Computer Science and Engineering Department, SVNIT, Surat
Social Network Analysis

Assignment 1

U20CS005
BANSI MARAKANA

Choose a network of your choice from UCI network data repository.

1) Describe the dataset.

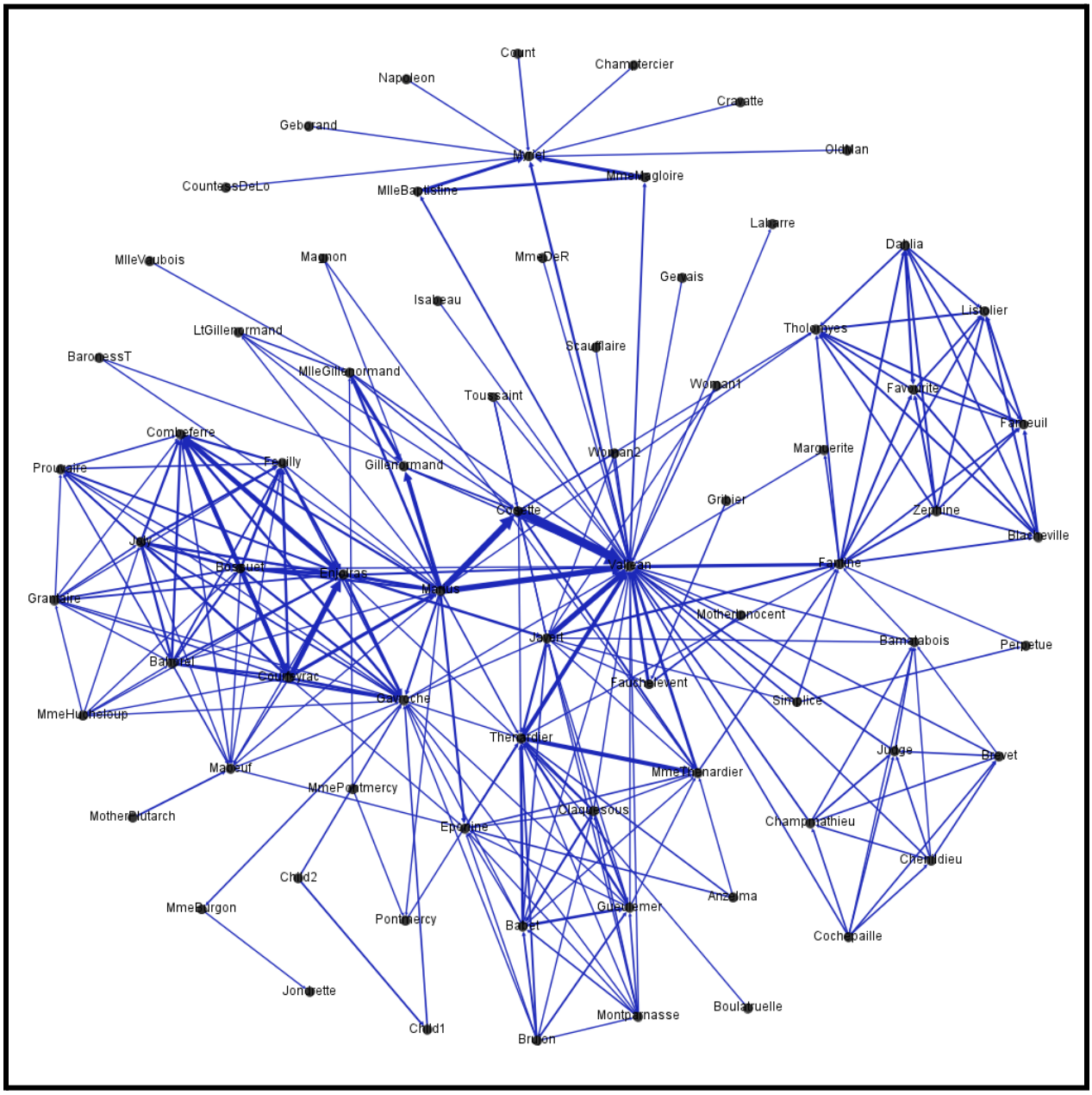
Dataset description:

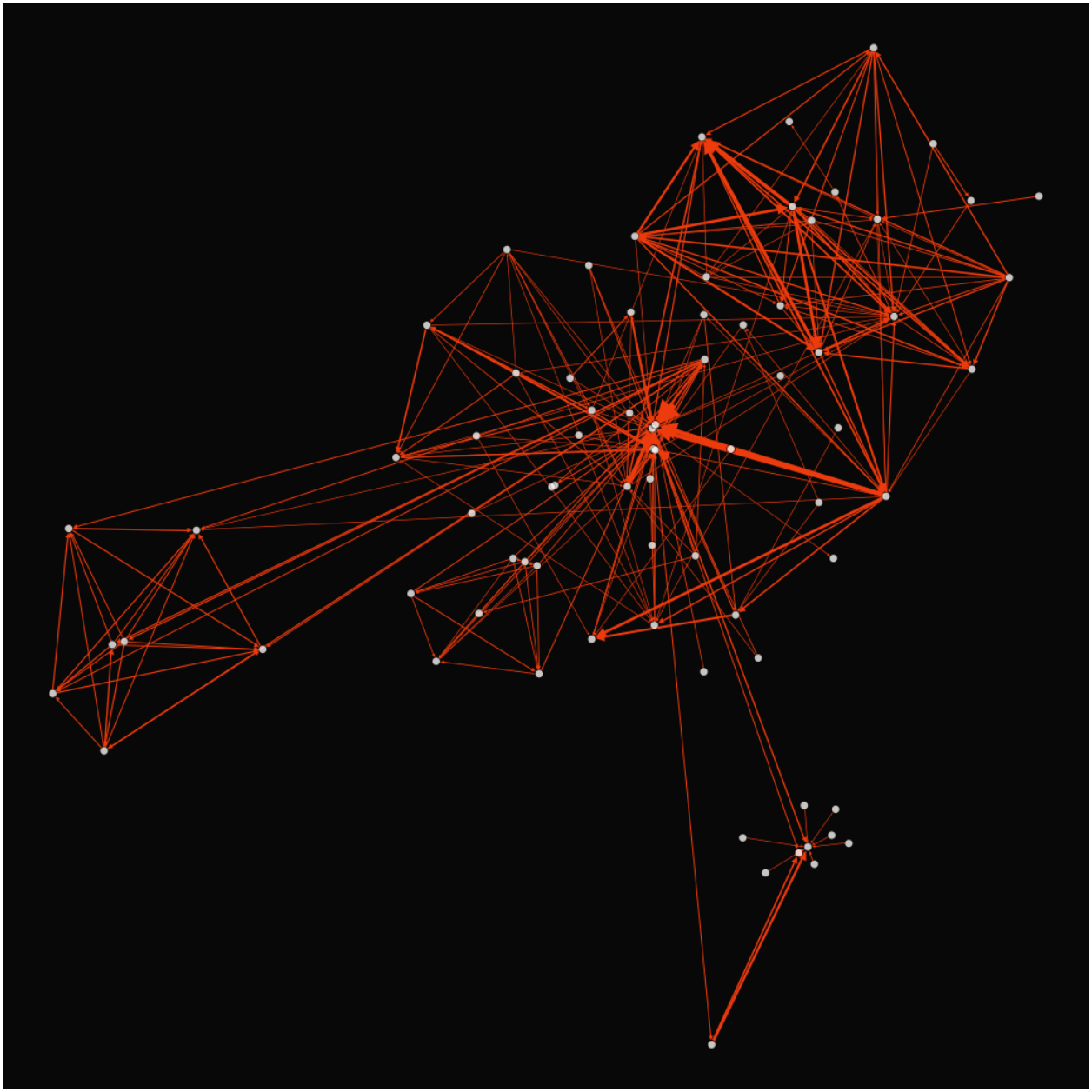
- I downloaded Les Miserables dataset from the github repository of Gephi which contains a co-appearance weighted network of characters in the novel Les Miserables. D. E. Knuth, The Stanford GraphBase.
- The dataset contains two spreadsheets of nodes and edges.
- The graph is weighted and directed.
- This dataset contains 76 nodes and 254 edges.

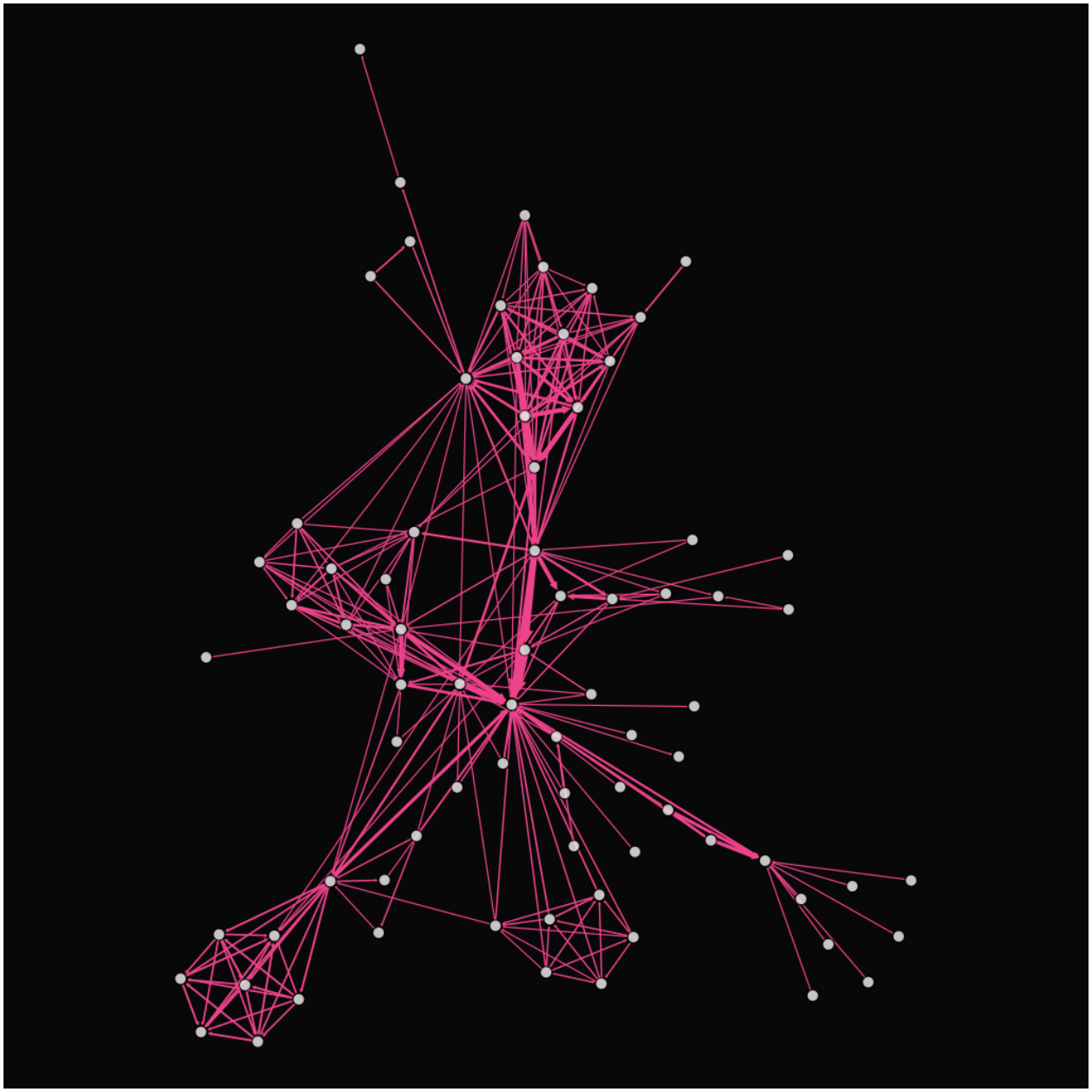
Gephi description:

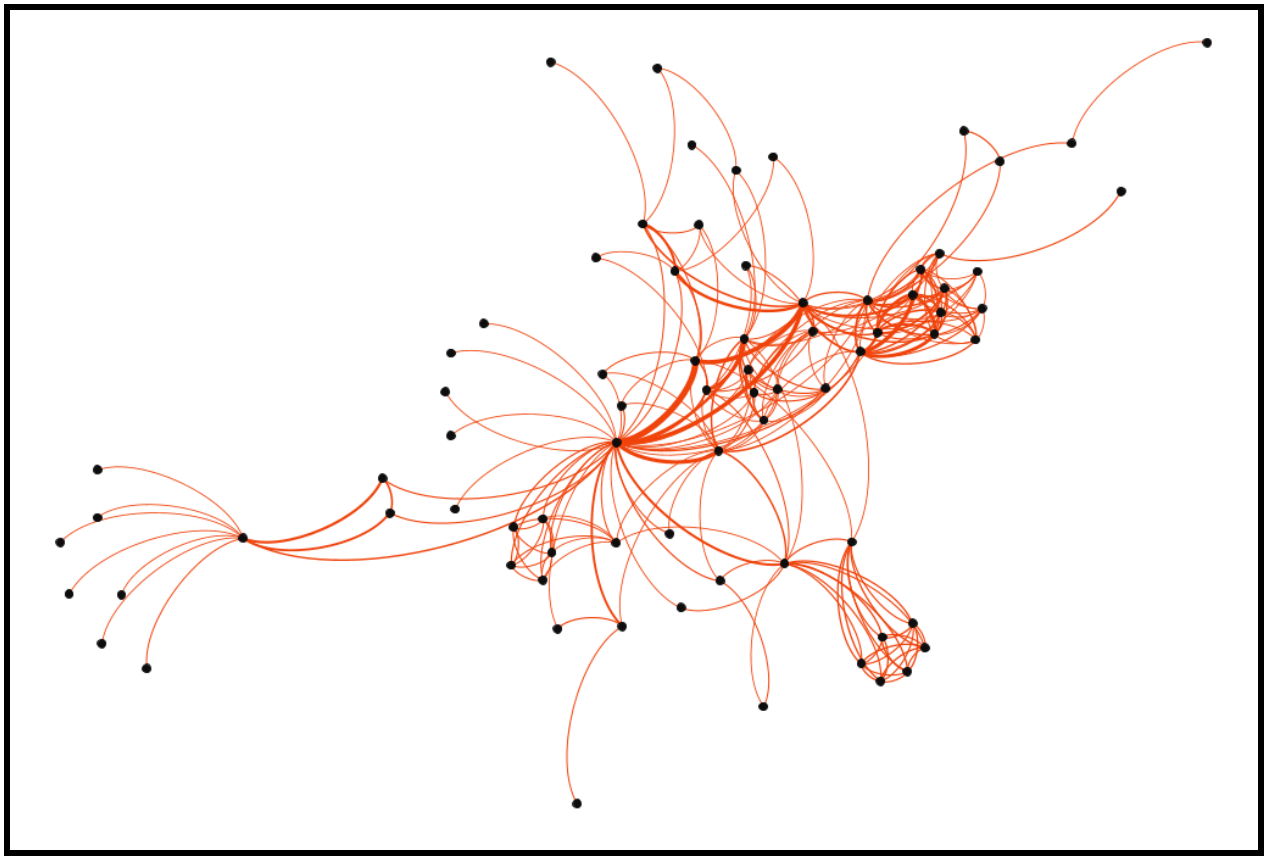
- Gephi is one of the most popular open-source software for network analysis.
- The point-and-click software can handle basic and advanced network analytics, and customizable graph visualizations. One can analyze network structures, communities and key actors, and can design network graphs on static and dynamic networks, geo-located data, as well as multimode/multiplex networks.
- However, Gephi is more limited analytically than software available in programming, such as R or Python.

2) Visualize it using Gephi.









3) Write a detailed report about your manual inspection about the structural property of the network.

The given network is directed network and it contains 77 nodes and 254 edges.

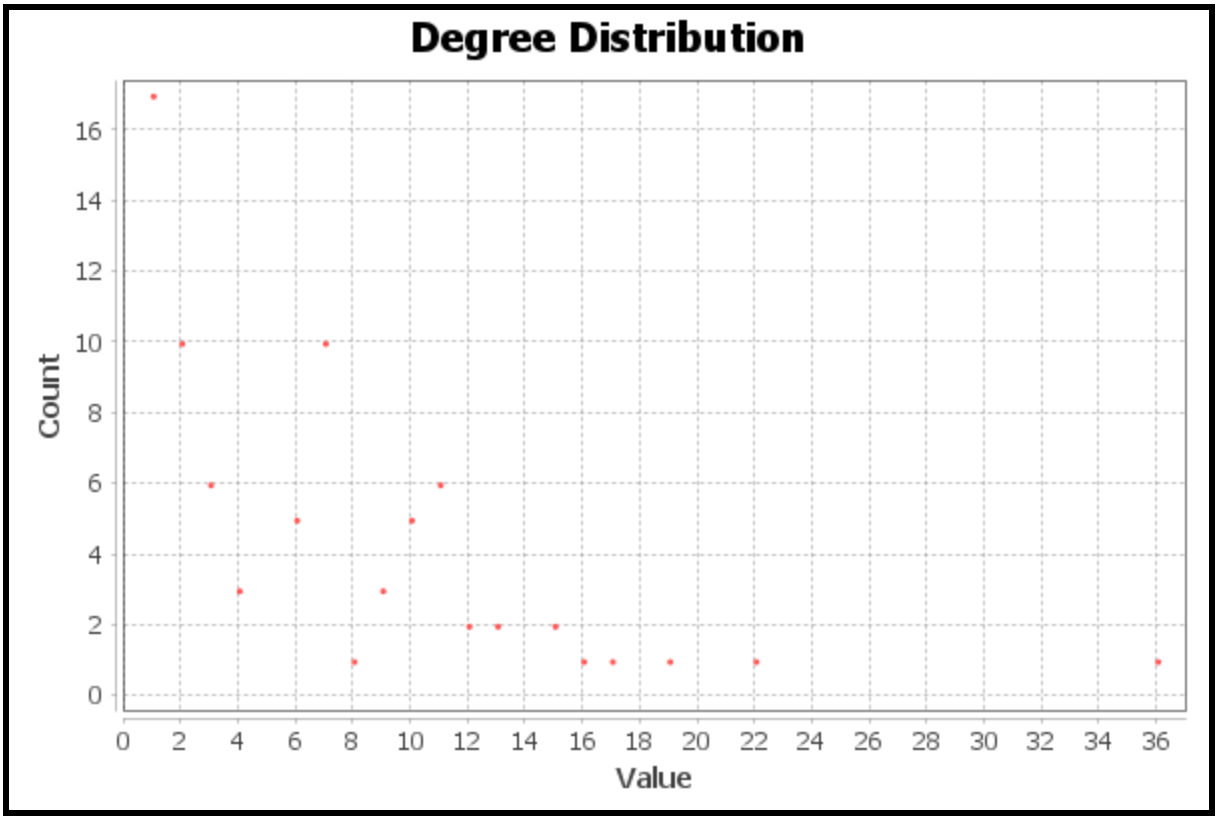
Nodes: 77

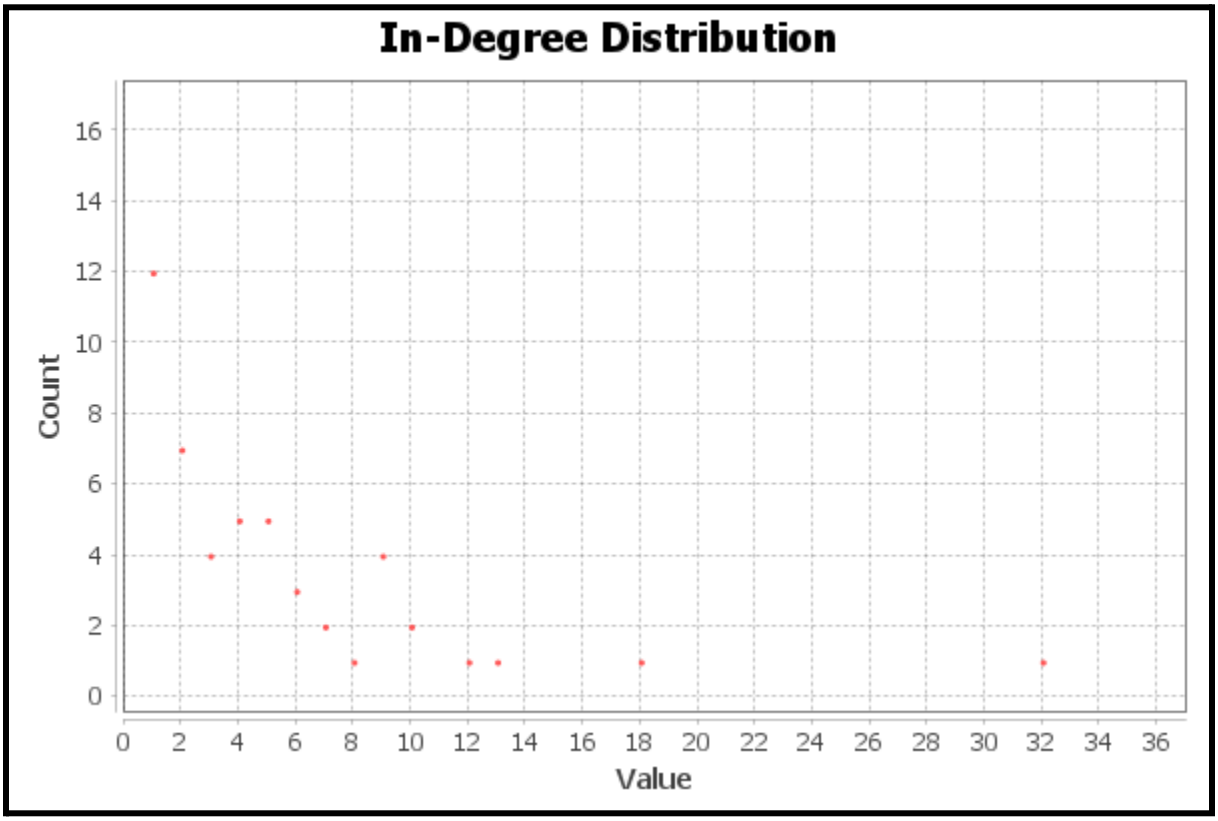
Edges: 254

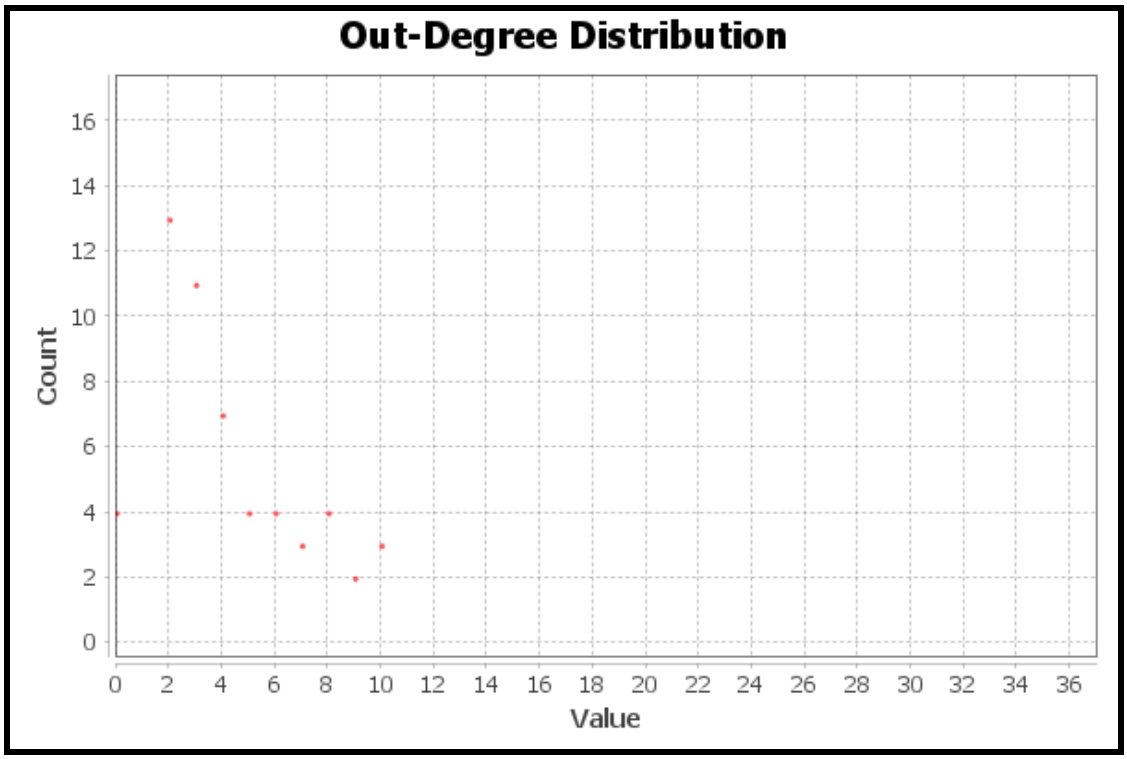
Directed Graph

Degree Distribution

- The degree of a vertex of a graph is the number of edges incident to the vertex.
- Degree distribution is a frequency count of the occurrence of each degree.
- Indegree of any node i is the number of nodes destined to i.
- Outdegree of any node i is the number of nodes originated at i.
- Average degree of this network is 3.3.

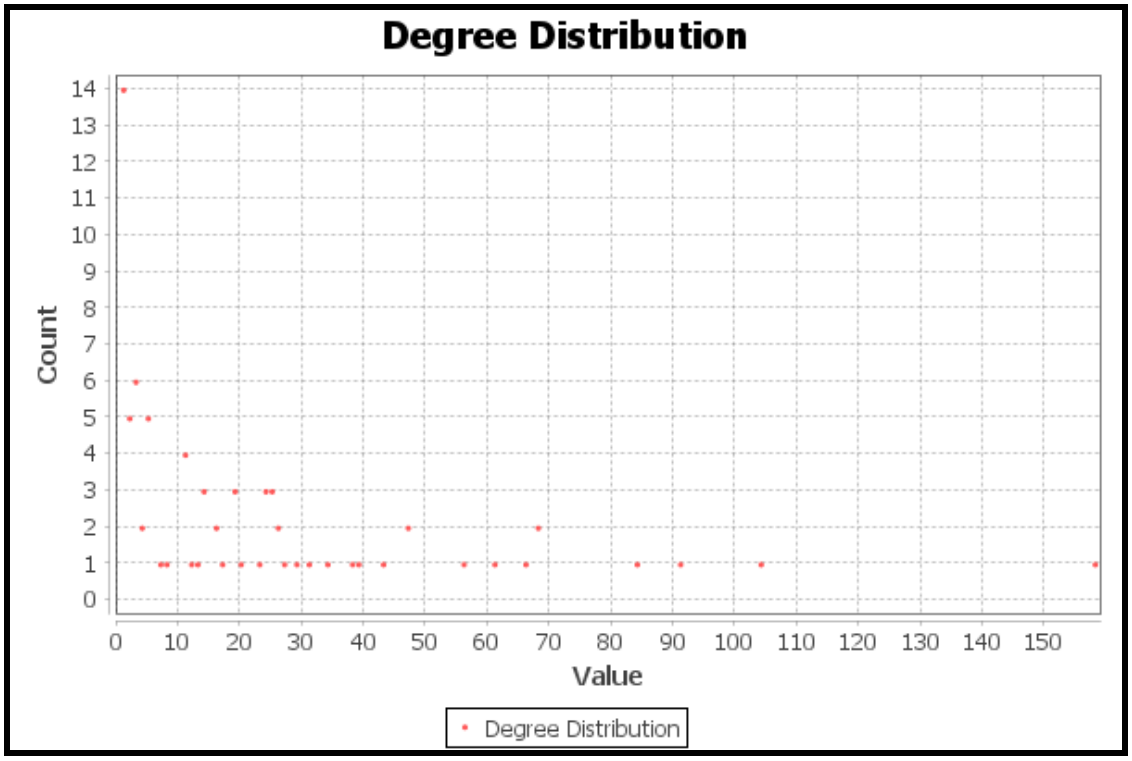

Degree Distribution
Count
Value
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36

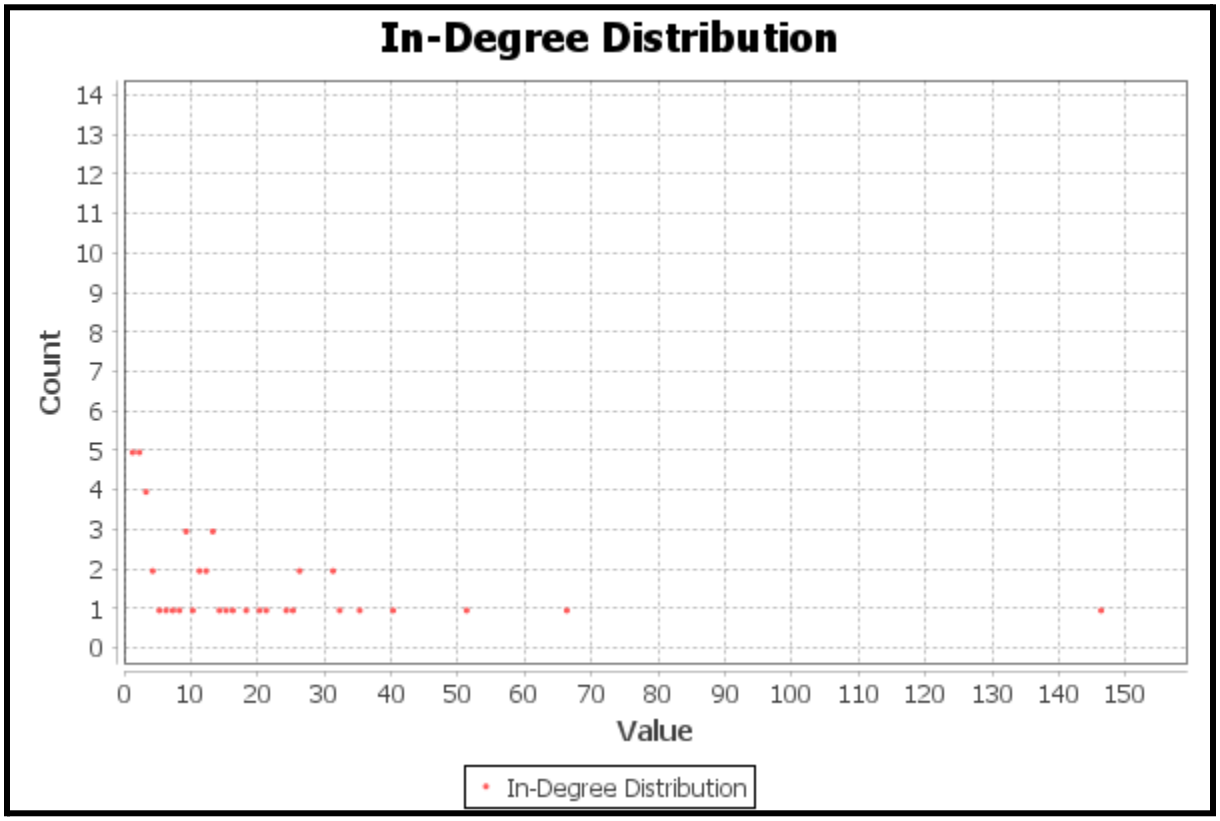

In-Degree Distribution
Count
Value
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36

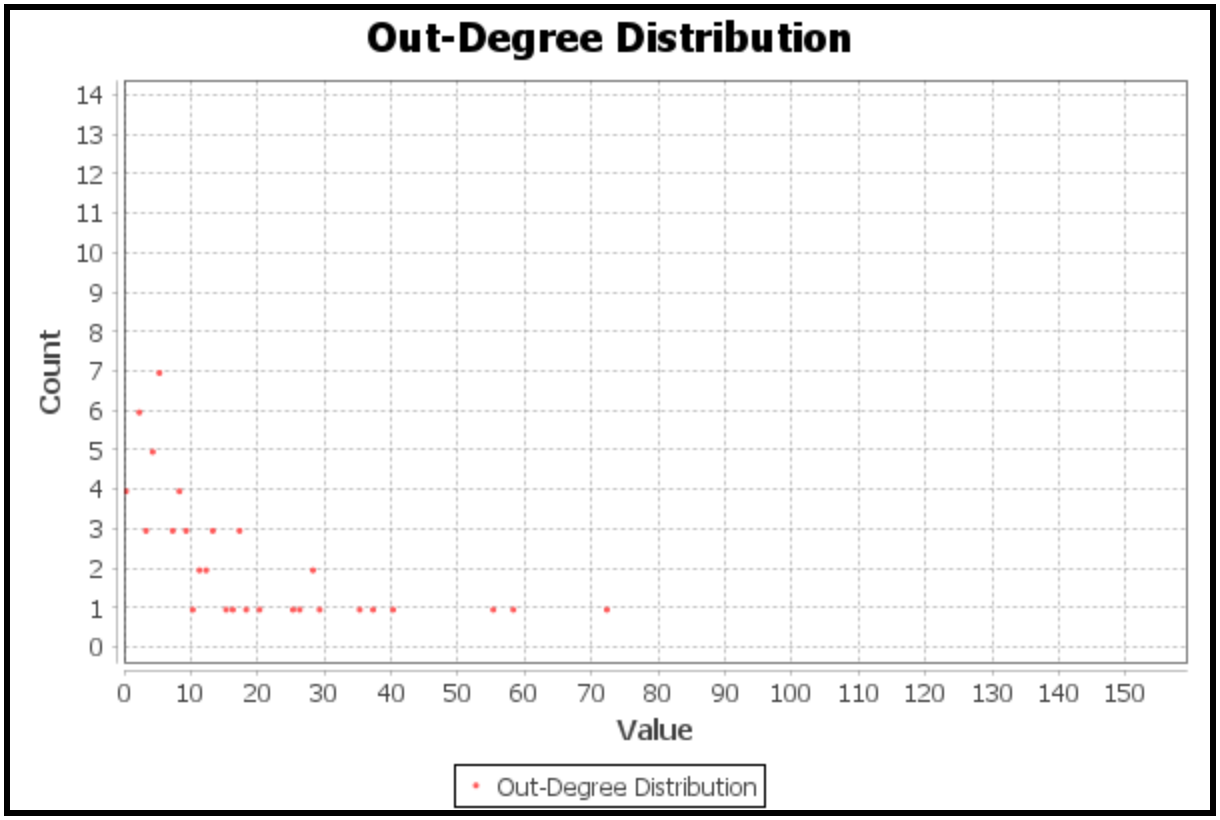


Weighted degree distribution:

- Average weighted degree of this network is 10.65.


Degree Distribution
Count
Value
Degree Distribution


In-Degree Distribution
Count
Value
In-Degree Distribution



Diameter of a network:

- It is the shortest distance between the two most distant nodes in the network.
- Diameter of this network is 5.

Density of a network:

- It measures how close the network is to completion.
- A complete graph has all possible edges and density equal to 1.
- Density of this network is 0.043.

Average path length:

- It is the average path distance between all pairs of nodes.
- Average path length is 2.4.