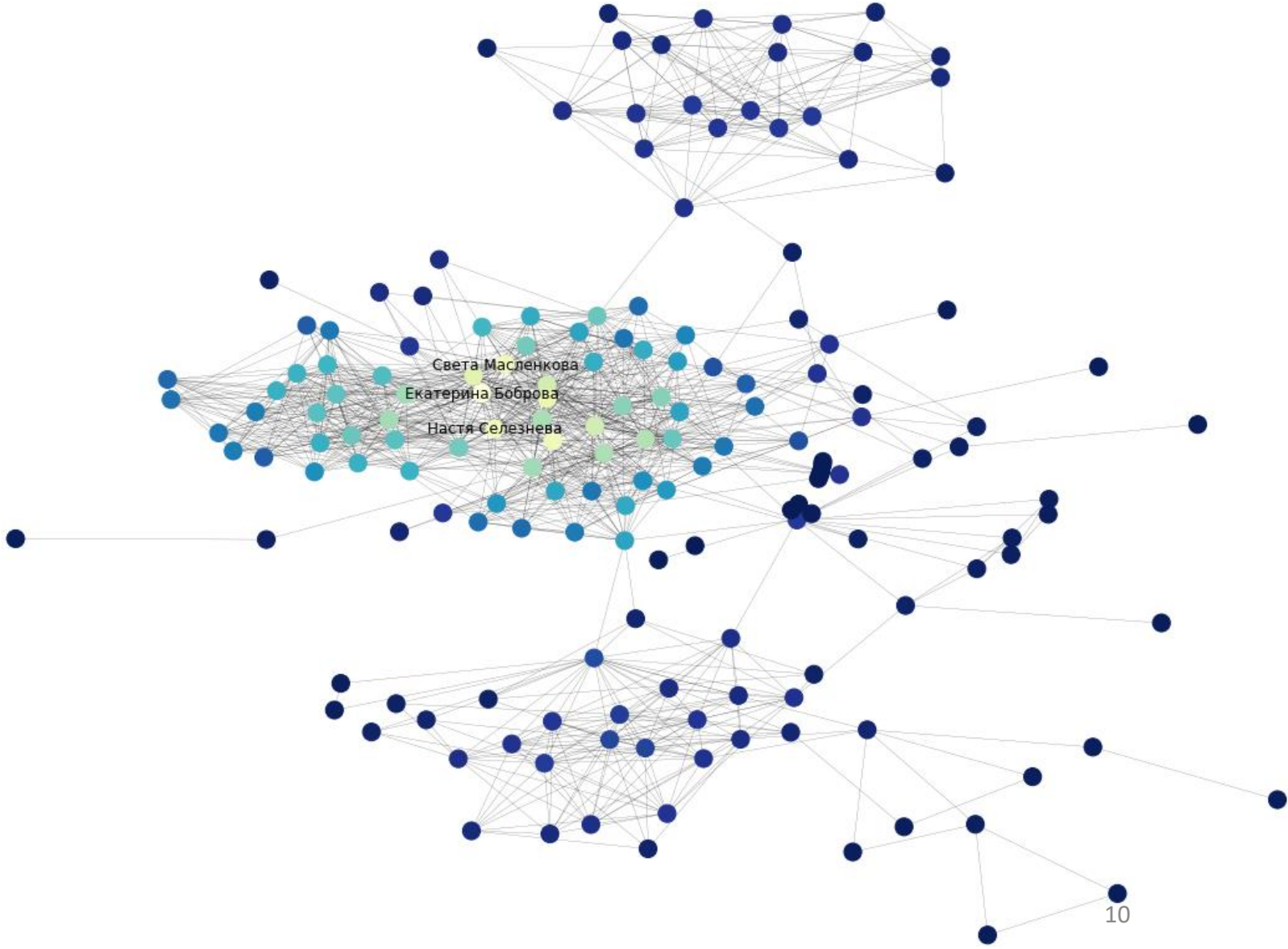
	id	Name	Betweenness centrality
0	86644662	Иван Краснов	0.284701
1	46978768	Алёна Свириденко	0.274095
2	148245859	Виктория Сафонова	0.193138
3	8469731	Александр Хижик	0.17854
4	182720390	Екатерина Михайлова	0.112447
5	156099664	Руслан Волошин	0.0791772
6	135946777	Саша Предущенко	0.0769845
7	124330617	Настя Селезнева	0.0662605
8	116568163	Диана Фелина	0.0502946
9	66917770	Полина Татарская	0.0473895



Katz centrality

Top-10 nodes

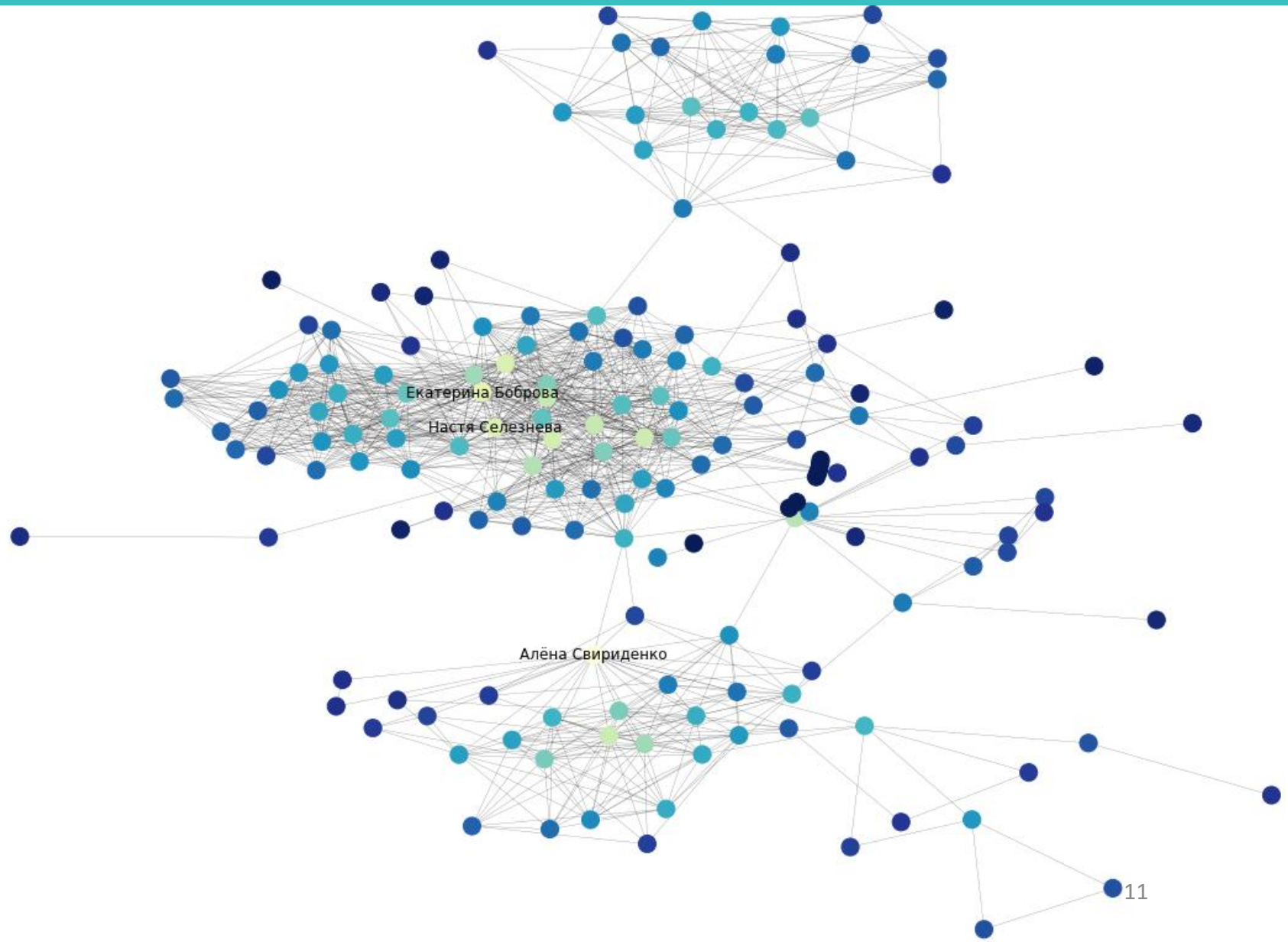
	id	Name	Katz centrality
0	152573209	Екатерина Боброва	0.190874
1	124330617	Настя Селезнева	0.176366
2	117958697	Света Масленкова	0.175823
3	156099664	Руслан Волошин	0.175452
4	16546514	Гриша Моргунов	0.171661
5	161085240	Вячеслав Саушкин	0.164656
6	143103666	Настя Бердникова	0.156382
7	197625757	Егор Ребриков	0.156008
8	66917770	Полина Татарская	0.14465
9	23506333	Елизавета Ильичёва	0.142949



Page-Rank

Top-10 nodes

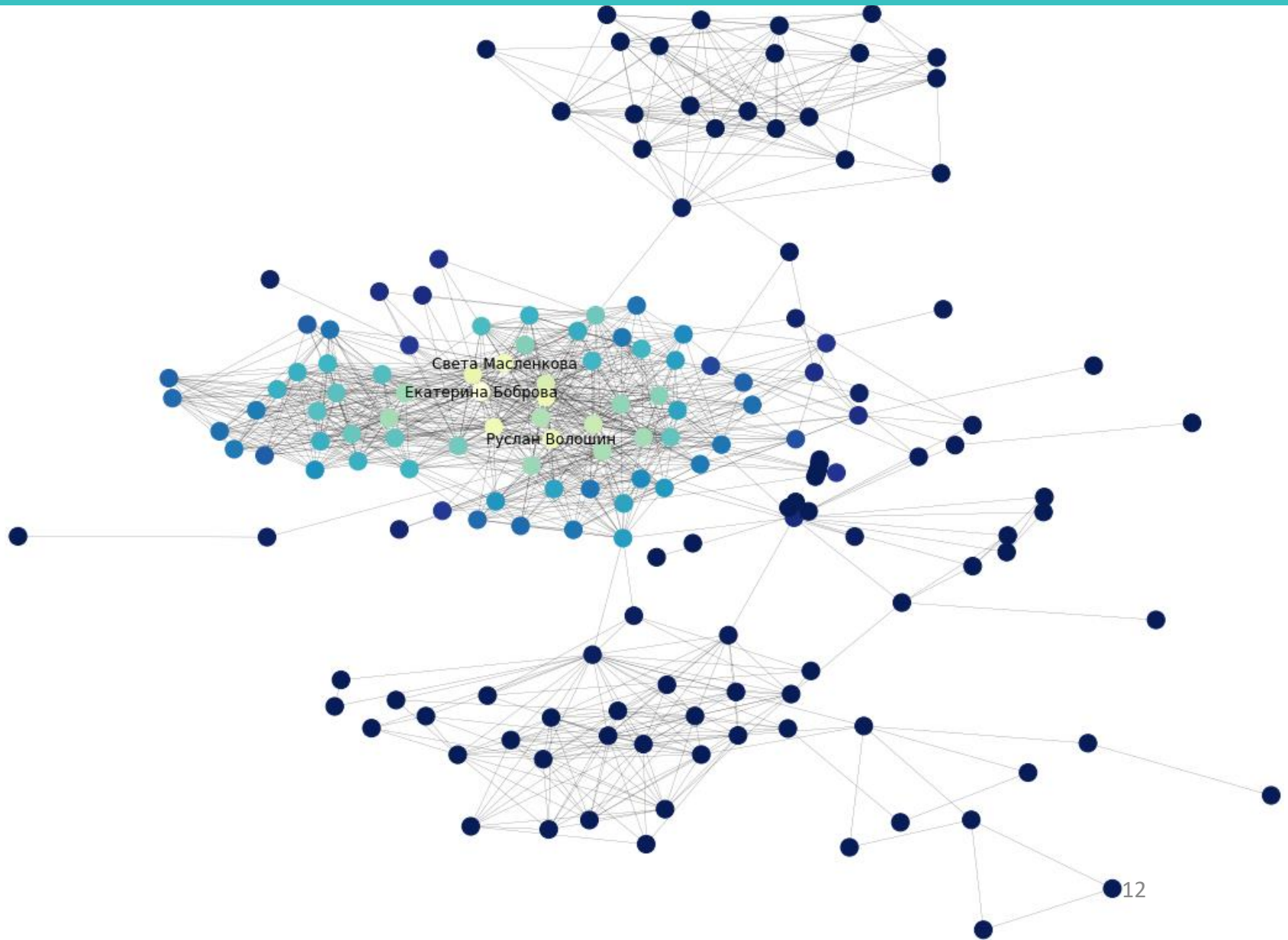
	id	Name	Page-rank
0	46978768	Алёна Свириденко	0.0166567
1	152573209	Екатерина Боброва	0.0145438
2	124330617	Настя Селезнева	0.013823
3	117958697	Света Масленкова	0.0137172
4	156099664	Руслан Волошин	0.0134243
5	66917770	Полина Татарская	0.0130736
6	61401572	Юлия Кретова	0.0129762
7	197625757	Егор Ребриков	0.012777
8	16546514	Гриша Моргунов	0.0127583
9	182720390	Екатерина Михайлова	0.0123849



HITS

Top-10 nodes

	id	Name	HITS
0	152573209	Екатерина Боброва	0.0292558
1	117958697	Света Масленкова	0.0265634
2	156099664	Руслан Волошин	0.0264799
3	124330617	Настя Селезнева	0.0263877
4	16546514	Гриша Моргунов	0.0259136
5	161085240	Вячеслав Саушкин	0.0252126
6	143103666	Настя Бердникова	0.023505
7	197625757	Егор Ребриков	0.0224994
8	217299615	Вика Троянова	0.0208381
9	23506333	Елизавета Ильичёва	0.0205132



Correlation comparison

Pearson correlation coefficient

Page-Rank

Degree centrality: 0.87
Closeness centrality: 0.53
Betweenness centrality: 0.44
Katz centrality: 0.7

Page-Rank is mostly correlated with degree centrality and they have quite similar top-10 nodes

HITS

Degree centrality: 0.88
Closeness centrality: 0.67
Betweenness centrality: 0.16
Katz centrality: 0.99

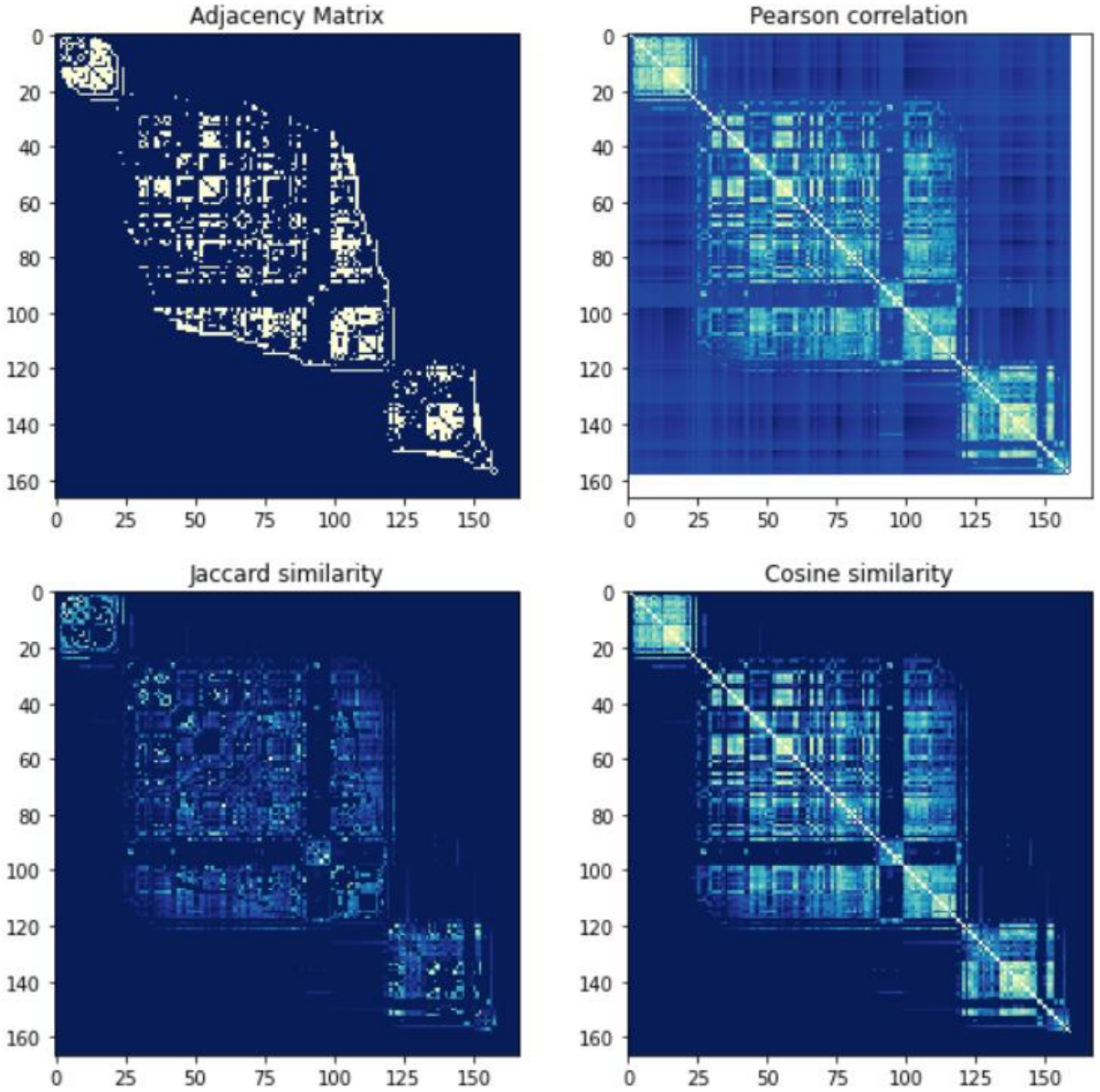
HITS is almost linearly dependent on Katz centrality and they have almost equivalent top-10 nodes.

Assortativity

	sex	city	education
assortativity coefficient	0.09198	0.216578	0.189979

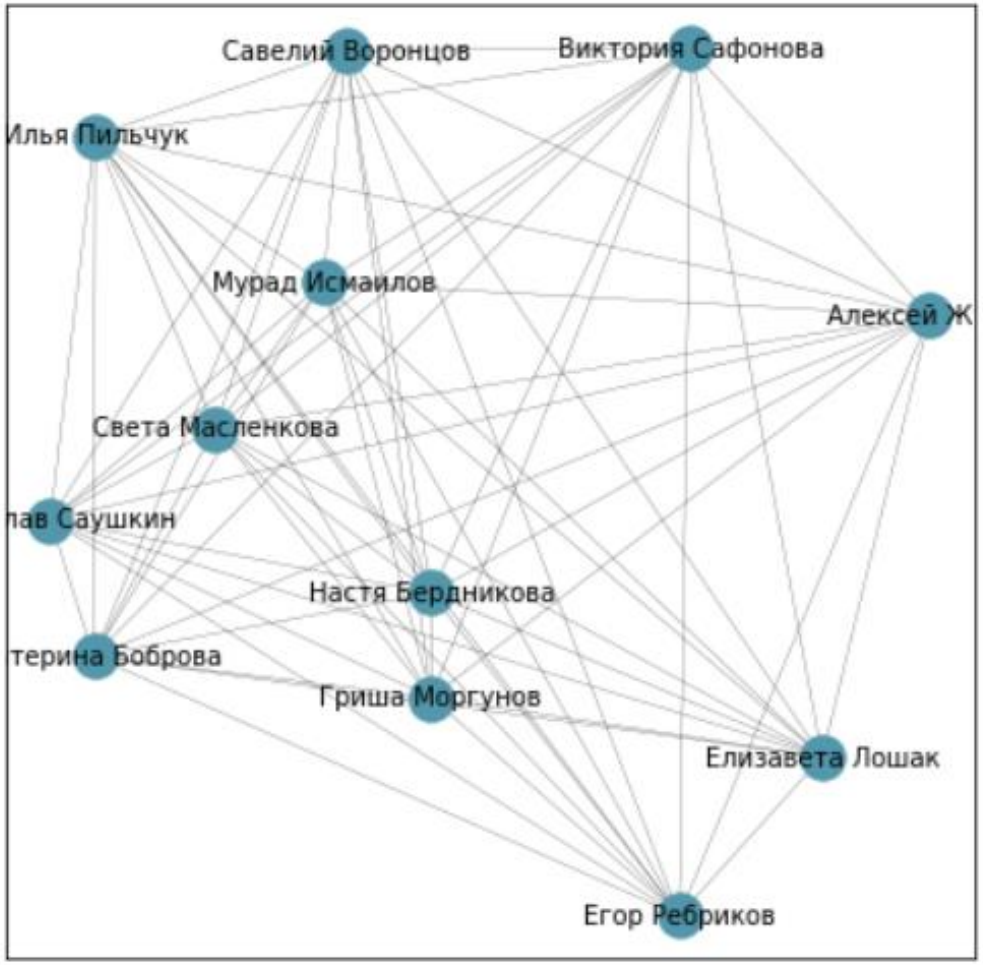
My graph is non-assortative by sex attribute but tends to be assortative by city and education attributes. Not everyone specify city and university in their VK profiles. So I think my graph could be more assortative by city and education if there was no missing attributes

Node structural similarity



My graph has a lot of structurally similar nodes. That is why my friends are well separated into communities, which do not intersect and majority of friends from one community know each other

Clique search

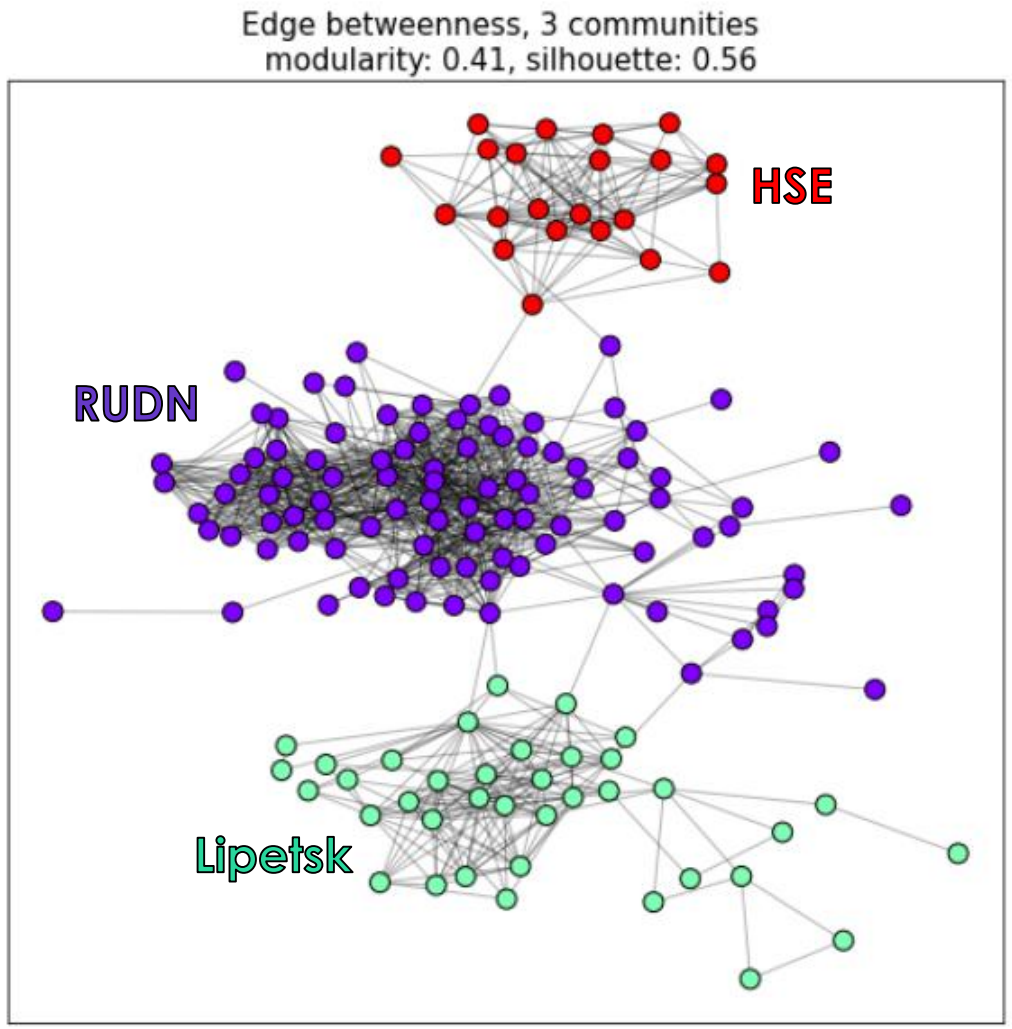


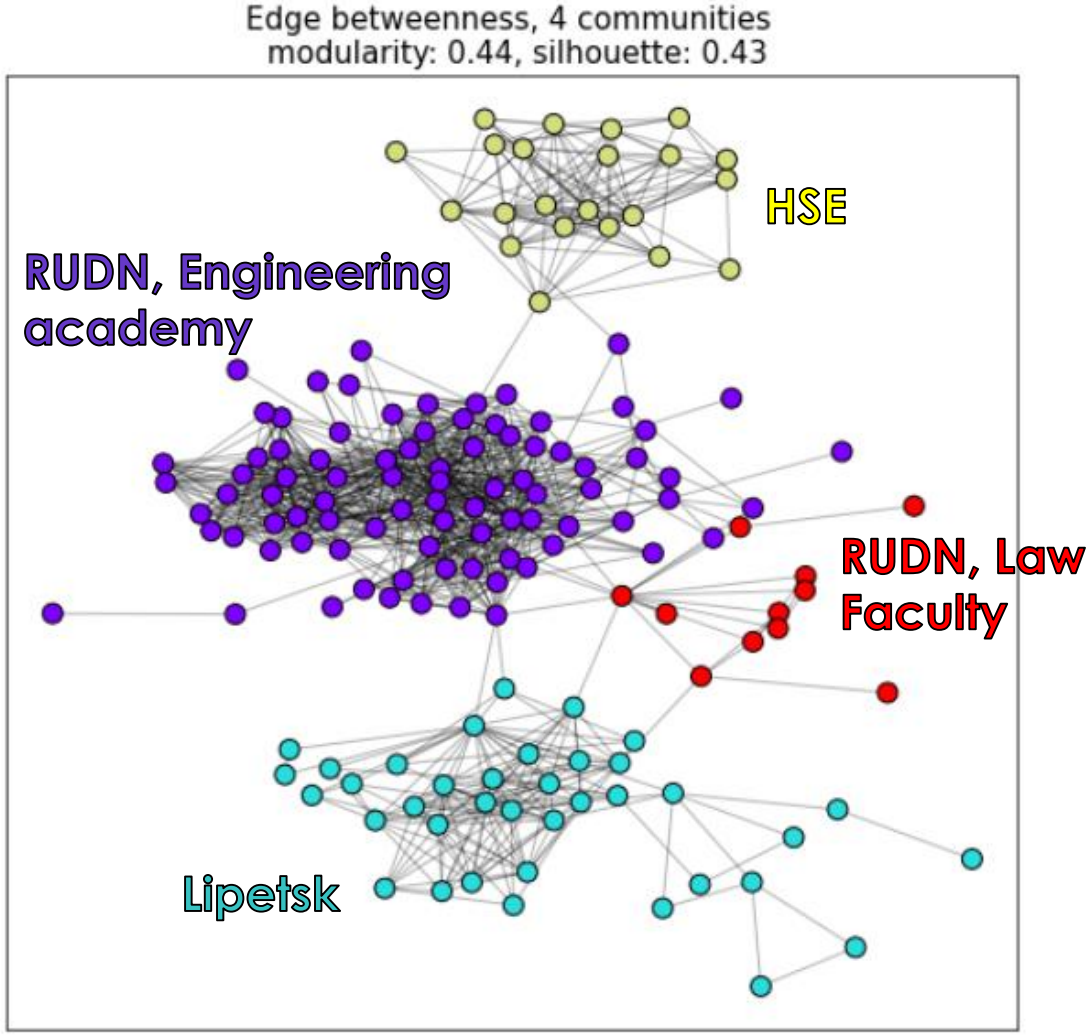
Number of cliques: 509

Number of cliques of maximal size: 10
Maximal size: 12

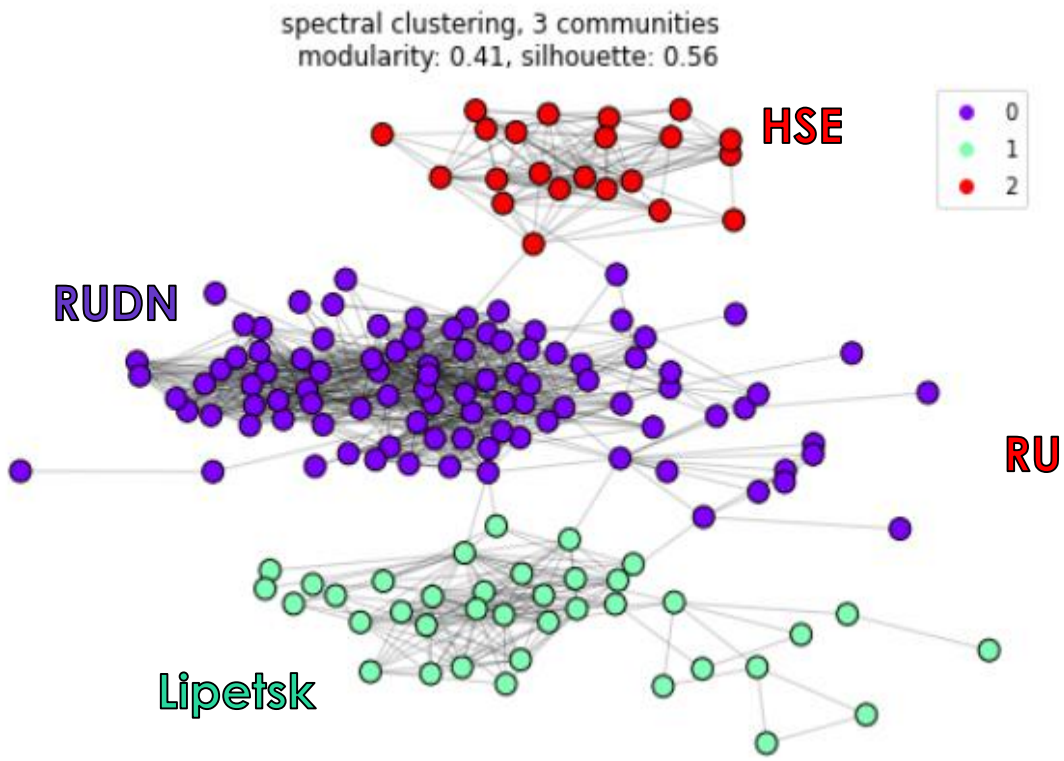
All cliques of maximal size are networks of my groupmates from RUDN

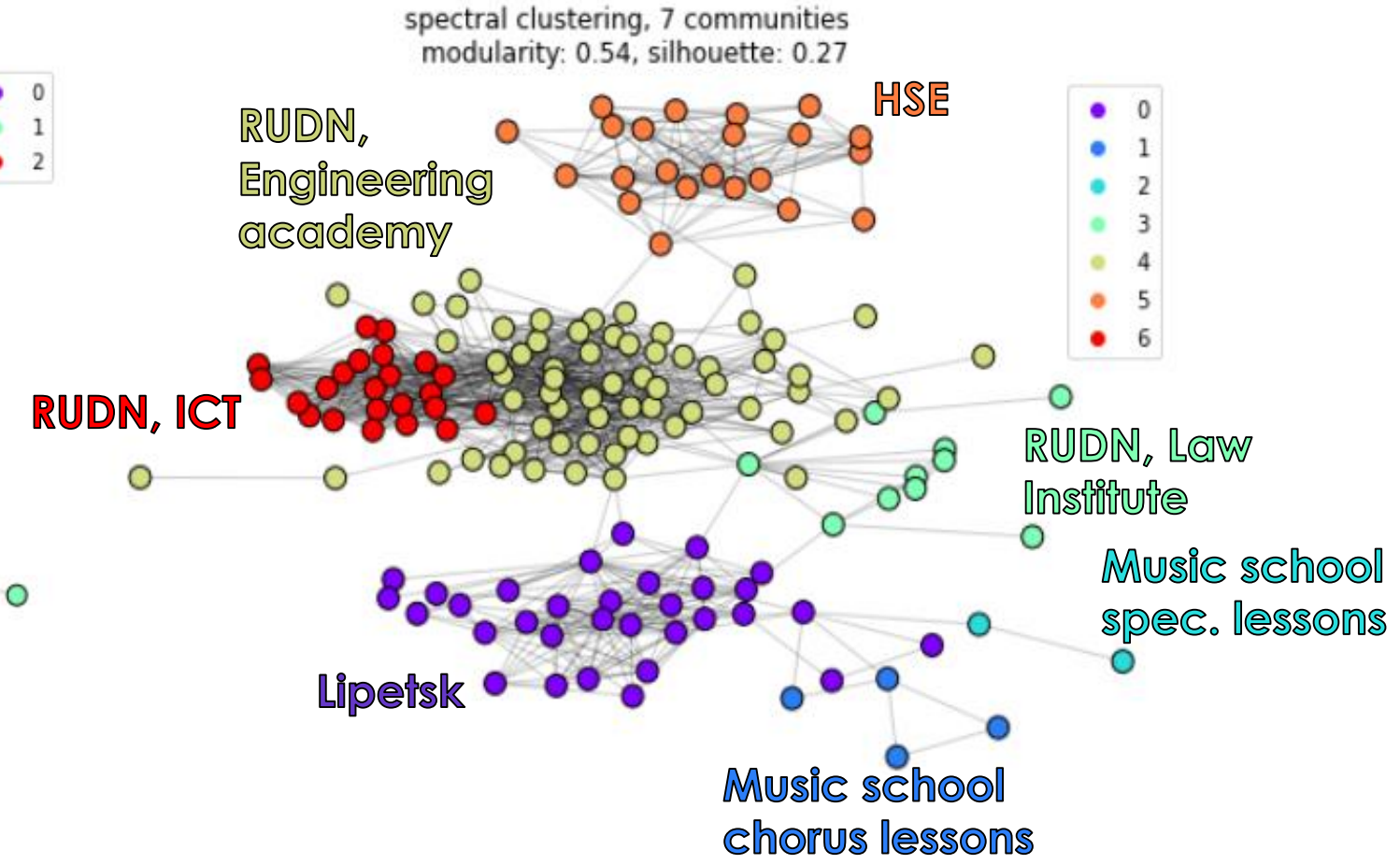
Communities



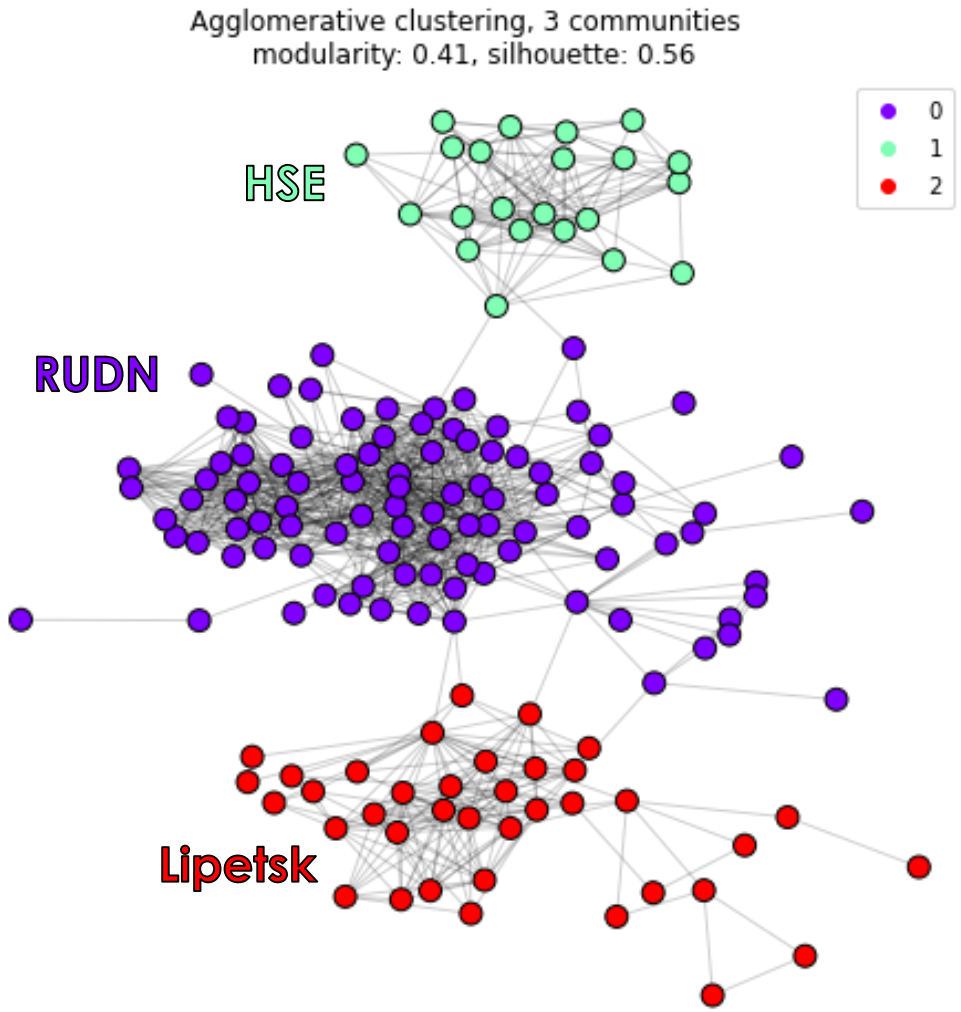


Communities

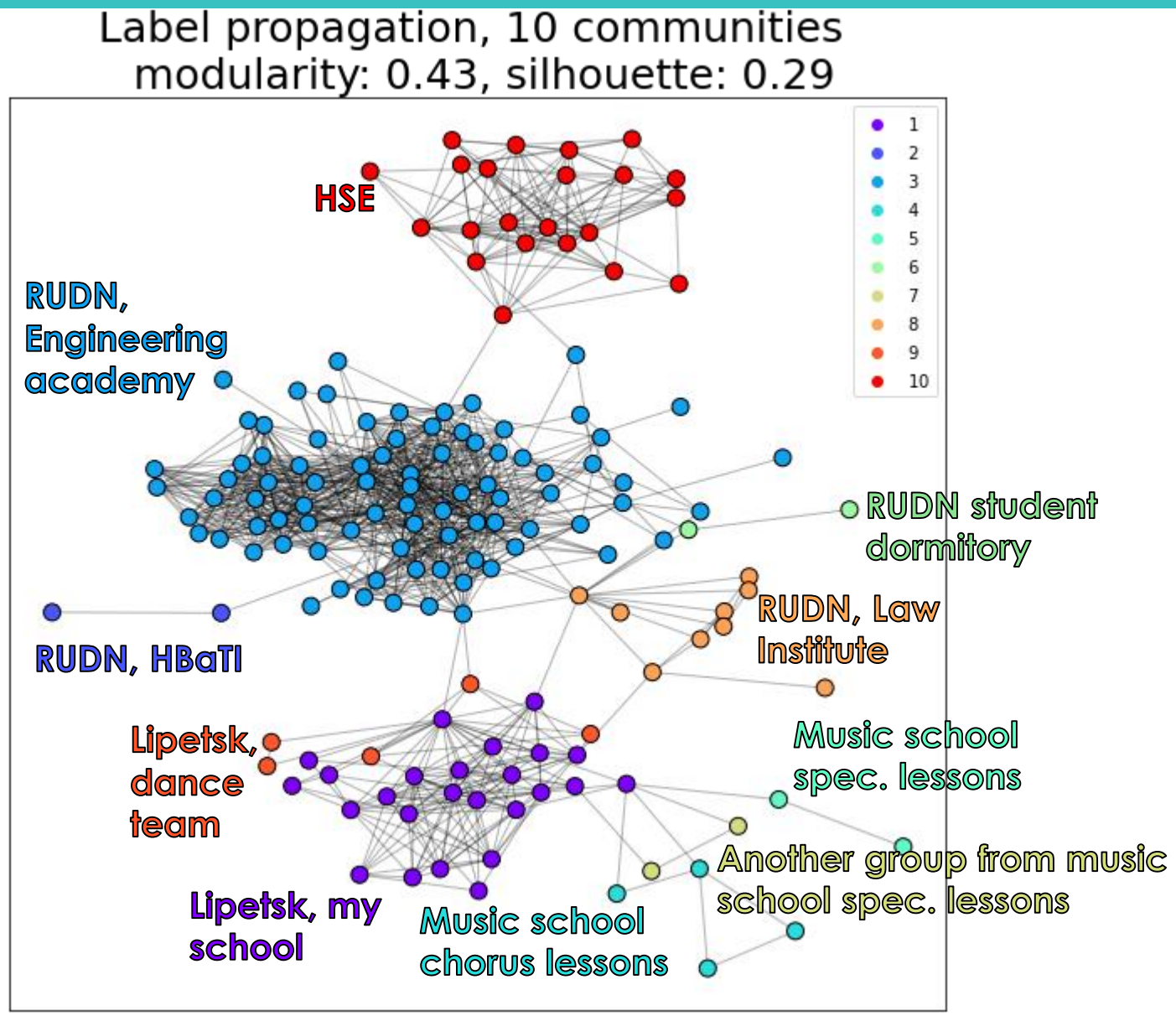




Communities



Communities



Thank you for attention!