НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Structural Analysis and Visualization of Social Network

Marina Ananyeva

National Research University
Higher School of Economics

March 30, 2017

Outline

- **Network Source and Preprocessing**
- **Network Summary**
  - Node/edge attributes
  - Order, size, layout
  - Degree distribution, diameter, Cluster Coefficient
- **Structural Analysis**
  - Degree/Closeness/Betweenness centralities
  - Comparison of PageRank and centralities
  - Assortative Mixing according to node attributes
  - Node similarity
  - The closest similar random graph
- **Community Detection**
  - Clique search
  - Results of detection algorithms
  - Incremental algorithm

Network Source and preprocessing

Source: *vk.com*
Data access: via VK API requests
Unweighted undirected graph.
The network is egocentric.
Preproccessing:
- 'Deactivated' users were manually deleted from the dataset
- Decoding (e.g. city code into name: '0' -> 'Moscow')

Network Summary

Node Attributes:

- First Name
- Last Name
- Sex
- City
- Birthday date

Edge attributes: None (weight, relationship)

Order: 341 (nodes)

Size: 3302 (edges)

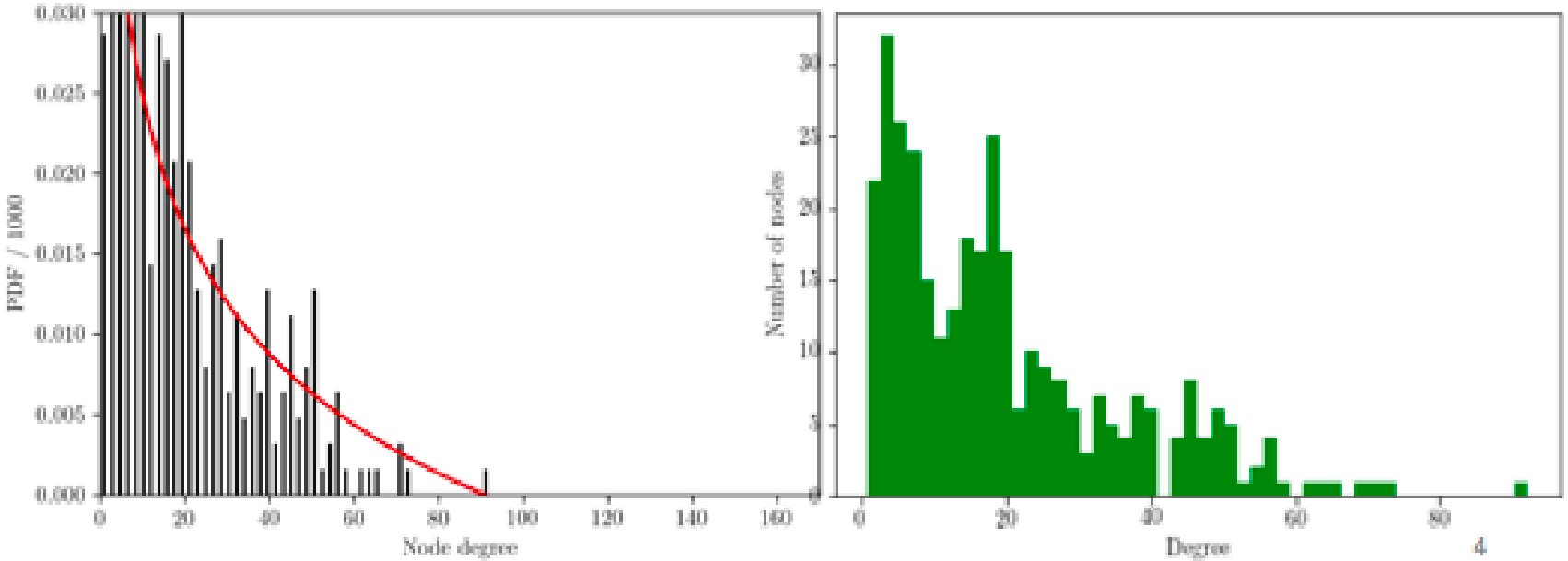
Degree distribution

Average clustering coefficient: 0.5

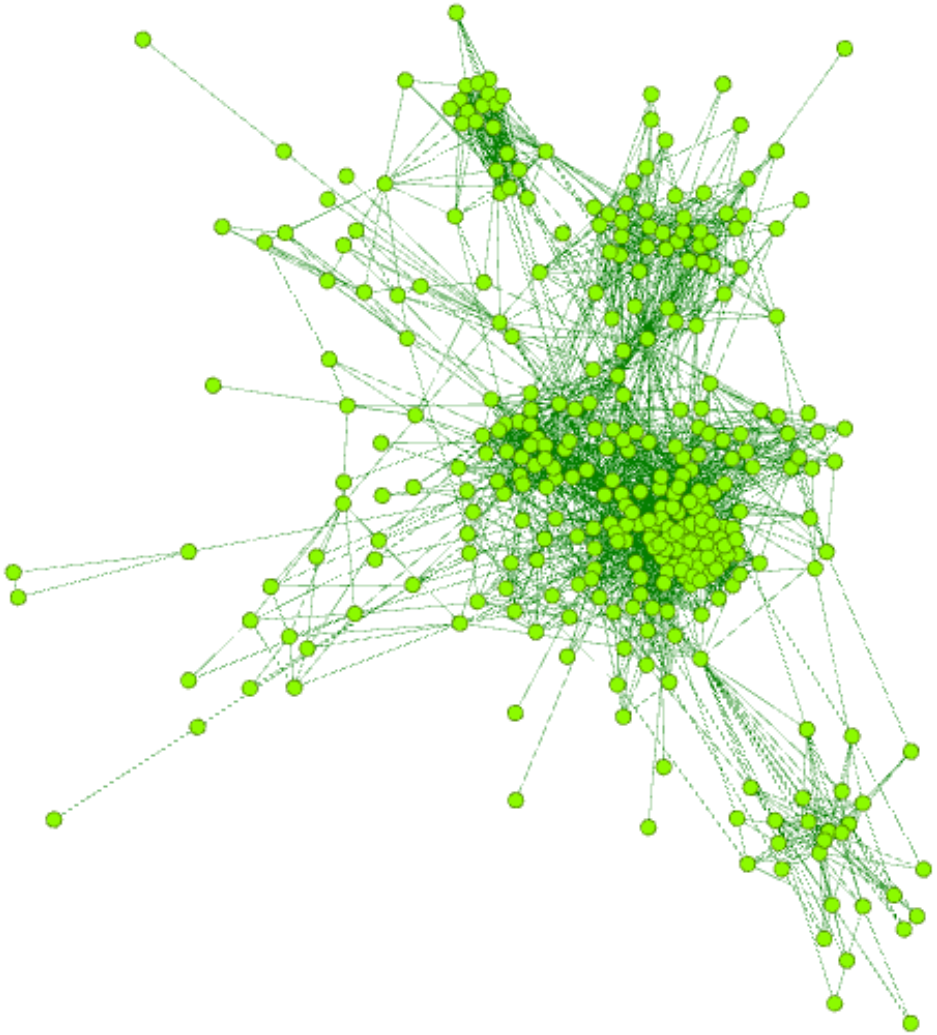
Average path length: 2.9

Radius: 4

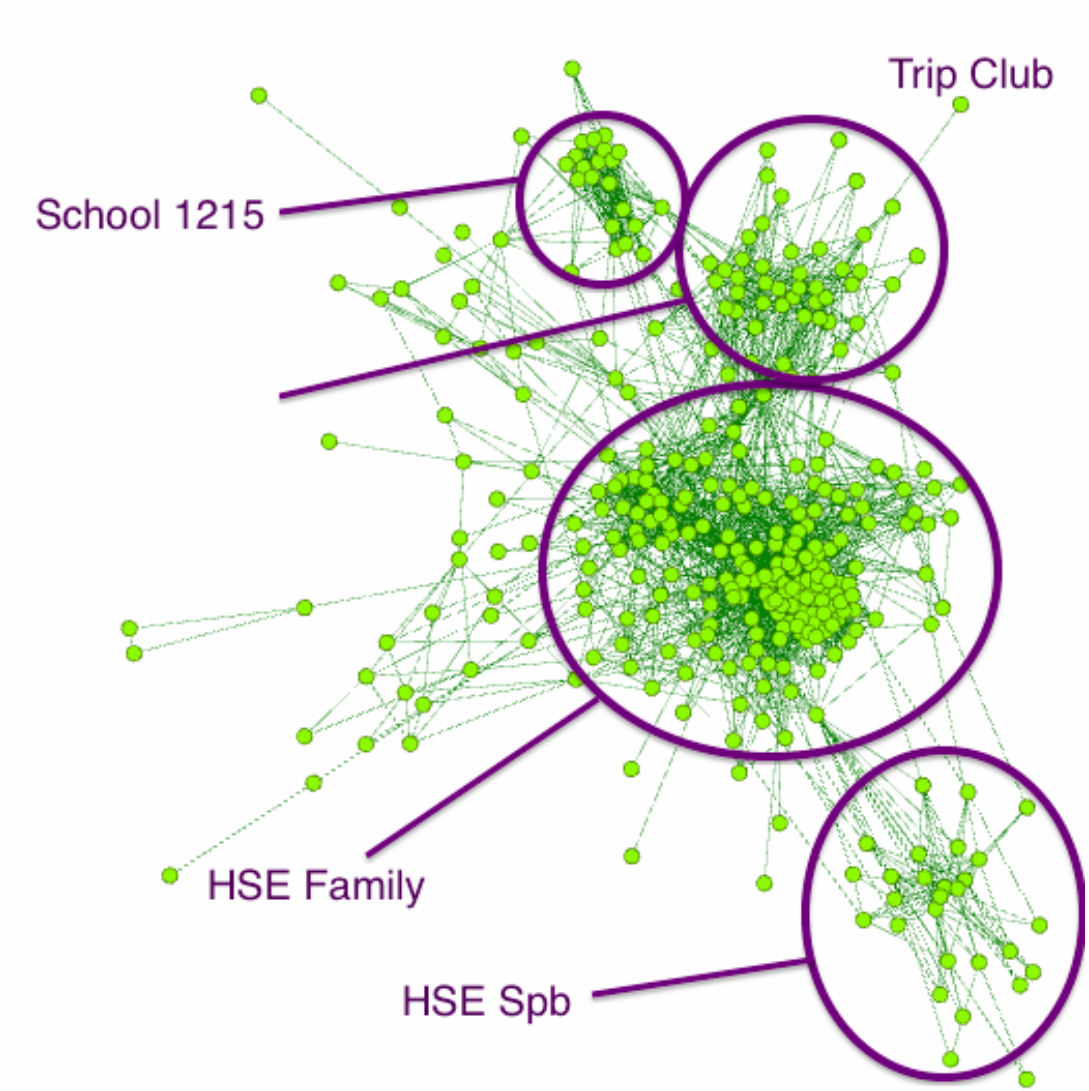
Diameter: 7



Network layout



Interpretation

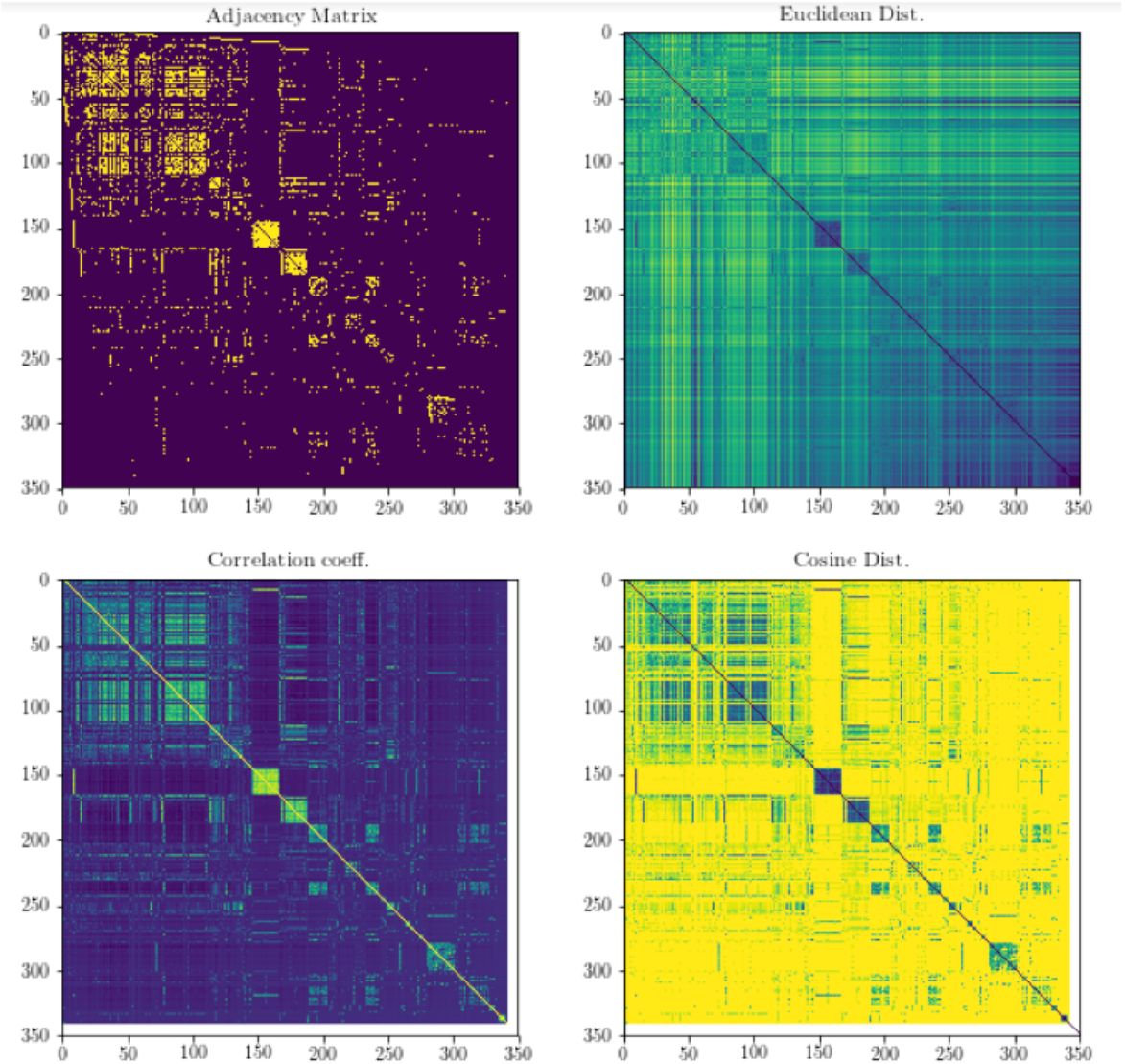


Centralities and PageRank

- **Degree centrality** identifies those friends who share with me a lot of other friends.
- **Closeness centrality** shows the closest friends who indirectly know my other friends (the length of the shortest paths between the node and all other nodes in the graph).
- **Betweenness centrality** helps to see those who connect diverse groups with each other

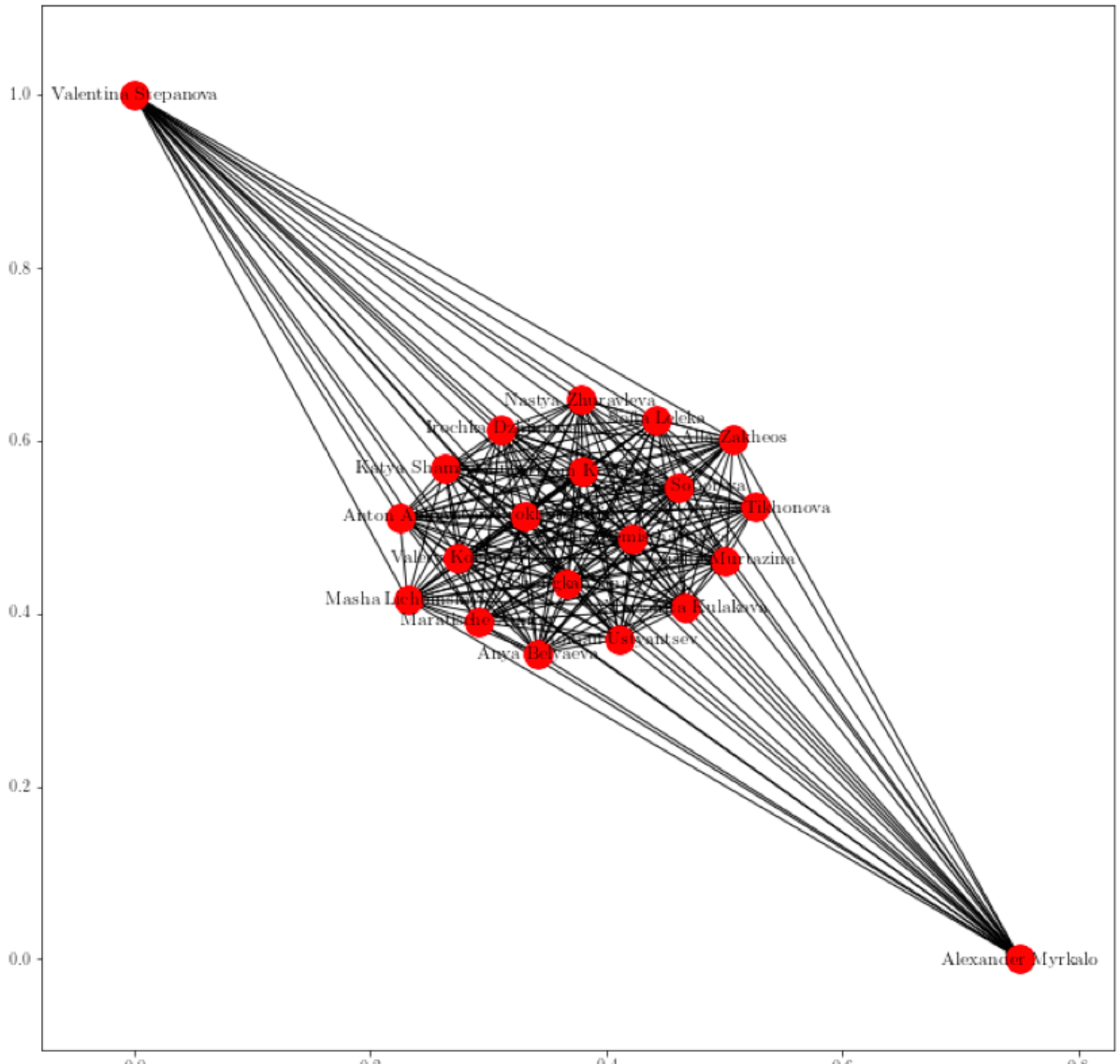
	Degree	Closeness	Betweenness	PageRank
1	**Ksenia Soborova** (0.27)	**Ksenia Soborova** (0.48)	Dasha Ukolova (0.348)	**Ksenia Soborova** (0.01)
2	Roman Ustyancev (0.22)	Valery Okunev (0.475)	Olga Ivanova (0.082)	Dasha Ukolova (0.009)
3	Sofia Leleka (0.21)	Dasha Ukolova (0.466)	Valery Okunev (0.07)	Dana Smoliakova (0.001)
4	Anton Akhiev (0.21)	Roman Ustyancev (0.463)	**Ksenia Soborova** (0.07)	Roman Ustyancev (0.001)
5	Marat Amirov (0.19)	Sofia Leleka (0.46)	Dana Smoliakova (0.065)	Sofia Leleka (0.001)

Assortativity Mixing and Structural similarity

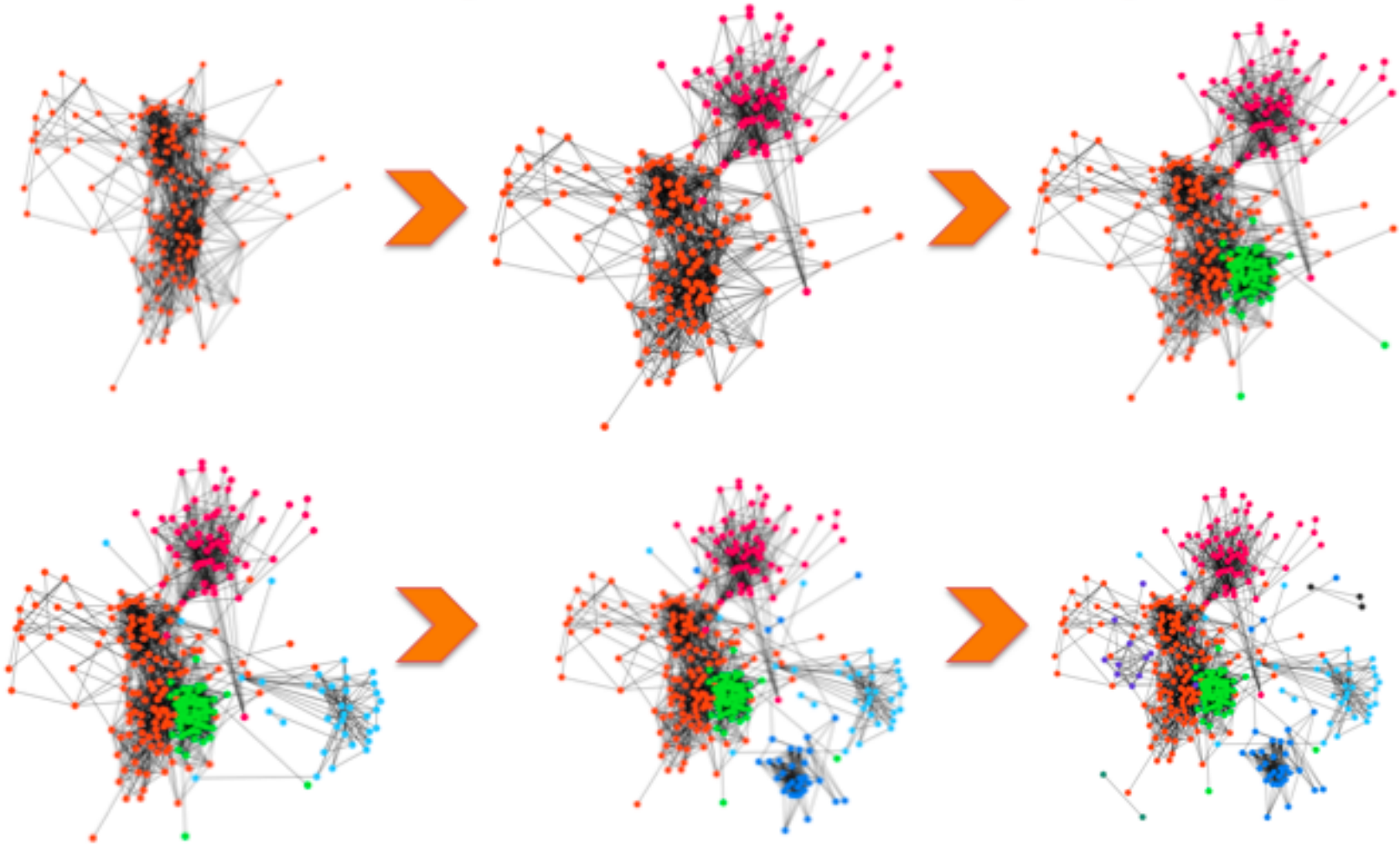


Clique search

- Largest clique
- Max edges: 162



Community detection*: step by step



* via igraph.community_fastgreedy

Thank you!

Any questions?