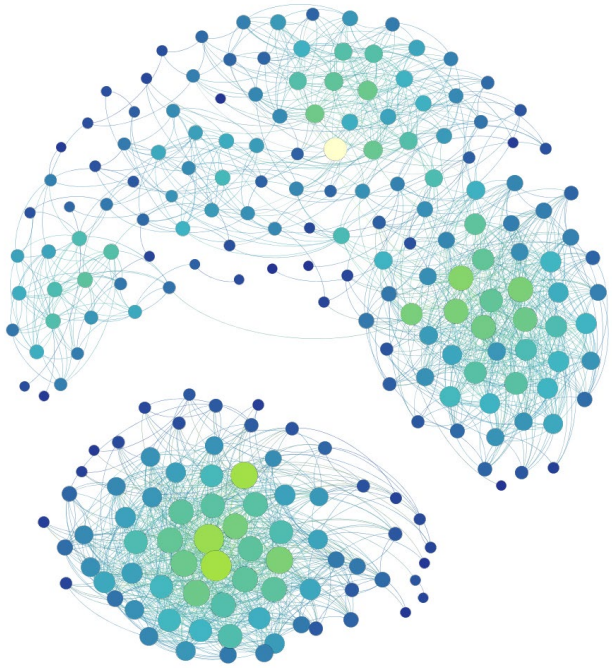
VK analysis
Network visualization

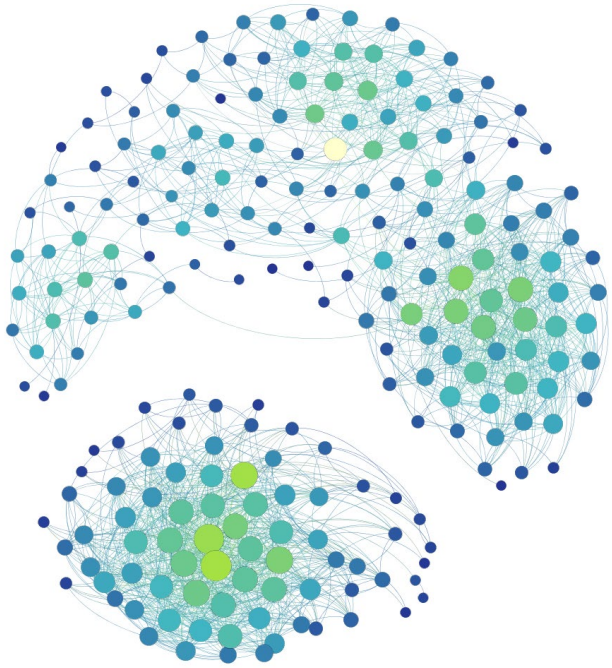
Konstantinov D.N.

M05-015a

Contents



Contents



Network Summary

Statistics:

- Nodes - 227 (30 deleted)
- Edges - 1960
- Average node degree - 17.27
- Average clustering coefficient – 0.55
- Components - 2

Largest:

- Nodes – 151
- Edges - 1093
- Diameter – 8
- Average path length - 3

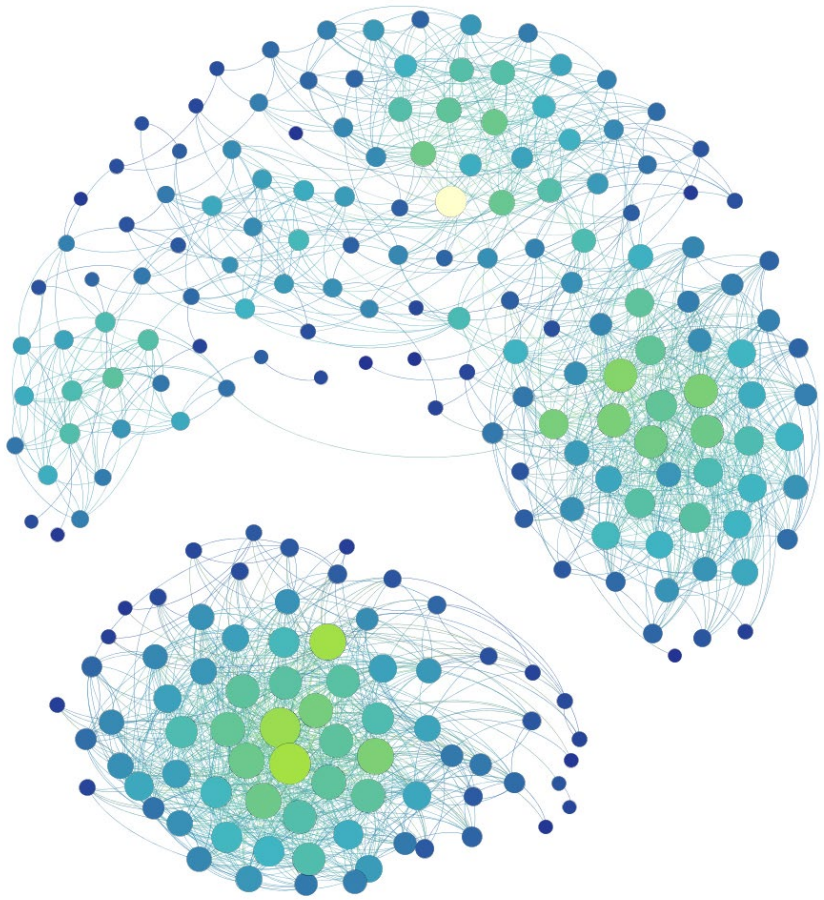
Smallest:

- Nodes – 76
- Edges - 867
- Diameter – 4
- Average path length – 1.8

Attributes of nodes:

- name
- sex
- city

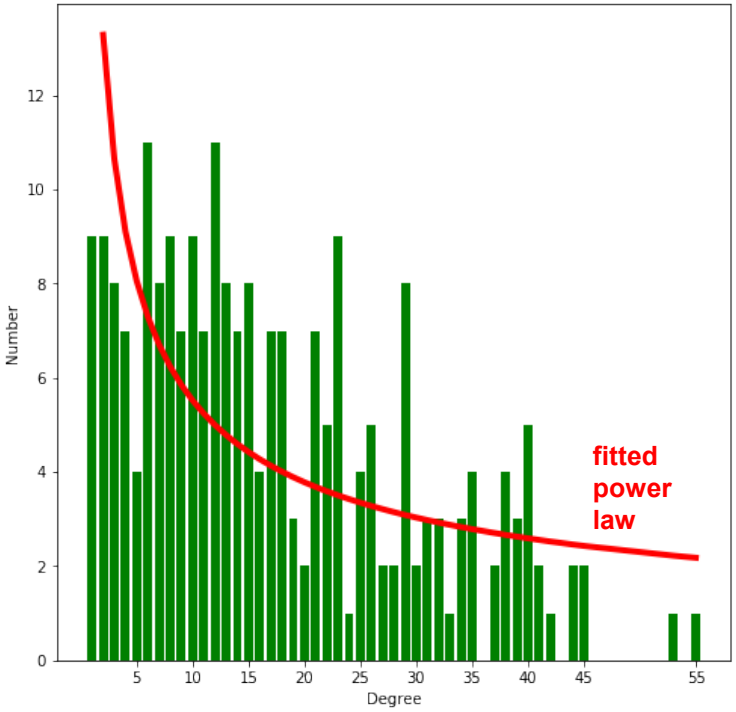
VK graph



vk.api
preprocessed;
sorted by degree

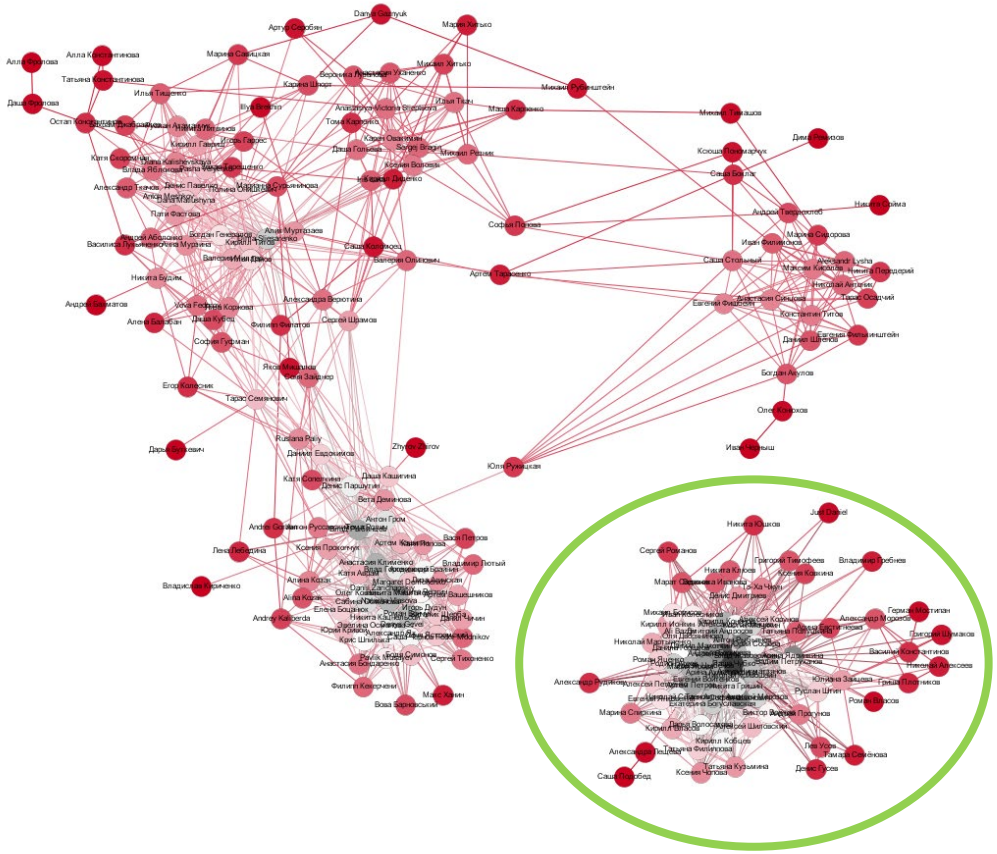
Network Summary

Degree distribution



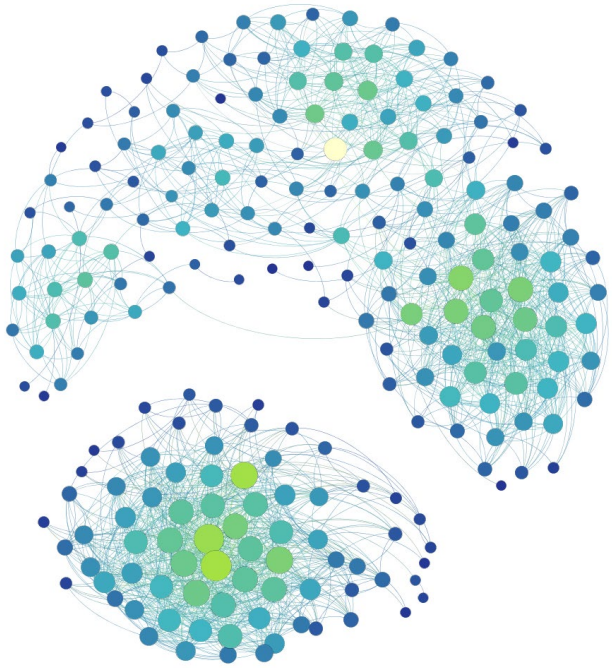
- min degree – 1
- max degree – 55
- not similar to power law

VK graph



MIPT DAFE

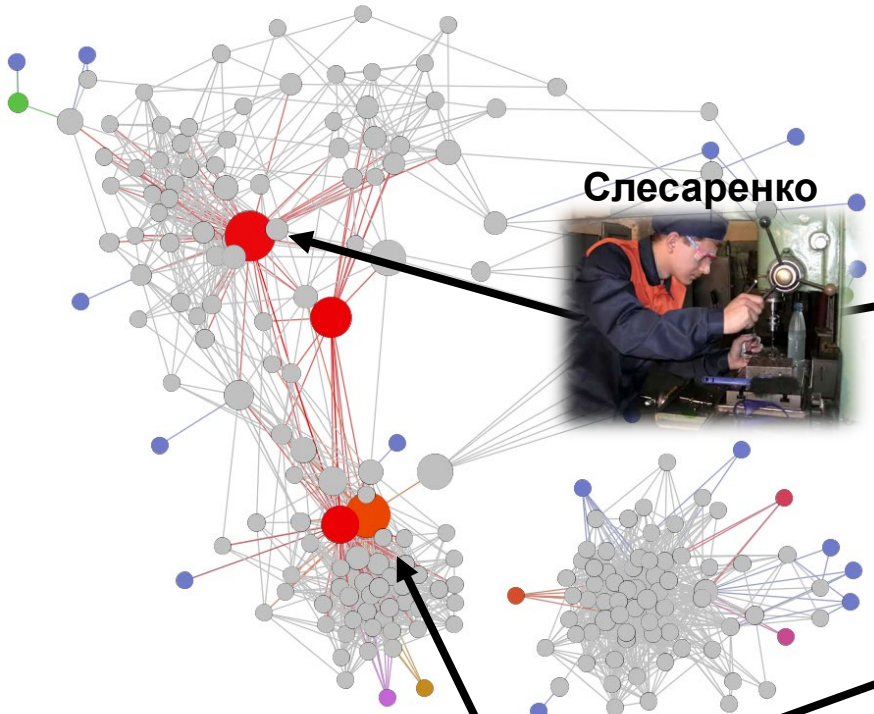
Contents



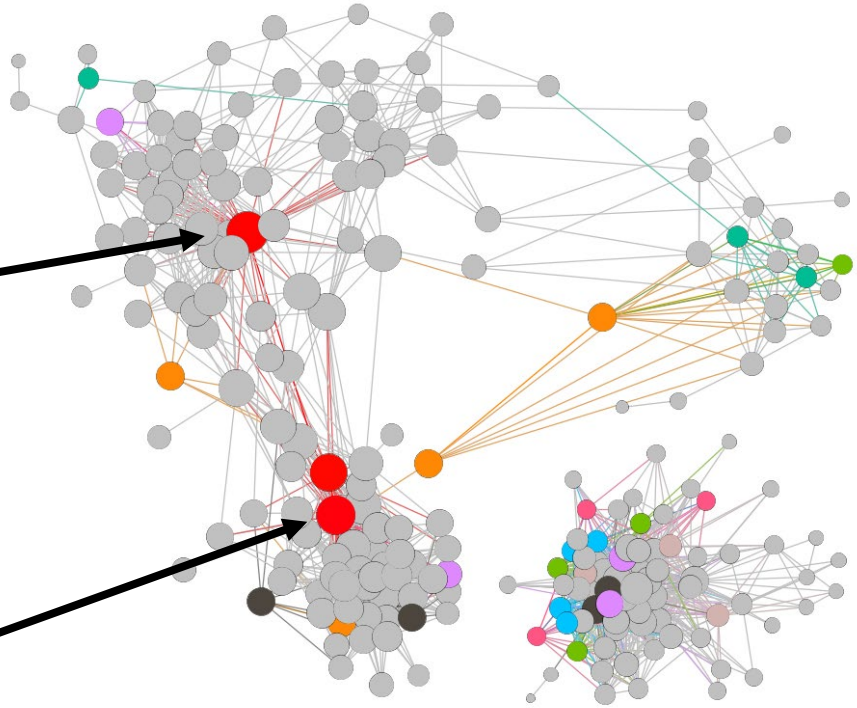
Degree/Closeness/Betweenness centralities

Betweenness

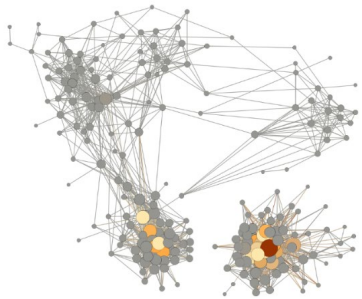
Closeness



Слесаренко



Рыба



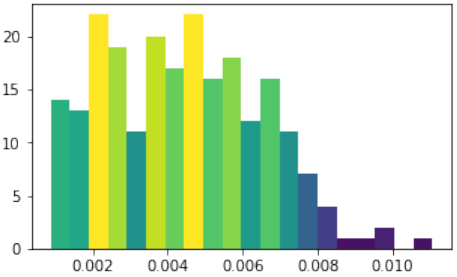
Degree

Structural Analysis

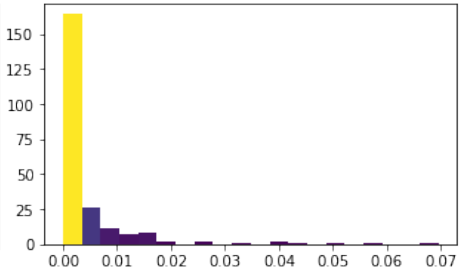
Top nodes interpretation

Degree centrality	Closeness centrality	Betweenness centrality	Pagerank
Паша Чубко	**Дима Слесаренко**	**Дима Слесаренко**	**Дима Слесаренко**
Ольга Борисова	**Влад Рыбинцев**	Антон Гром	Паша Чубко
Арина Ядринкина	Денис Паршутин	Сергей Шрамов	Арина Ядринкина
Антон Рыбьянов	Александра Верютина	**Влад Рыбинцев**	Ольга Борисова
Андрей Волков	Антон Гром	Валерия Олинович	**Влад Рыбинцев**

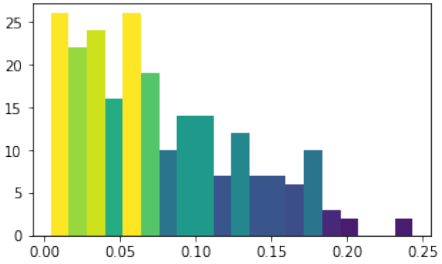
Pagerank



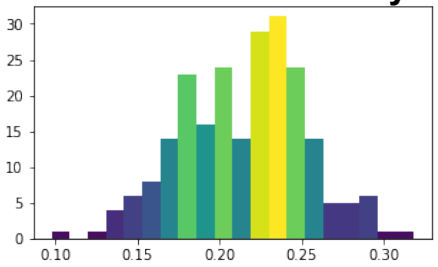
Betweenness centrality



Degree centrality

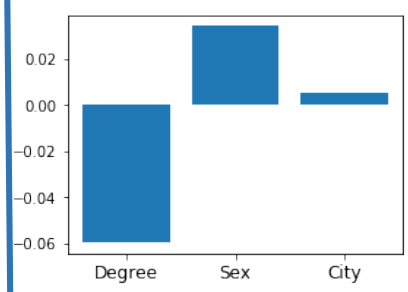


Closeness centrality



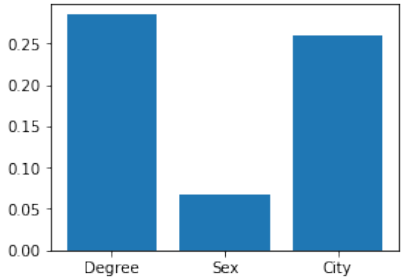
Assortative Mixing

MIPT



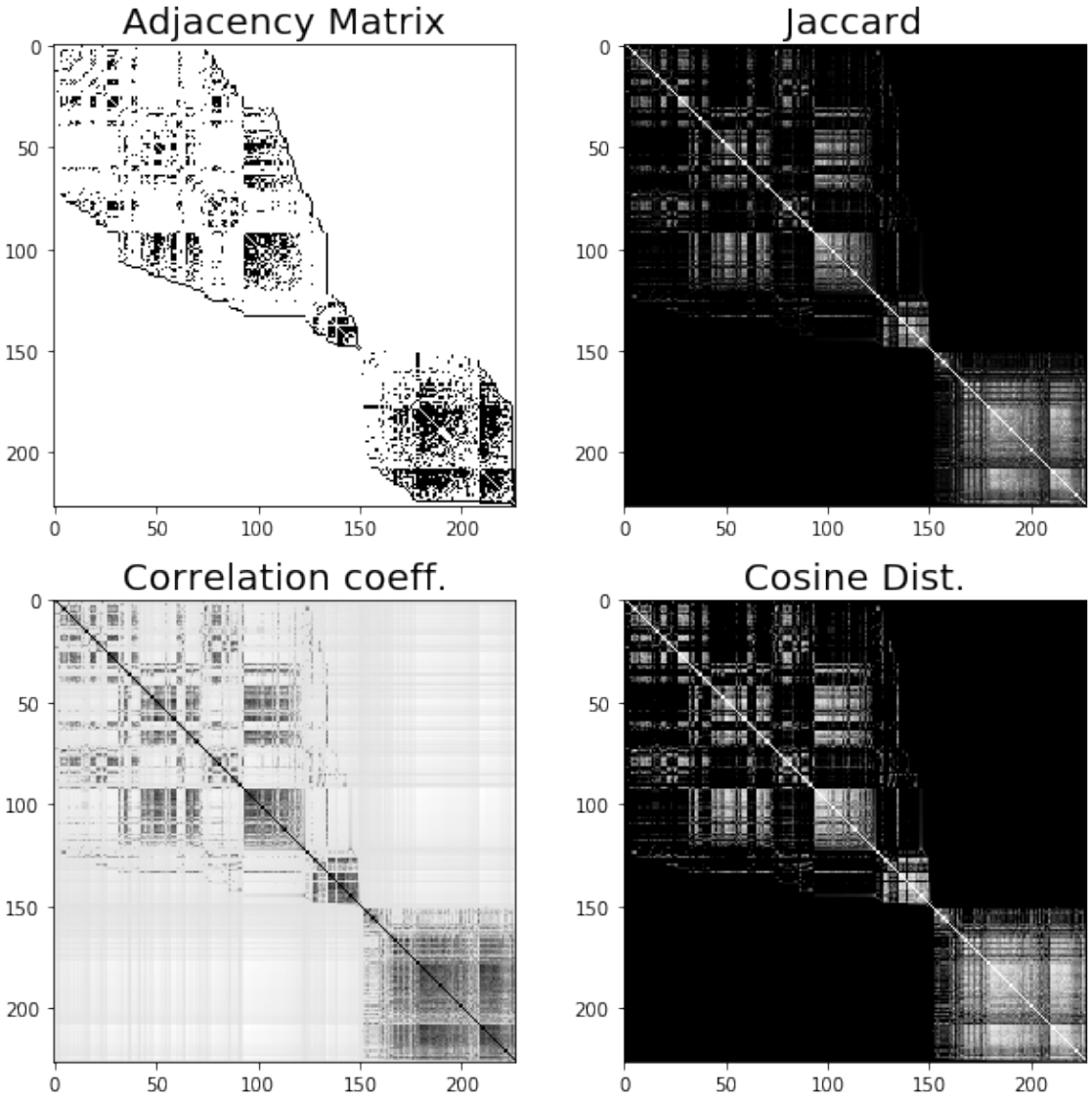
Disassortative by degree

Large component



Assortative by degree and city

Node structural similarity



- Used the reverse Cuthill-McKee heuristic
- 3 metrics reveal ~3 clusters

Closest random graph model similar to our network

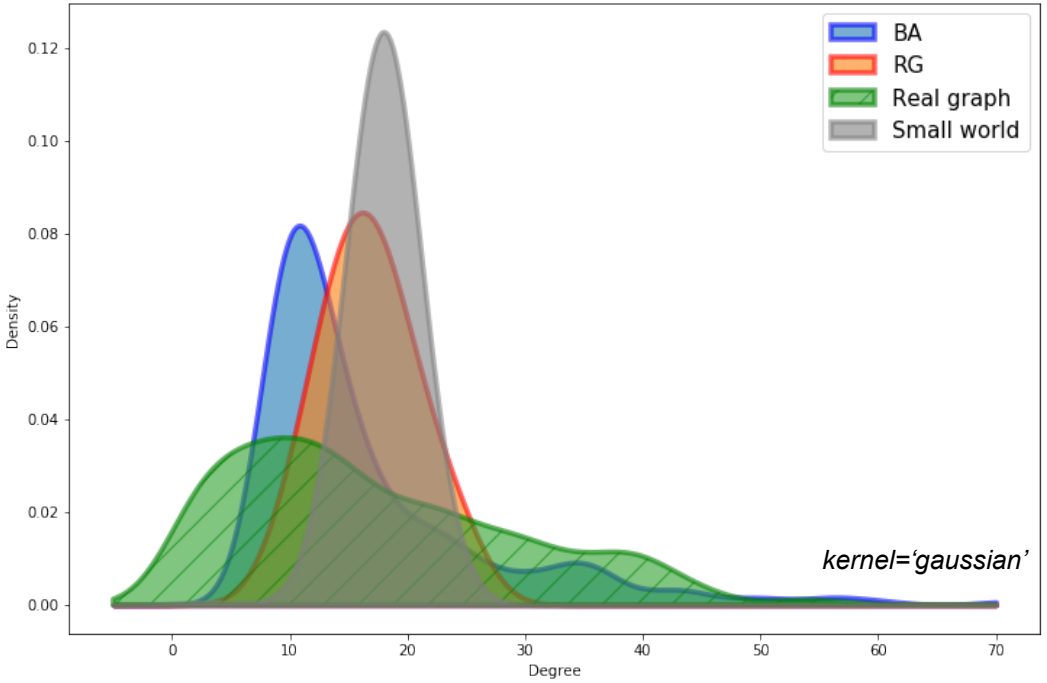
Assumptions:

- same number of nodes and edges
- $p \approx \langle k \rangle / n$ for random graph (RG)
- $m = \langle k \rangle / 2$ for Barabase-Albert (BA)
- probability parameter for small world preferential attachment is optimized

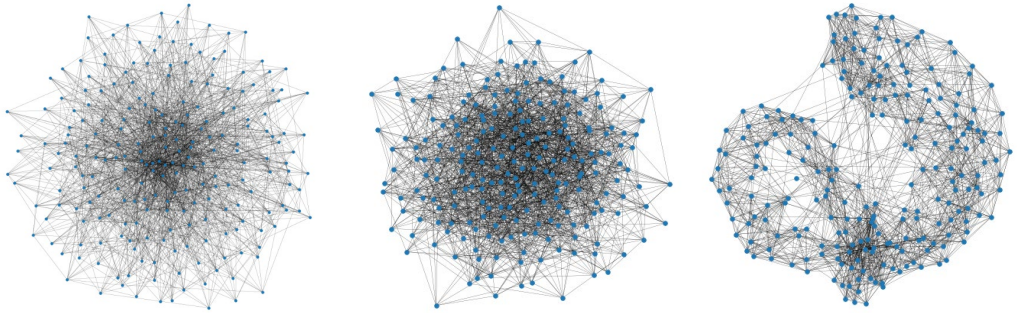
	BA	RG	SW	real
nodes	227	227	227	227
edges	1962	1898	2041	1960
<C>	0.16	0.08	**0.55**	0.55
<L>	2.16	2.19	**2.62**	3; 1.8
<k>	**17.3**	16.7	17.9	17.3
D	3	3	**5**	8; 4

Small world has the best metrics, but still bad…

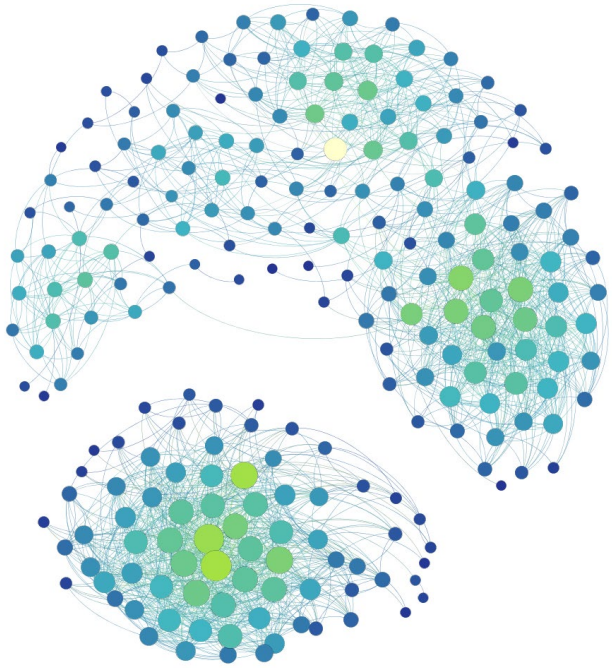
Probability density function



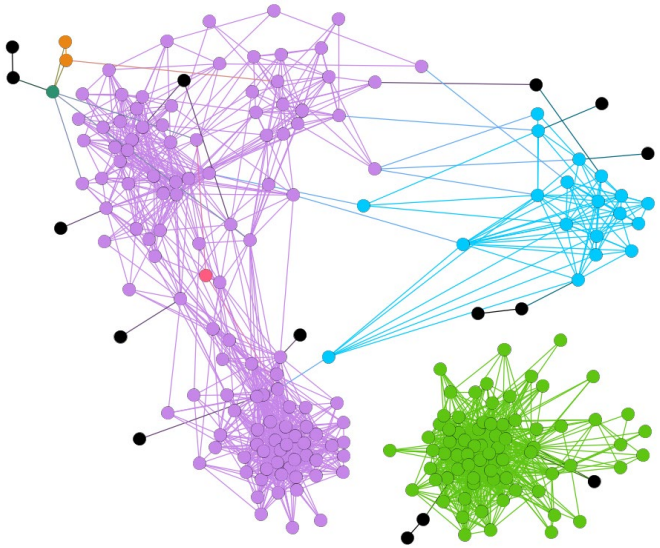
Barabási–Albert | Random graph | "Small world"



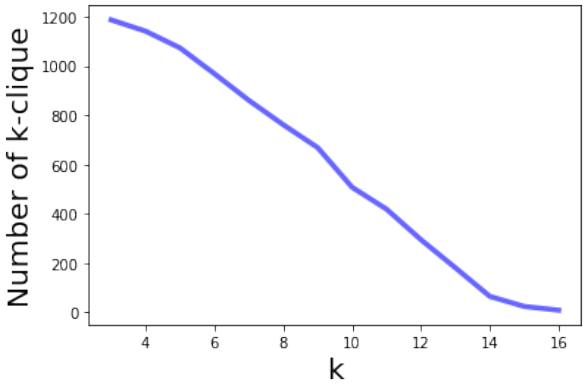
Contents



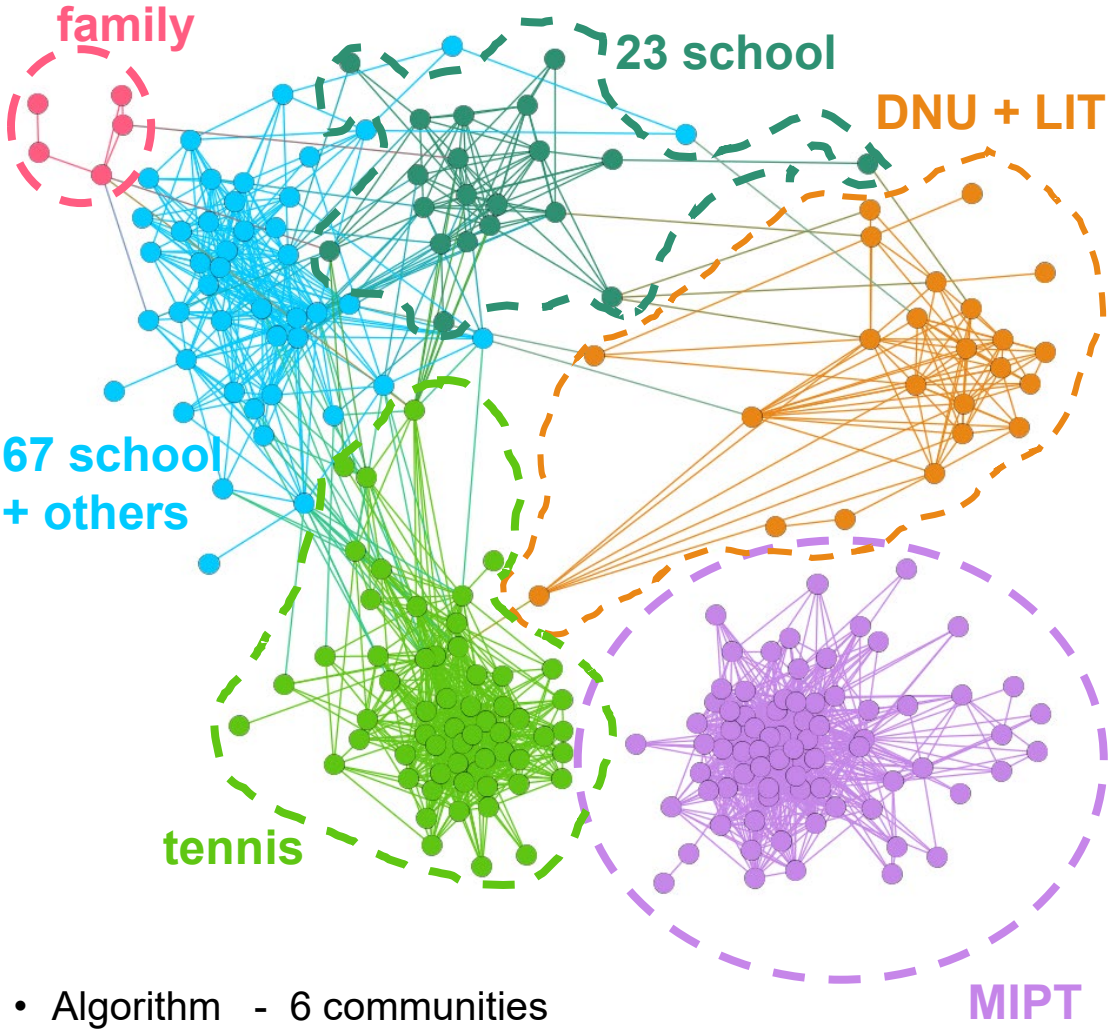
k-Clique search



#k-clique



Best results of various community detection algorithms



- Algorithm - 6 communities
- Real graph - 8 communities

igraph.community_label_propagation()
modularity = 0.35

Thank you for attention