NetworkX (Python)

Agata Skibińska
Maksym Telepchuk

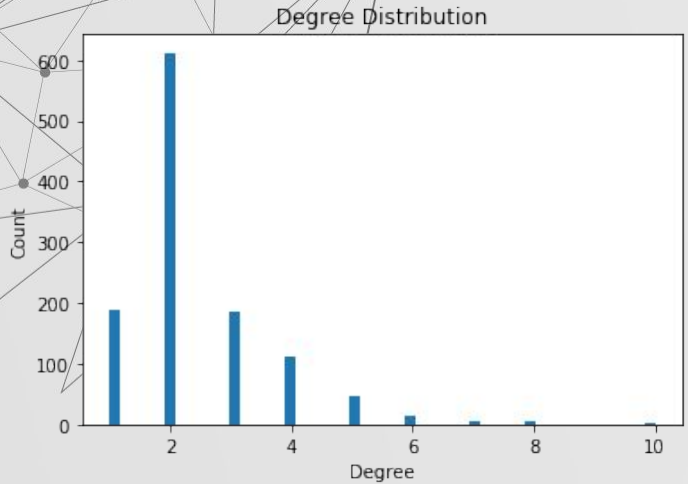
Networks

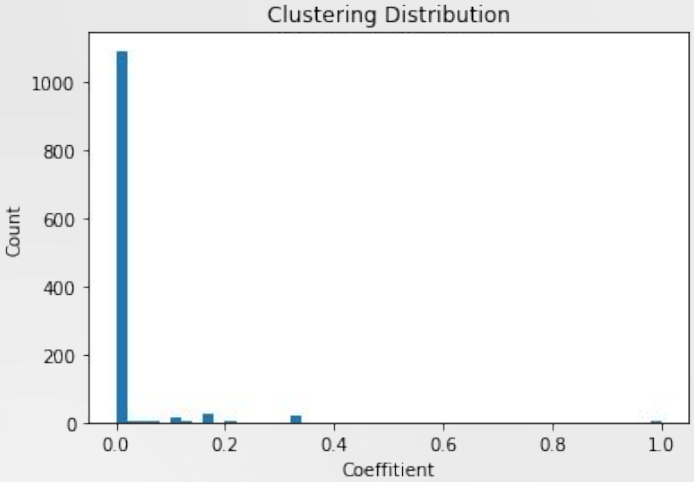
1. Euroroad
    - International E-road network, a road network located mostly in Europe.
    - **Undirected**
    - **1 174** vertices
    - **1 417** edges
2. arXiv astro-ph
    - Authors of common papers from the arXiv's Astrophysics section.
    - **Undirected**
    - **18 771** vertices
    - **198 050** edges
3. Livemocha
    - Social network of Livemocha, an online language learning community.
    - **Undirected**
    - **104 103** vertices
    - **2 193 083** edges

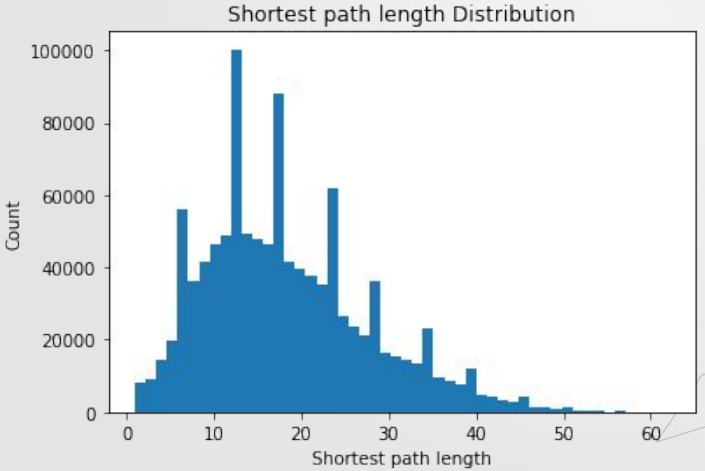
NetworkX time performance (in secs)

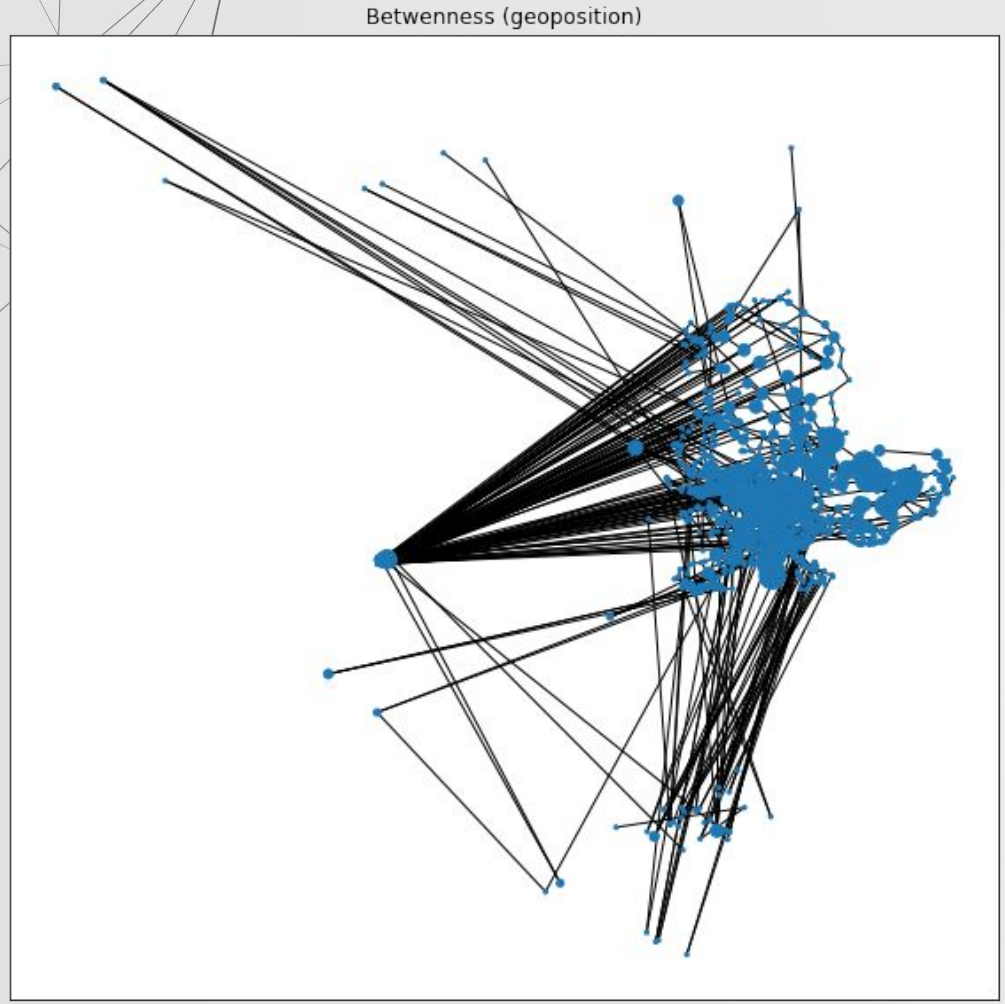
	Euroroad	arXiv astro-ph	Livemocha
Degree	0.0008	0.017	0.09
Betweenness	5.469	-	-
Closeness	3.236	-	-
Clustering Coefficient	0.016	4.061	239.934
Page Rank	0.21	8.344	103.78
Shortest Path	7.14	-	-
Diameter	0.0008	0.225	-
Connected Components	1.5e-05	1.597e-05	1.835e-05
Density	0.0004	0.0124	0.073

Euroroad

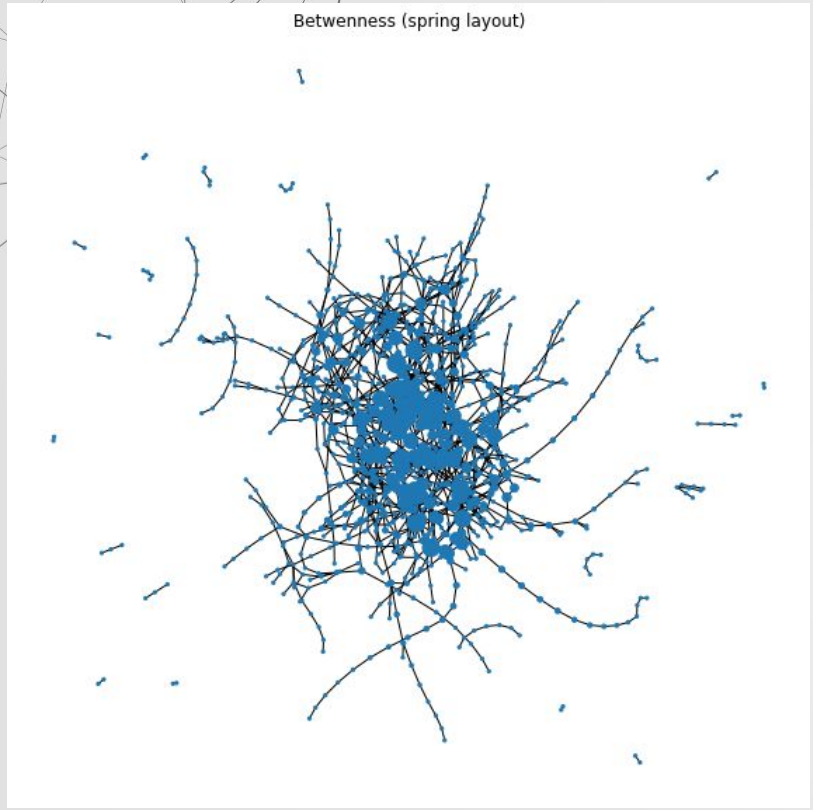

Degree Distribution
Count
Degree
0
100
200
300
400
500
600
2
4
6
8
10

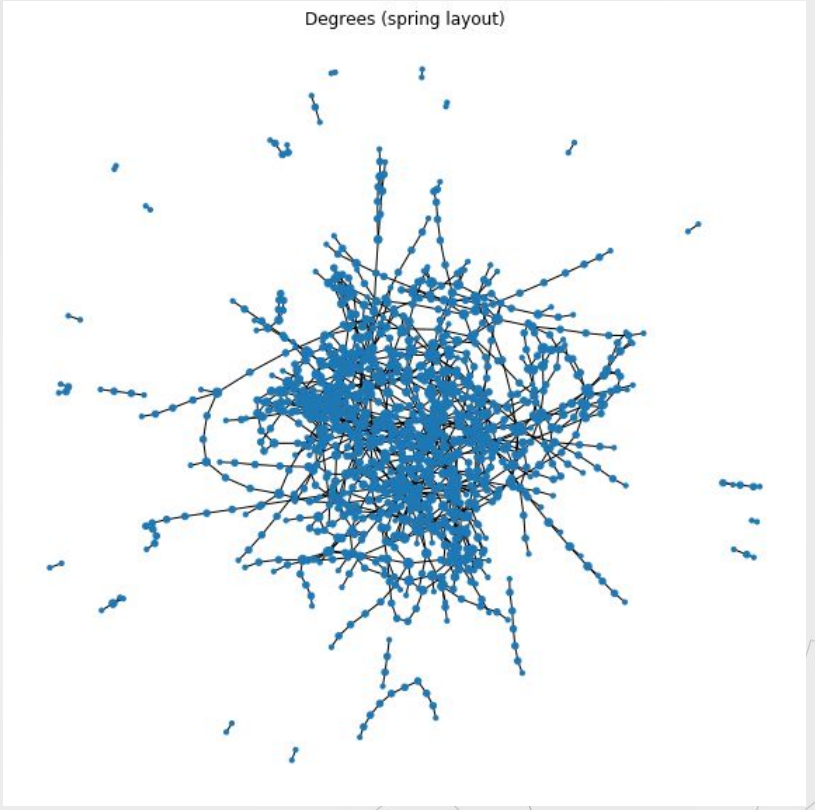

Clustering Distribution
Count
Coeffitient
0
200
400
600
800
1000
0.0
0.2
0.4
0.6
0.8
1.0


Shortest path length Distribution
Count
Shortest path length
0
20000
40000
60000
80000
100000
0
10
20
30
40
50
60



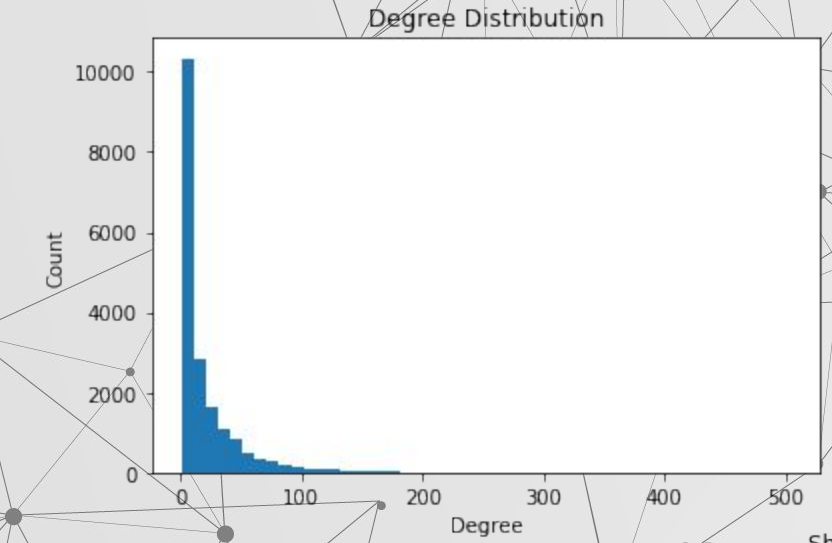
- Size of nodes related to betwennes
- Layout based on geoposition of city
- geocode.xyz API

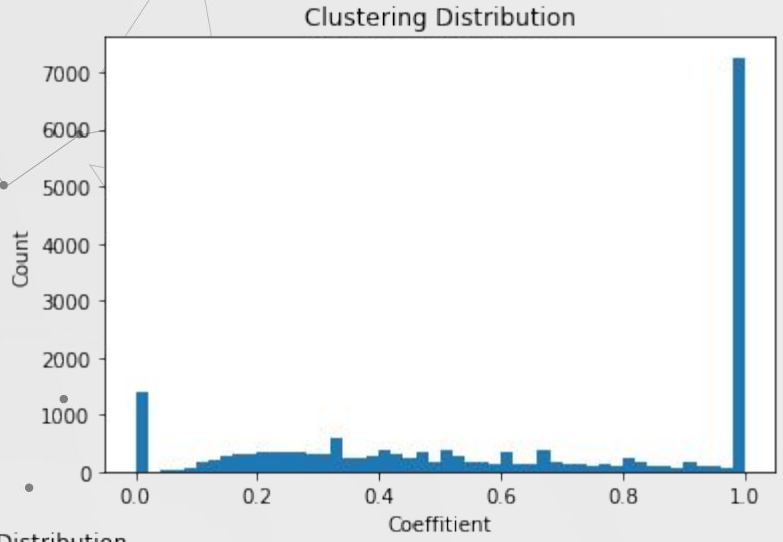


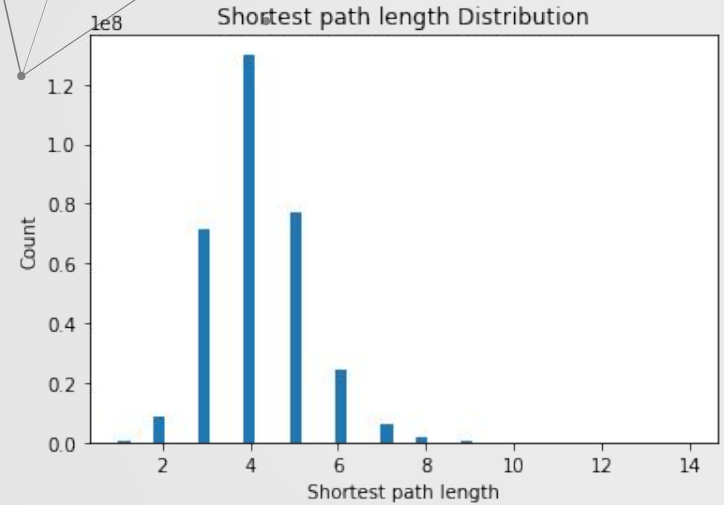


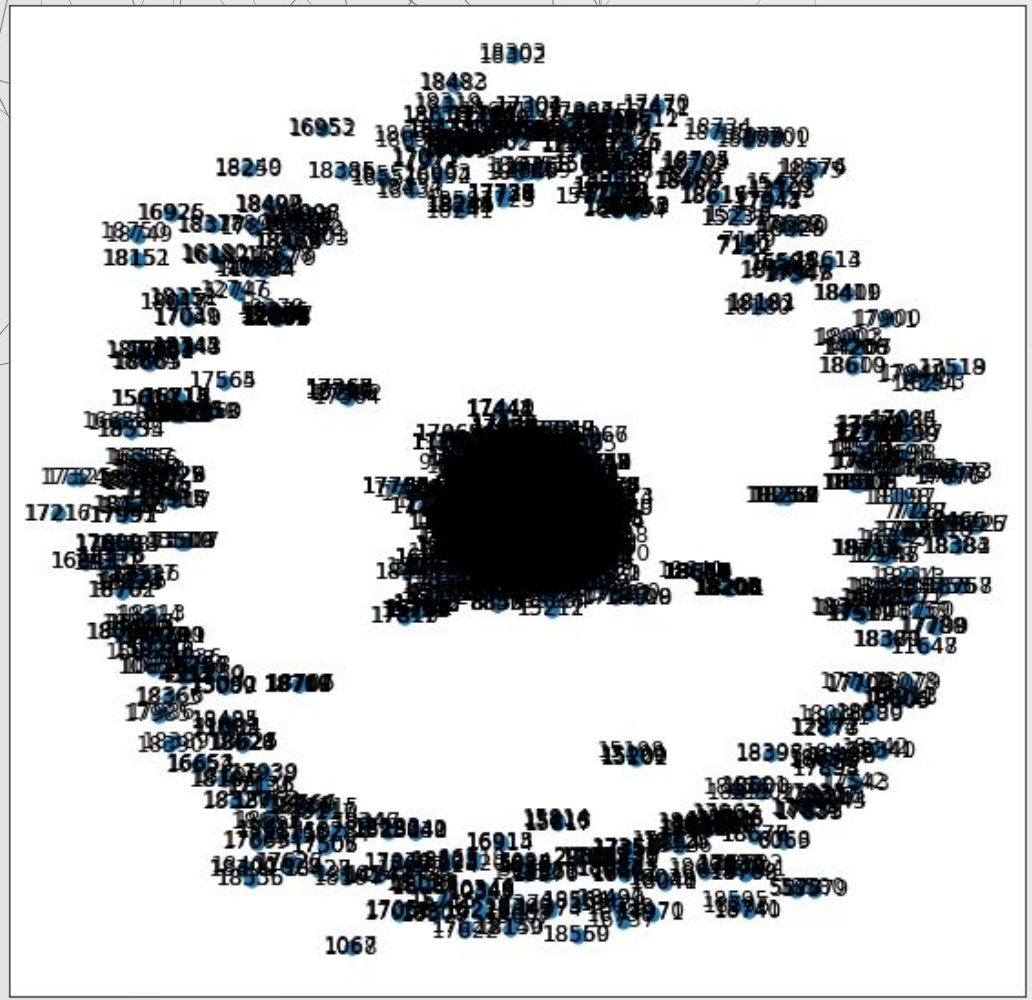
https://github.com/gephi/gephi/wiki/Fruchterman-Reingold

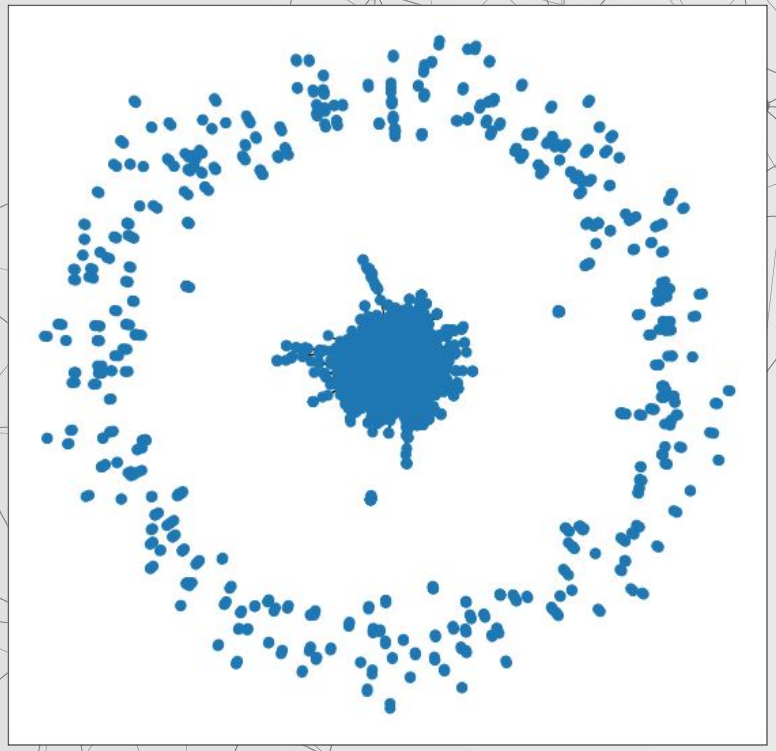
Astroph


Degree Distribution
Count
Degree
0
2000
4000
6000
8000
10000
0
100
200
300
400
500

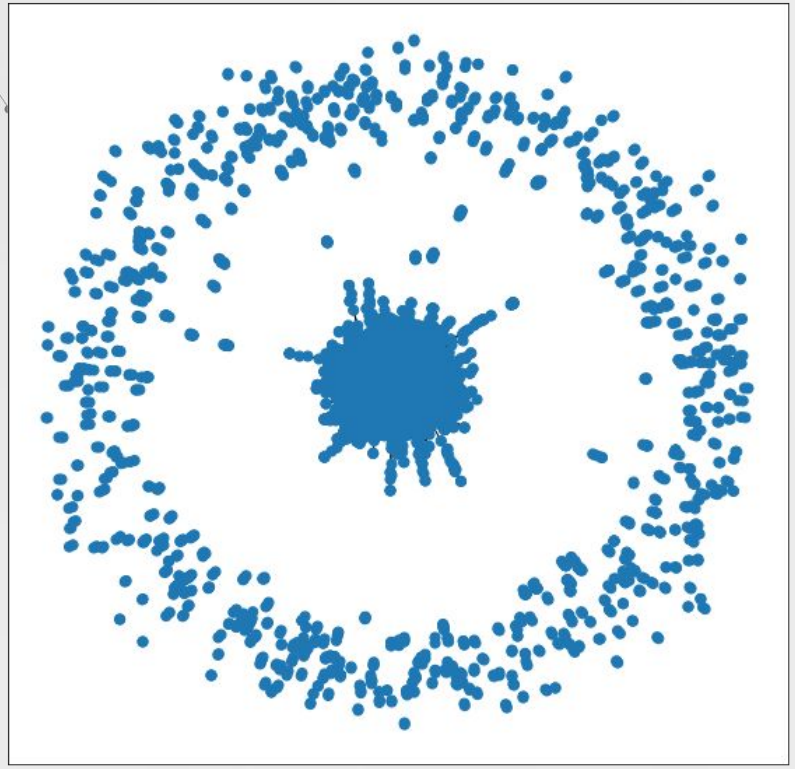

Clustering Distribution
Count
Coeffitient
0
1000
2000
3000
4000
5000
6000
7000
0.0
0.2
0.4
0.6
0.8
1.0


Shortest path length Distribution
1e8
Count
Shortest path length
0.0
0.2
0.4
0.6
0.8
1.0
1.2
2
4
6
8
10
12
14

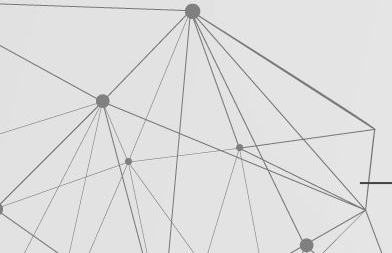




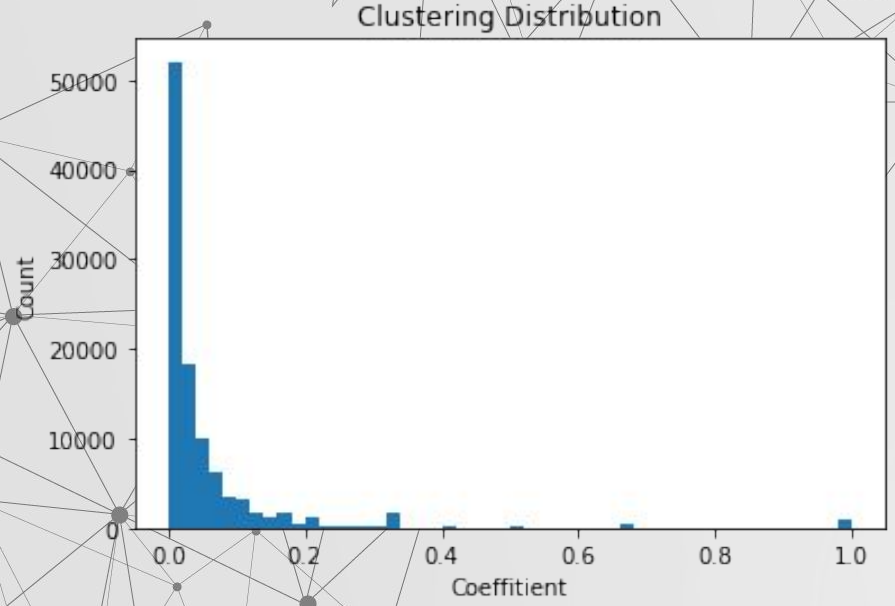
Degree based

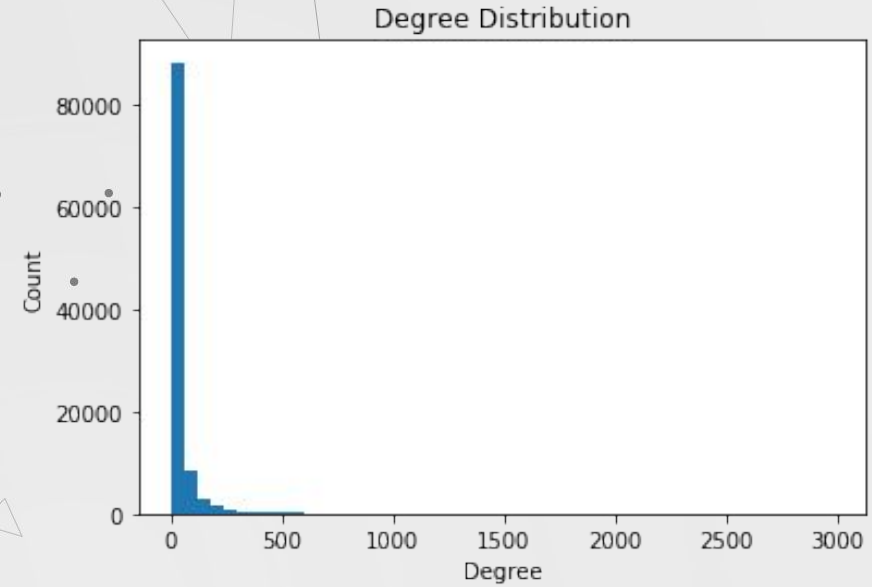


Clustering coefficient based



Livemocha


Clustering Distribution
Count
50000
40000
30000
20000
10000
0
0.0
0.2
0.4
0.6
0.8
1.0
Coeffitient


Degree Distribution
Count
80000
60000
40000
20000
0
0
500
1000
1500
2000
2500
3000
Degree

Advantages / Disadvantages

+ Easy access to data -> control over histograms

+ Library -> Process configurability

+ Control over network parameters (size and color of nodes, edges per instance)

+ A wide selection of analysis algorithms (which work slowly on large networks)

- In most cases, slow calculation of network statistics

- A small count of available layouts

- Slow drawing and no control over the process

- No own drawing libraries

THANKS

Does anyone have any questions?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Please keep this slide for attribution.