Marvel Heroes

Martin Jenc

Dataset Description

- All Marvel Comics and Heros since 1961
- Heroes names
- Comics names
- Heroes and lists of comics where they appeared
- Source
  - http://syntagmatic.github.io/exposedata/marvel/
    - Also with awesome analysis using Gephi
  - https://www.kaggle.com/csanhueza/the-marvel-universe-social-network

Basic description

- Comics: 12,651
- Heroes: 6,439
- Hero → Comic: 96,106
- No additional properties
- Heroes with most appearances
  - SPIDER-MAN (PETER PARKER): 1577
  - CAPTAIN AMERICA : 1334
  - IRON MAN (TONY STARK): 1150
- Comics with most heroes
  - Contest of Champions: 111
  - The Infinity War #3: 91
  - The Infinity War #2: 90

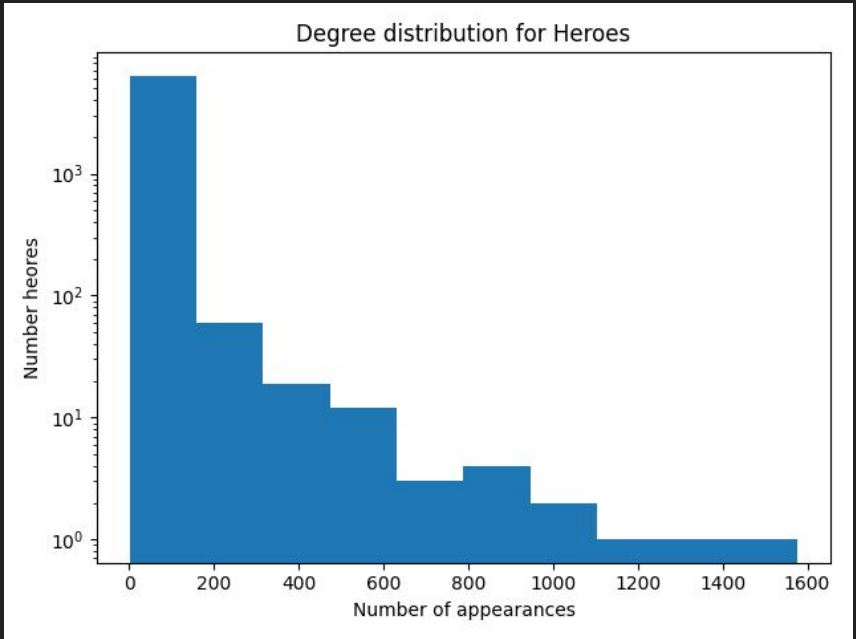
Goal of the analysis

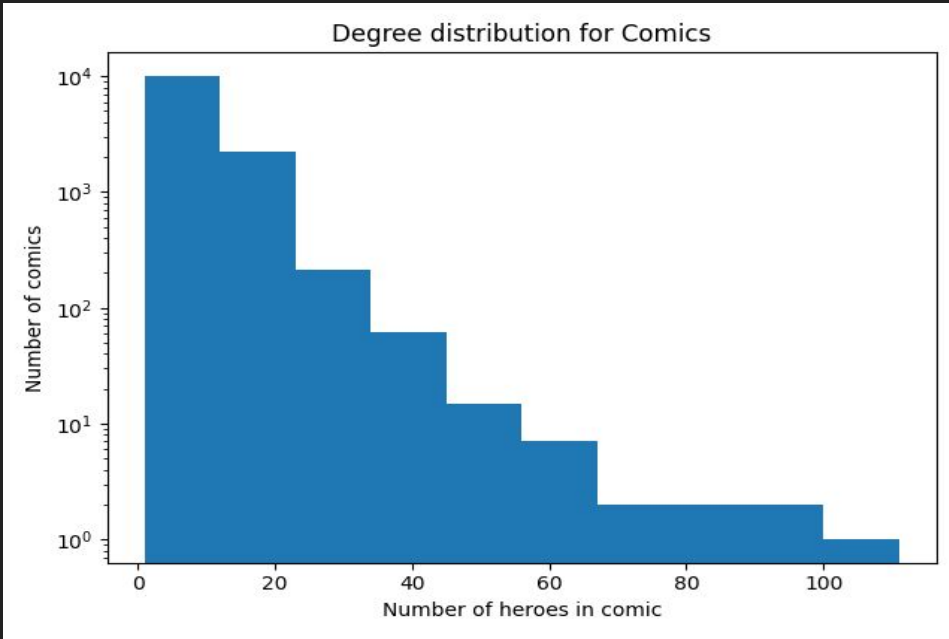
- Reveal basic properties - real world network?
- Then focus on the heroes..
- Find interesting communities
- Find key heroes
- Predict co-appearances of heroes in future comics

Basic Properties

- Bipartite graph
- Nodes: 19,090
- Edges: 96,106
- Connected Components: 22
  - 19029, 11, 8, 4, 3, 3, 2, 2, …
  - Heroes appearing only, in few comics
  - So let’s consider just largest component
- Degree: mu=3.37, q0.5 = 1, q0.75 = 3, q0.9 = 6, q0.99 = 11, max = 1577

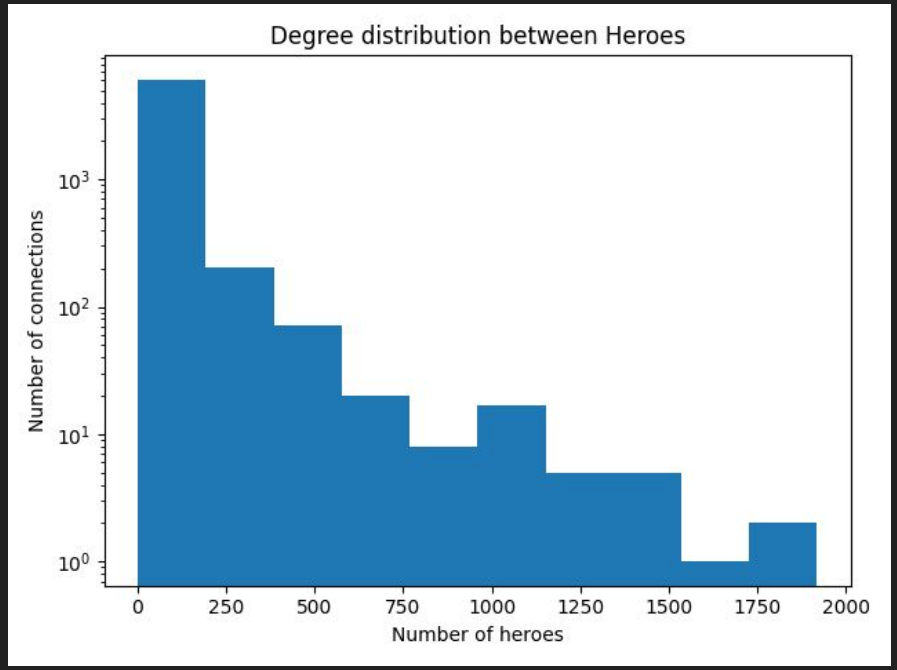
Degree Distribution





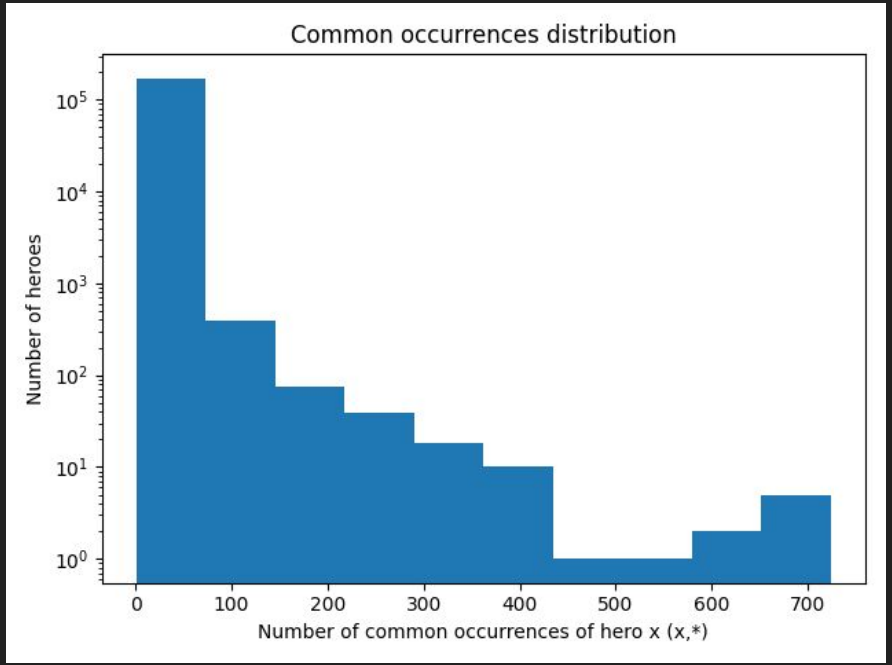
Bipartite → Heroes Graph

- Weighted undirected graph
- Edge weight := # of appearances of H1 and H2 together
- Tightly connected network
- Global clustering coeff.: 0.79
- Average path length: 2.73
- Diameter: 4, Radius: 2
-



Finding Communities

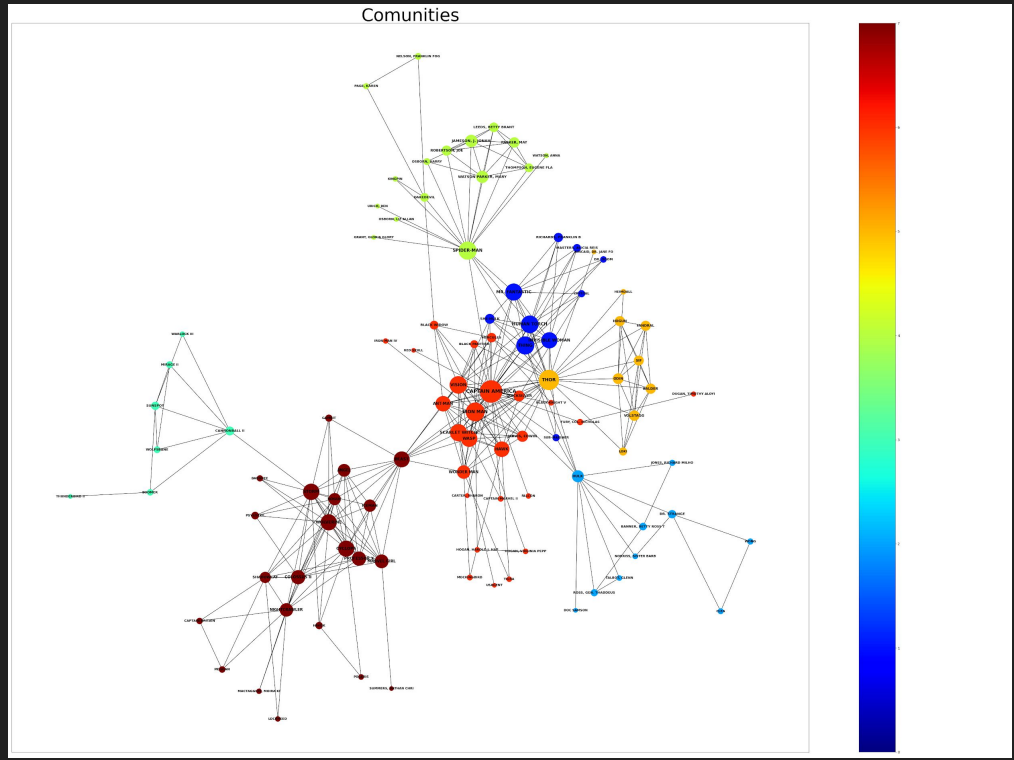
- Consider only edges with high weight (100+)
- Limit the result to top 100 most known heroes
- Plot the result
- Visually interpret the results
- Weight distribution:



Communities (Louvian Method)

https://raw.githubusercontent.com/jencmart/mff-sna-nail116/master/Assignment-4-term-project/images/communities2.png

- 7 communities
- Comparable sizes:
  - 22
  - 26
  - 16
  - 10
  - 10
  - 10
  - 7

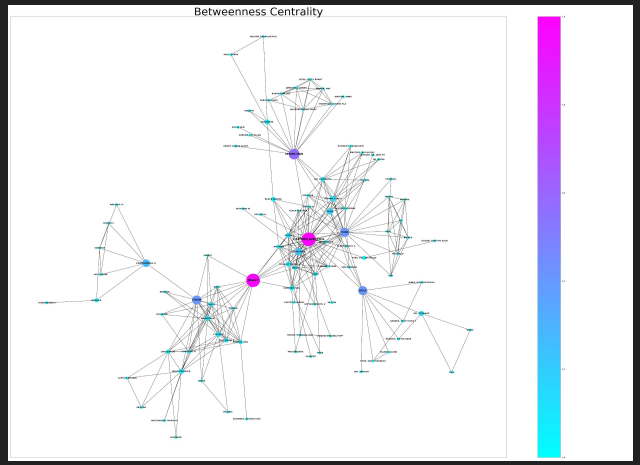


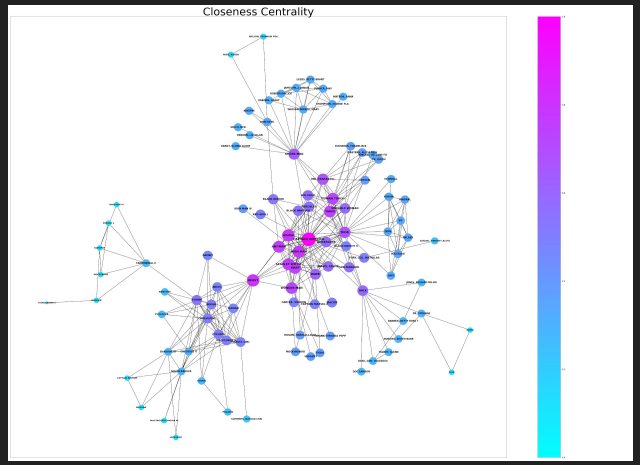
Centralities

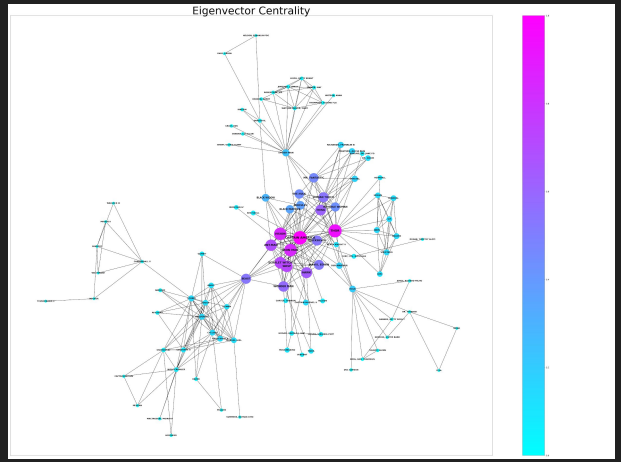
Betweenes: https://raw.githubusercontent.com/jencmart/mff-sna-nail116/master/Assignment-4-term-project/images/betweenes.png

Closeness: https://raw.githubusercontent.com/jencmart/mff-sna-nail116/master/Assignment-4-term-project/images/closeness.png

Eigenvector: https://raw.githubusercontent.com/jencmart/mff-sna-nail116/master/Assignment-4-term-project/images/eigenvector.png

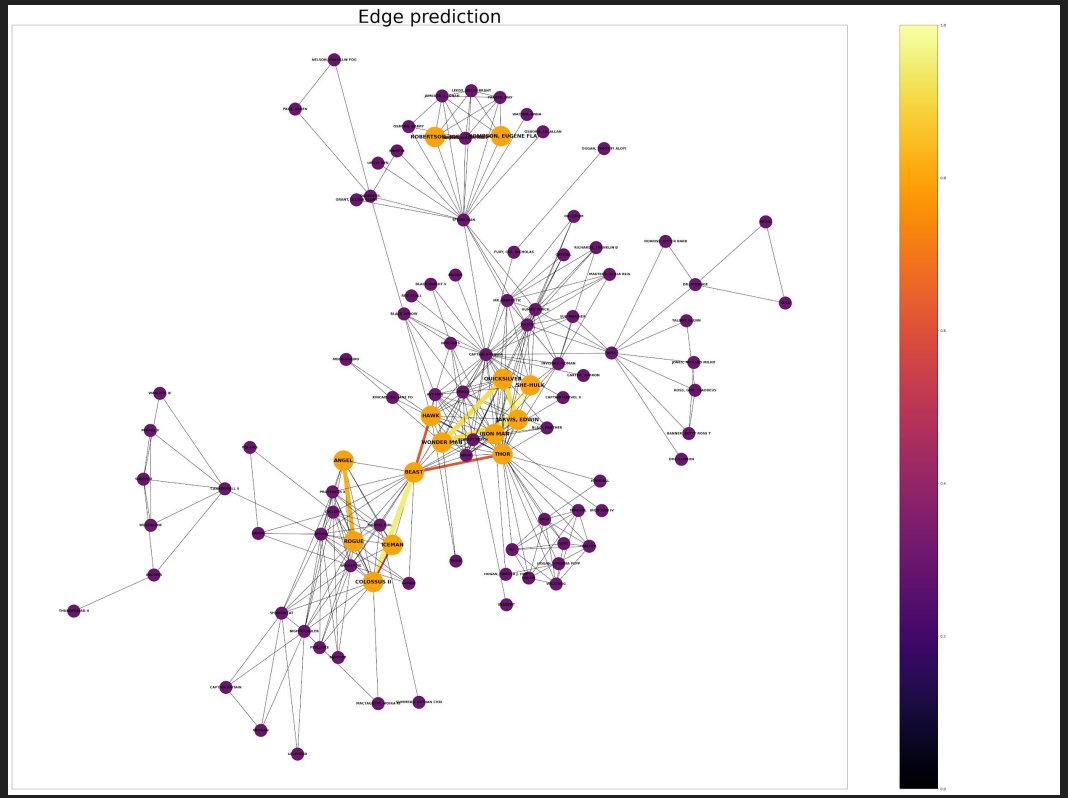






Co-appearance Prediction (Adamic-Adard Measure)

('SHE-HULK', 'IRON MAN') : 2.91
('COLOSSUS II', 'BEAST') : 2.86
('WONDER MAN', 'JARVIS, EDWIN') : 2.83
('WONDER MAN', 'QUICKSILVER') : 2.83
('QUICKSILVER', 'JARVIS, EDWIN') : 2.83
('ANGEL', 'ROGUE') : 2.74
('HAWK', 'BEAST') : 2.55
('THOR', 'BEAST') : 2.55
('ROBERTSON, JOE', 'THOMPSON, EUGENE FLA') : 2.49
('COLOSSUS II', 'ICEMAN') : 2.45



https://raw.githubusercontent.com/jencmart/mff-sna-nail116/master/Assignment-4-term-project/images/edge_prediction.png

Question?

Thank You !