

Collaboration network analysis of scientific production at UB-SEE

Jovan Stojković, Marko Mišić, Jelica Protić

University of Belgrade, School of Electrical Engineering

27th Telecommunications Forum TELFOR 2019

Introduction (1)

- ▶ Scientific production is an important indicator of the development of scientific society
 - ▶ Affects ranking of individuals and institutions
 - ▶ Used in evaluation of project proposals, etc.
- ▶ Scientific production at University of Belgrade, School of Electrical Engineering (UB-SEE) is the focus of this paper
 - ▶ UB-SEE is mainly educational institution
 - ▶ The other important activity is research
 - ▶ Research results published in scientific conferences and journals

Introduction (2)

- ▶ Usually, researchers work in groups that publish papers together
 - ▶ They are called coauthors
- ▶ The coauthors can work in the same institution or they can be affiliated to different institutions
 - ▶ In our case, co-authors from UB-SEE and other institutions
- ▶ The goal of our research was to analyze collaboration patterns of UB-SEE employees
 - ▶ Such analysis can show the directions of the future research
 - ▶ We can find the institutions having the most similar research interests to that of UB-SEE

What we did

- ▶ Analysis of the collaboration network of coauthors of research papers published in journals where at least one author is affiliated to UB-SEE.
- ▶ Raw data from the school's database and processed 2000 papers published in journals from 2000 to 2017.
- ▶ Python scripts for data cleansing and conversion of names of the authors in canonical form
- ▶ The network of coauthors is created using MS Excel and Python scripts, while the visualization and further analysis is performed using Gephi software tool.
- ▶ The network is analyzed based on the metrics used in social network analysis

Network modeling

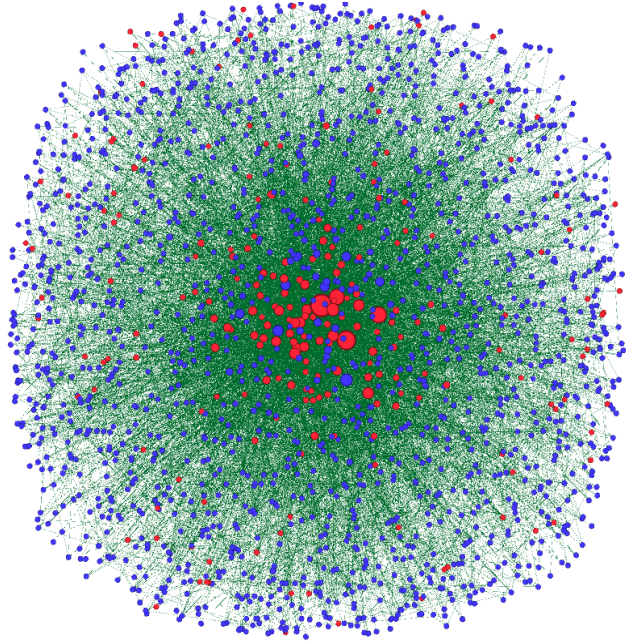
- ▶ Python scripts with NetworkX library for collaboration network modeling
- ▶ Nodes - authors of the papers published in journals.
- ▶ Edges - connection between coauthors
- ▶ Weight - number of papers where two authors collaborated
- ▶ Network is weighted and undirected.

New database

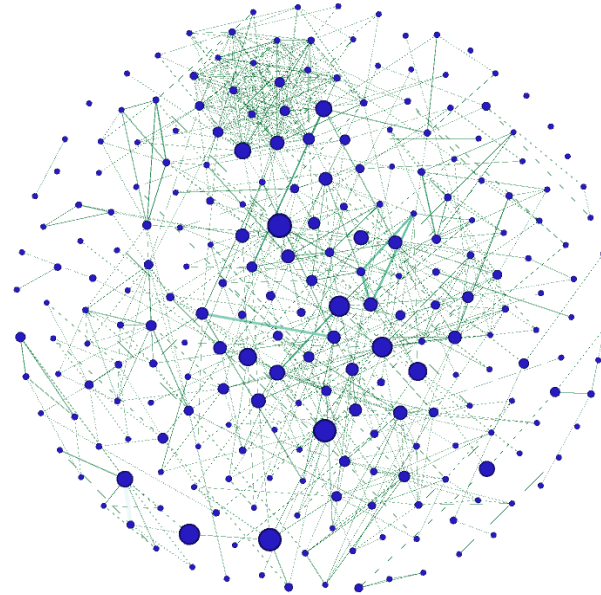
- ▶ We extracted all the authors, e.g. the nodes of our network, from the database with published papers. There are two possible types of nodes: UB-SEE's researcher and non UB-SEE's researcher.
- ▶ Our edge has five components: author A, indicator if author A is UB-SEE researcher, author B, indicator if author B is UB-SEE researcher and number of papers on which A and B collaborated.

Author_A	Flag_A	Author_B	Flag_B	Weight
B. Reljin	1	P. Kostic	0	8

Visualization of the network -> Gephi



Collaboration network visualized by Gephi software tool. Blue nodes present non-UB-SEE's researchers, while the red nodes stand for UB-SEE's researchers



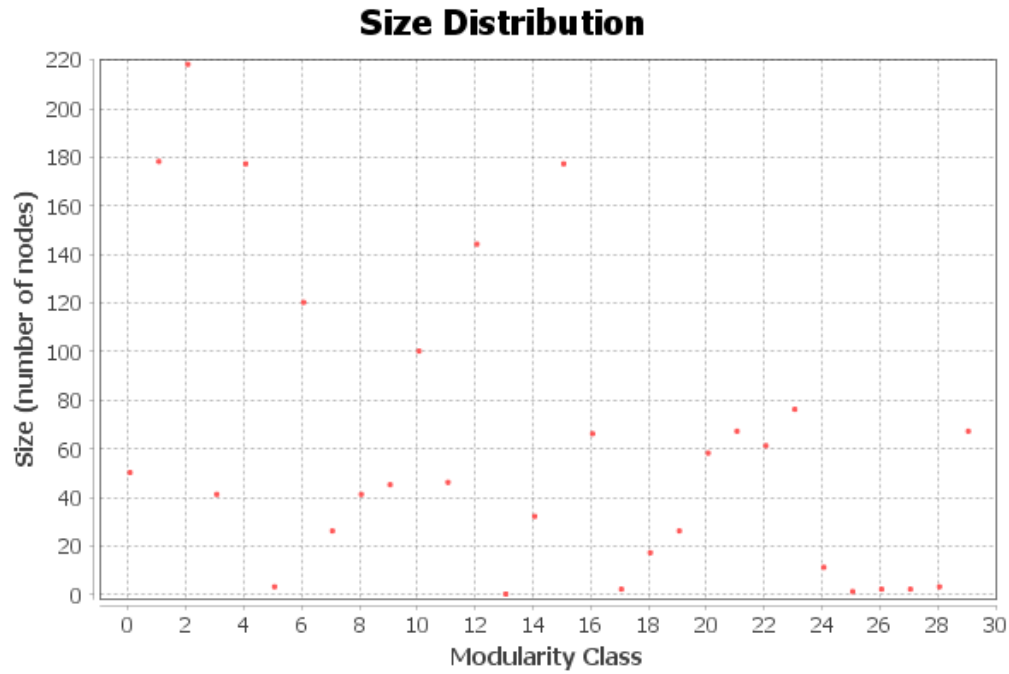
Collaboration network consisting only of UB-SEE's researchers visualized by Gephi software tool

Metrics

Metric	Value
Graph density	0.004
Average degree	7.585
Average weighted degree	14.082
Network diameter	11
Average path length	4.315
Connected components	10
Average clustering coefficient	0.82

Network characteristics

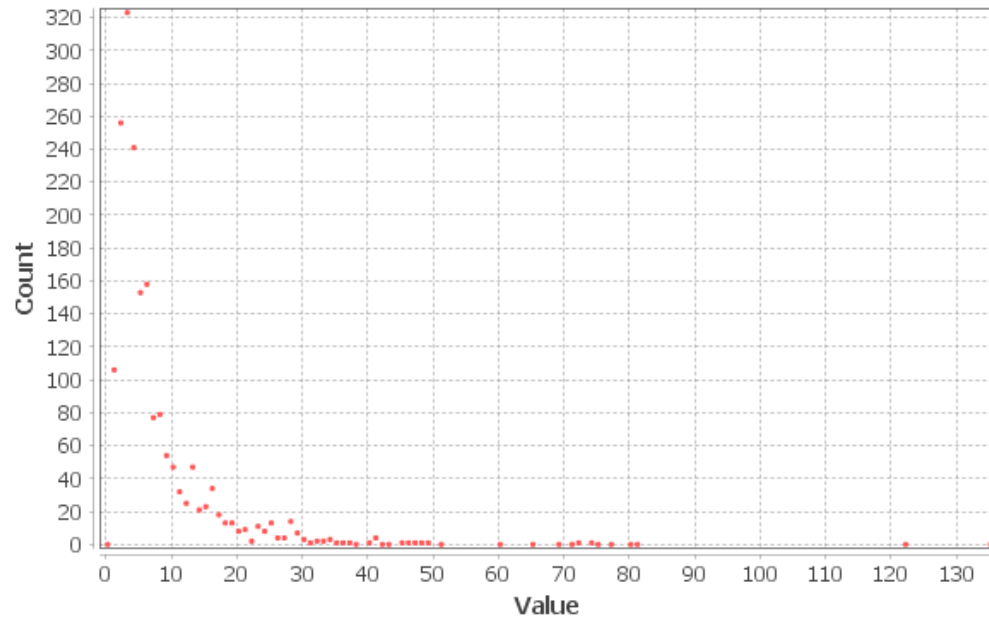
- ▶ Analyzed network is sparse
 - ▶ Lot of edges are not connected
- ▶ UB-SEE collaboration network exhibits small-world phenomenon
 - ▶ Relatively small diameter and average path length
 - ▶ Similar to Facebook and other social networks
- ▶ The network consists of one giant component - core of the network
 - ▶ There is one big connected component consisting of 1839 nodes
 - ▶ Other 9 components have all together 55 nodes
- ▶ High clustering coefficient show us “all-my-friends-know-each-other” property
 - ▶ Co-authors work together in well-established groups
- ▶ Modularity suggests that our network has dense connections between the nodes within modules but sparse connections between nodes in different modules
 - ▶ Revealing community structure of the network



Size Distribution

30 classes having from 1 to 220 nodes. Researchers grouped into communities based on the department.

Degree Distribution



Degree Distribution

Our network follows a power law, i.e. the network is scale-free

Authors ranking

- ▶ All 5 highest ranked authors based on the degree are UB-SEE's researchers except of P. Nikolic, who was a member of Serbian Academy of Sciences and Arts (SASA).
- ▶ Mirjana Popovic, D. Popovic, and A. Djordjevic also serve as bridges in their scientific fields, not only that they have many co-authors.

Researcher	Degree
D. Popovic	122
Mirjana Popovic	111
A. Djordjevic	81
P. Nikolic	79
K. Stankovic	77

Researcher	Betweenness centrality
D. Popovic	0.117621
A. Djordjevic	0.098955
Mirjana Popovic	0.098258
S. Stankovic	0.092957
V. Milutinovic	0.061902

Conclusion

- ▶ Institutional databases offer great possibilities to analyze scientific production and co-authorship (collaboration) patterns of its employees.
- ▶ In this paper, we have analyzed collaboration network of UB-SEE employees and their collaborators from other institutions. We showed that analyzed network exhibits the properties of a social network and point out the most important researchers in terms of their collaboration patterns.



Thank you!

Questions?

Jovan Stojković, jovan.stojkovicvr1997@gmail.com

Marko Mišić, marko.misic@etf.bg.ac.rs

Jelica Protić, jelica.protic@etf.bg.ac.rs