Homework 4: Network analysis and visualization

Florian Brandl, Dennis Ledwon, Clément Bernier, Pia Koller

Group 8

November 2021

1 Report

- Create a node-link diagram differentiating the two types of nodes.

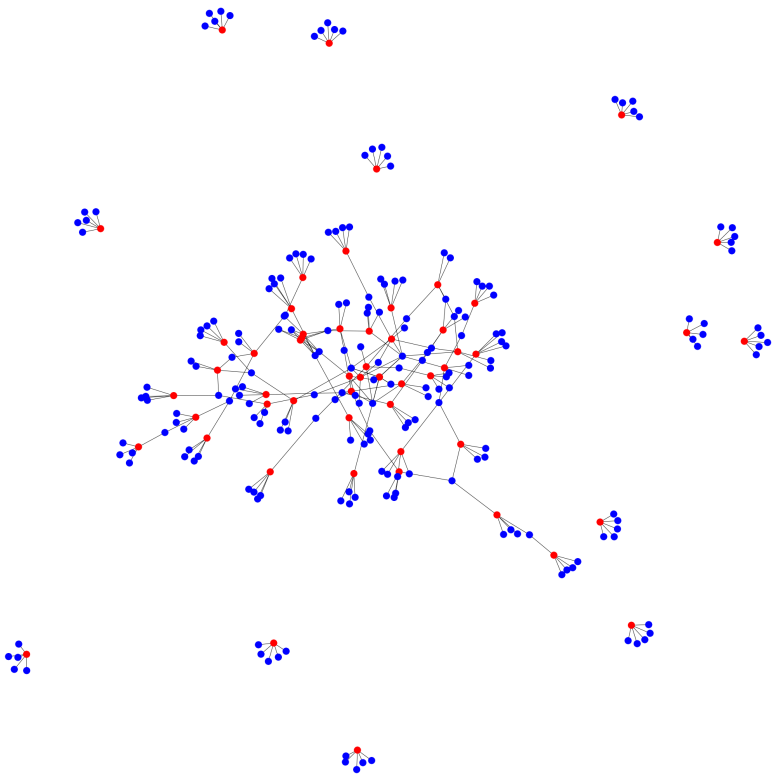

Figure 1: Node-link diagram differentiating the tow types of nodes (Persons and Datasets)

- "How many components are there in this graph?"
  - 14
- Extract the largest component:
  - 'Bernier, Clement',
  - 'Brandl, Florian Lukas',

  - 'Chai, Kyuwook ',
  - ...
  - 'https://www.kaggle.com/zygmunt/goodbooks-10k'

- Names of people who do not belong to the main component:
  - Kim, Jaehee
  - Kim, Junmo
  - Ham, Yoonhee
  - Koller, Pia
  - Shin, Yoon Jae
  - Bansal, Parth
  - Sluimer, Jasper
  - Kim, Chongmin
  - Kim, Hyunwoo
  - Oh, Sewon
  - Na, Doori
  - Doret, Norman
  - Buisson, Antoine

- Visualize both persons graph and datasets graph using node-link diagrams
  - Persons graph

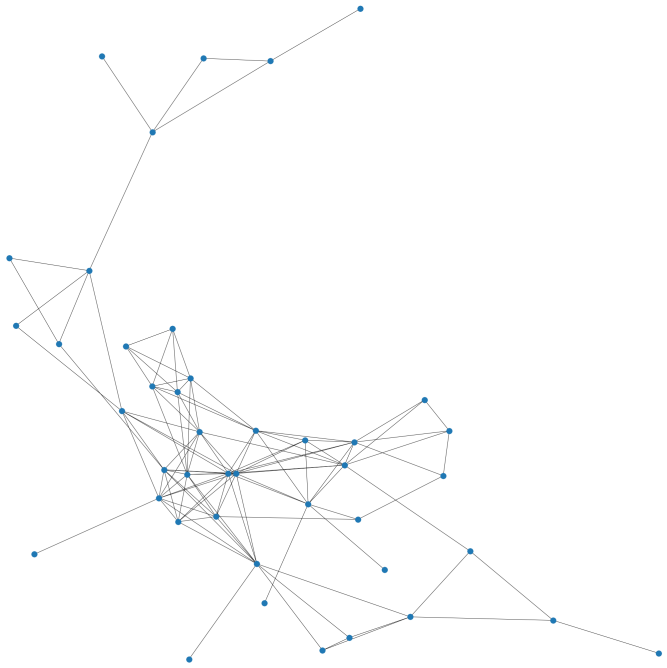

Figure 2: Node-link diagrams of persons graph

– Dataset graph

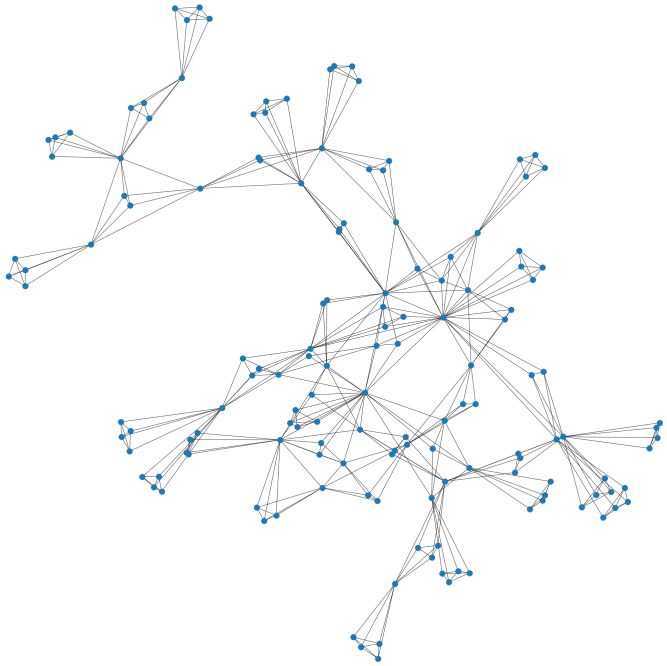

Figure 3: Node-link diagrams of dataset graph

- Compute closeness centralities for all nodes in the persons graph
  - 'Kim, Myungdong': 0.4883720930232558,
  - 'Lee, Sumin': 0.3684210526315789,
  - 'Logghe, Jubeline Jacqueline L': 0.32558139534883723,
  - ...
  - 'Chun, Ye Ji': 0.40384615384615385
- Compute betweenness centralities for all nodes in the persons graph
  - Kim, Myungdong': 0.1411841159227919,
  - 'Lee, Sumin': 0.029849012775842034,
  - 'Logghe, Jubeline Jacqueline L': 0.05533709418726842,
  - ...
  - 'Chun, Ye Ji': 0.40384615384615385
- Compute the correlation coefficient between the two centrality measures

  [1. , 0.52622229],

  [0.52622229, 1.]