

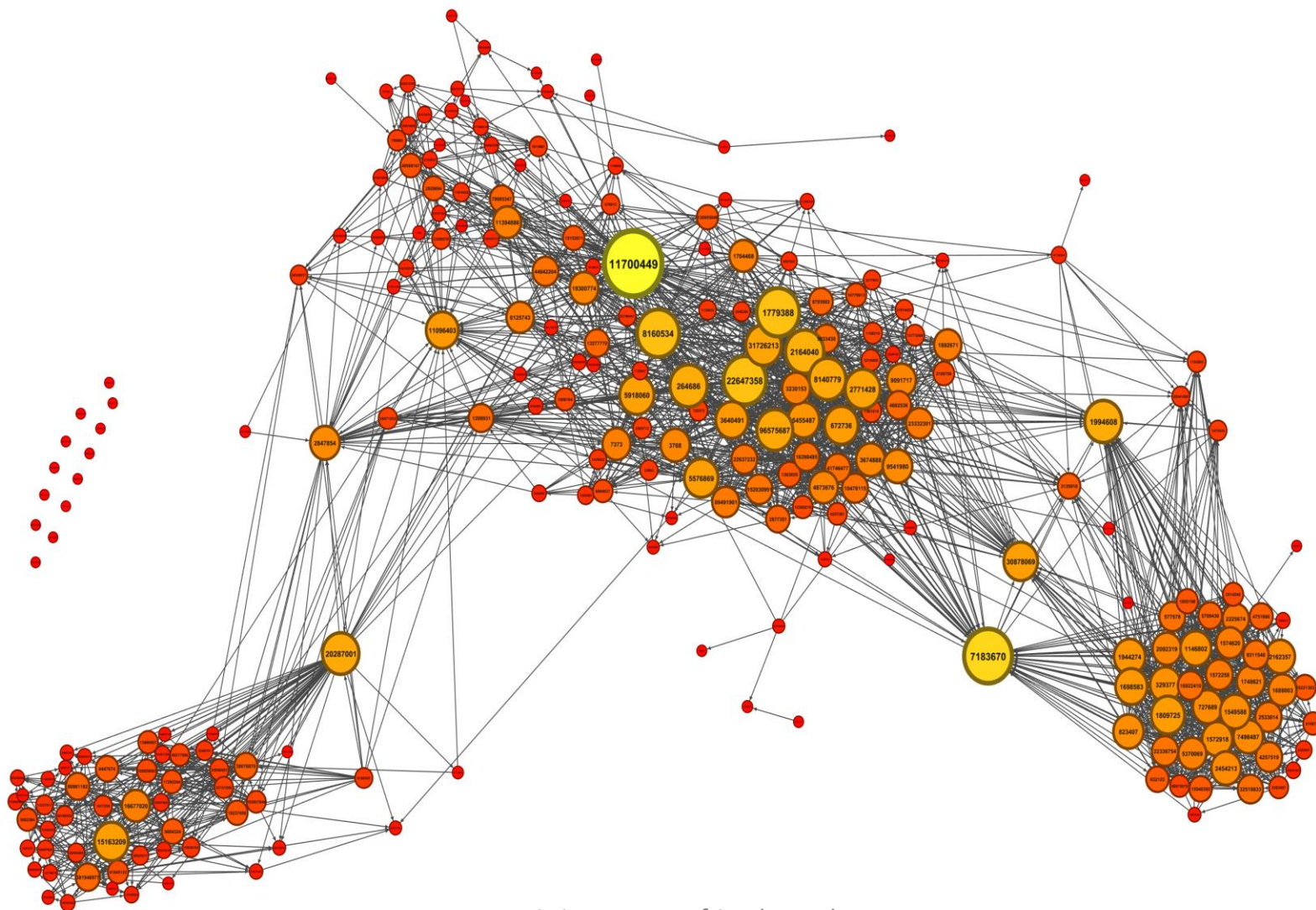
Network science, project 1

Dmitriy Meynster, Data Science
programme, 2017

Download

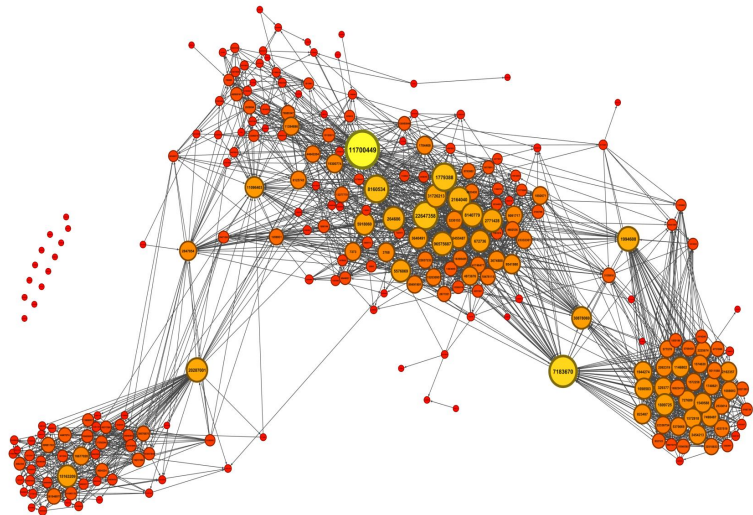
- Data taken from Vkontakte, used VK API
- Requests: friends.get + several friends.getMutual batches
- Deactivated accounts: deleted
- Used attributes: uid, name, city, sex, bdate, schools, universities
- No node attributes

Graph

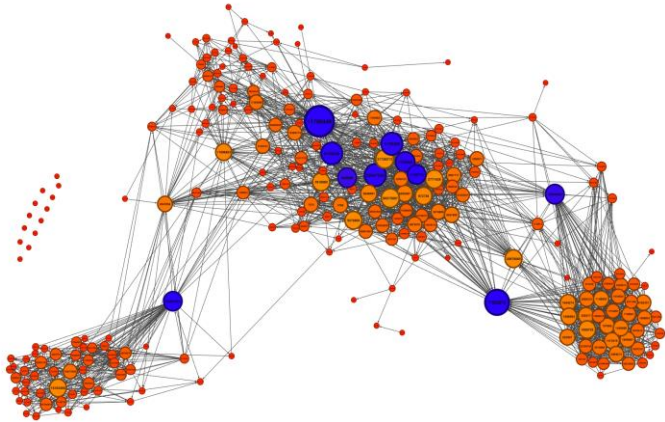


Summary

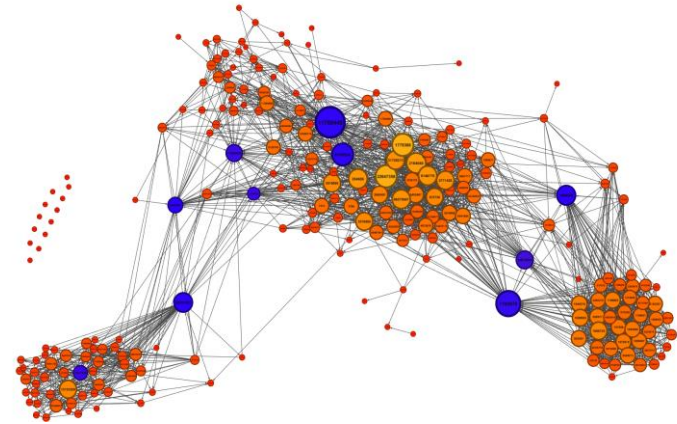
- 250 nodes, 2375 edges
- Far from scale-free
- Maximal connected subgraph: 237 nodes
- Diameter: 7
- 3 main communities



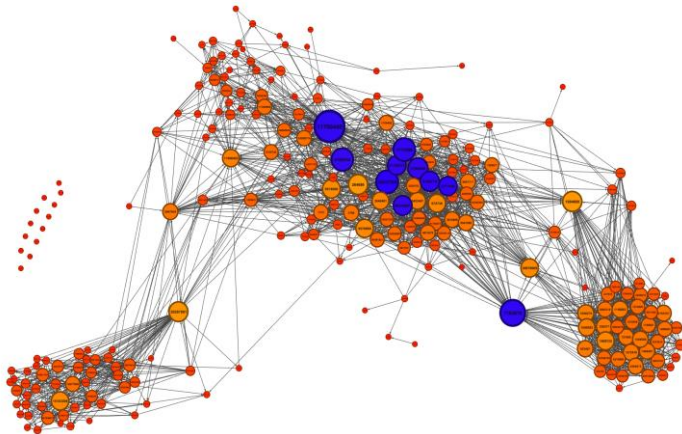
Centralities



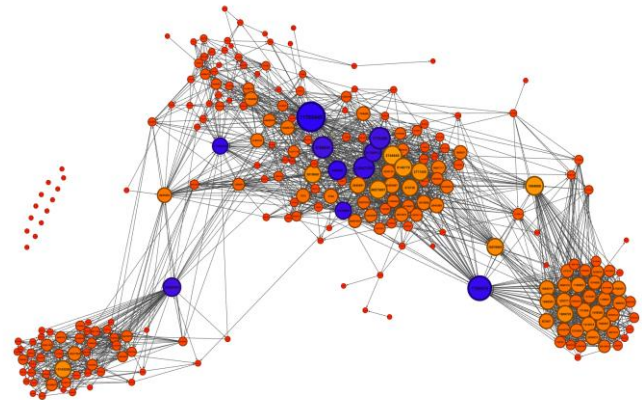
Degree



Betweenness

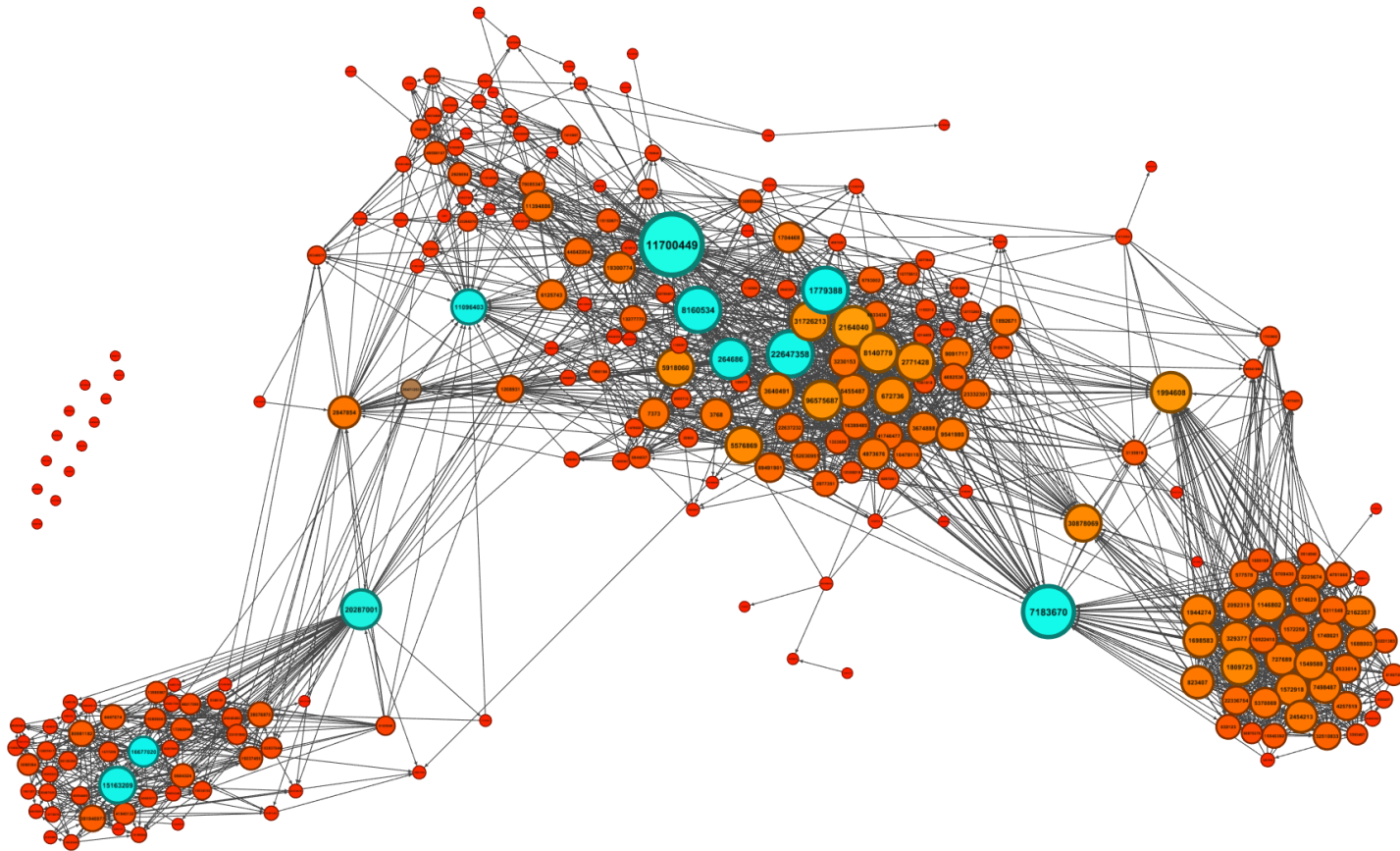


Closeness



Katz

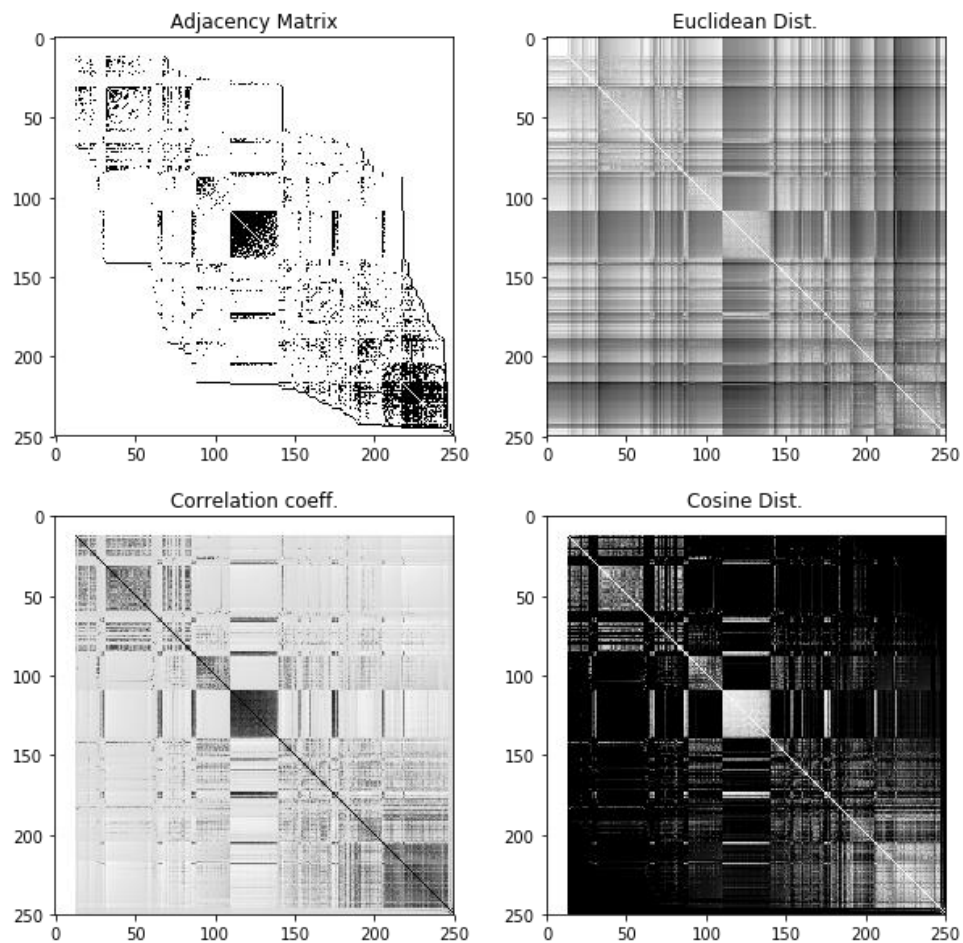
Pagerank



Assortative mixing

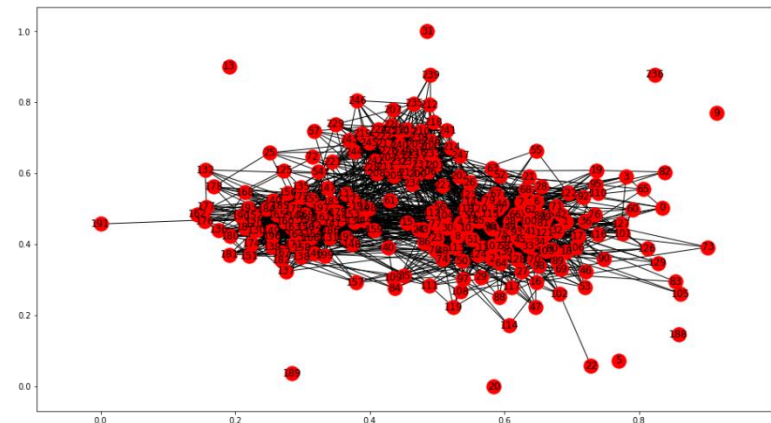
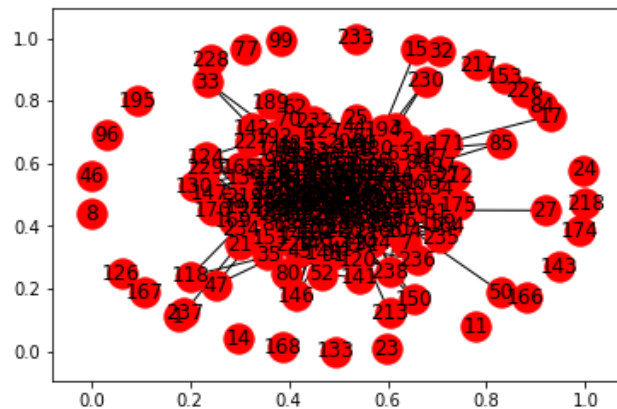
- By degree: 0.043659
- By city: 0.211636
- By sex: 0.135093
- By age: 0.027326

Node similarity



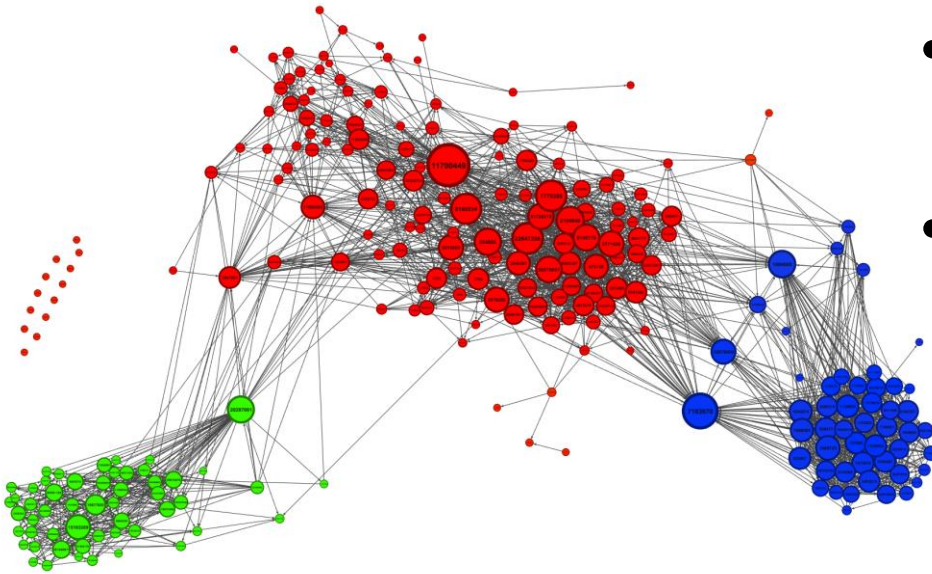
Random graph model

- Erdos-Renyi model – does not suite graph
- Model B (preferential attachment without growth) – much better, but only one cluster
- Solution: generate several Model B's, then merge and iterate a bit more

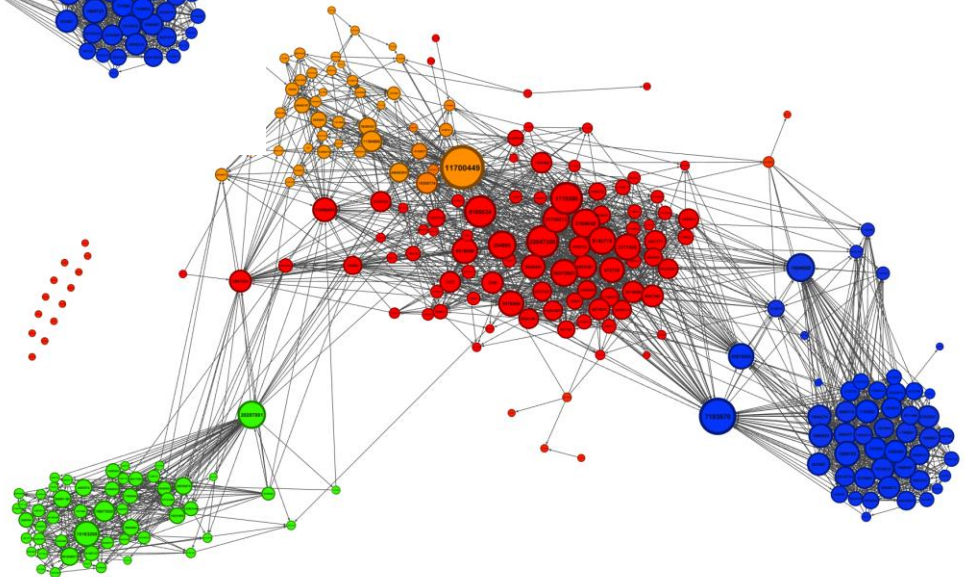


Community summary

- Comparison by modularity and visual appearance
- Louvain community detection was the best



- 3-4 communities overall



Thanks!

- Questions?