Gephi Report

Tag: #earthquake

Abstract:

To understand real life processes, we need to understand social behavior of other people using different techniques. Now a day a lot of people using social media platforms for different purposes like to spread information etc. Twitter is one of them. It allows their users to express their views and participate in discussions. To analyze the behavior, we need to analyze communication. Social Network Analysis is one of the best methods to analyze communications.

The purpose of this paper is how many users read and talk about Earthquake on twitter on earthquake hashtag in 2022 according to annotations, hashtags, replied to tweets, retweets and URL's.

In this research we tried to focus on analyzing these networks and extract important information form these networks. For dataset we get 220 nodes network of tweets. We analyze them using social network analysis methodologies like macroscopic level analysis, and also find important nodes using different centrality measures. These social network analysis measures can be used to analyze different kinds of discussions on twitter that can be helpful to understand people's behavior for multiple purposes in real life.

1. Introduction:

There are different kinds of tweets for earthquake hashtag in different countries. Every country has its own views for those country like Muzaffarabad Earthquake in 2005 at 8 October in Pakistan with the magnitude of 7.9 and over the 86000 of people death. Many peoples just view the tweets but some tweets are being watched by people with full attention. As a Human I'm want to see that how many people's discussed and tweeted about Earthquake.

An earthquake also known as a quake, tremor or temblor. An Earthquake is the shaking of the surface of the Earth resulting from a sudden release of energy in the Earth's lithosphere that creates seismic waves. Earthquakes can range in intensity, from those that are so weak that they cannot be felt, to those violent enough to propel objects and people into the air, damage critical infrastructure, and wreak destruction across entire cities. The seismic activity of an area is the frequency, type, and size of earthquakes experienced over a particular time period.

The seismicity at a particular location in the Earth is the average rate of seismic energy release per unit volume. The word *tremor* is also used for non-earthquake seismic rumbling.

Earthquake Fault Types

There are three main types of fault, all of which may cause an interpolate earthquake: **normal, reverse (thrust), and strike-slip**. Normal and reverse faulting are examples of dip-slip, where the displacement along the fault is in the direction of dip and where movement on them involves a vertical component. Many earthquakes are caused by movement on faults that have components of both dip-slip and strike-slip, this is known as oblique slip.

We need to understand social behavior of other peoples about natural disasters using different techniques. Now a day a lot of people are using social media platforms for different purposes like to spread information etc.

Twitter is one of them. It allows their users to express their views and participate in discussions. To analyze the behavior. we need to analyze tweets and retweets. Twitter is the best platform that provides facility to debate on different topics. We adopt a method and Social Network Analysis is one of the best methods to analyze communications on Twitter. By using this method. We can analyze human's behavior by different perspectives of natural disasters and climate change problems. Climate change, globalization and urbanization are driving factors behind flood events and their consequences

We can detect news, Earthquake Monitor, FM Global and WHO groups (World Health Organization) on Twitters. Similarly, by aiming this method we can predict earthquake risk. As we know that Twitter is one of the famous platforms in the world. We can find user's interest and use it for different purpose like news about earthquakes, natural disasters and environment etc. We can find top hashtags, top tweets, and users with maximum tweets with this method, users who replied to tweets related to this, users that get traffic by using this hashtag and so on. This paper is related to Pakistan hashtag tweets. In this paper I tried to analyze networks related to annotations, hashtags, retweets, replied on tweets and URL. By this analysis, I found most important hashtags, annotations. Relation between different tweets by using retweets network. Various communities are detected and analyze whole network. For this analysis we used Gephi Software. It's a very famous software about analysis Gephi has been used in a number of research projects in academia, journalism and elsewhere, for instance in visualizing the global connectivity.

2. Related Work:

Social media sites are designed to develop social relationships. Even people who know each other in real life can make virtual friends on social networks. People can share different types of contents, and others can interact with it. In recent years, Researchers and Businesses Companies have analyzed social media to better understand user behavior. In this section, I want to draw attention to the growing interest in social network analytics. The objective of this article, Analysis of the similarity of social networks: an example on Twitter.

3. Method:

We used Gephi Software. It's a very famous software about analysis Gephi has been used in a number of research projects in academia, journalism and elsewhere, for instance in visualizing the global connectivity. Gephi software with twitter API tool to extract tweets related with our hashtag and then analyzed them in terms of different factors. Like closeness centrality, in degree and out degree centrality and their weights. We assign colors to different nodes and edges according to our need. We took a sample of 200 plus nodes network on which we performed our analysis.

4. Evaluation and results:

From the above findings, the sentimental model gives accurate results. Outcomes of these models were compared, the model related to tweets and the hybrid model give accurate results but the model related to natural parameter gives less accurate results. According to this research the best one classifier is Logical regression for tweets related model, model related to natural parameters. Degree centrality and betweeness centrality measures give maximum accurate results, according to the logistic models that have been developed. In this way we were able find results using network analysis.

5. Datasets:

For evaluation the data is gathered from twitter. This dataset is related to Pakistan tweets. There are seven types of different networks extracted from tweets related to Pakistan like annotations (undirected network). Hashtags (directed and undirected networks), mentions (directed network), replies to (directed network), retweets (directed network), URLs (undirected network). Each network has different kind of dataset. This dataset contains source and target with weight attribute that shows how many times linked edges generated between source and target. Networks are provided in the form of edge list and extracted from tweets that were gathered by using Twitter Search API V2 endpoint. These are directed and un-
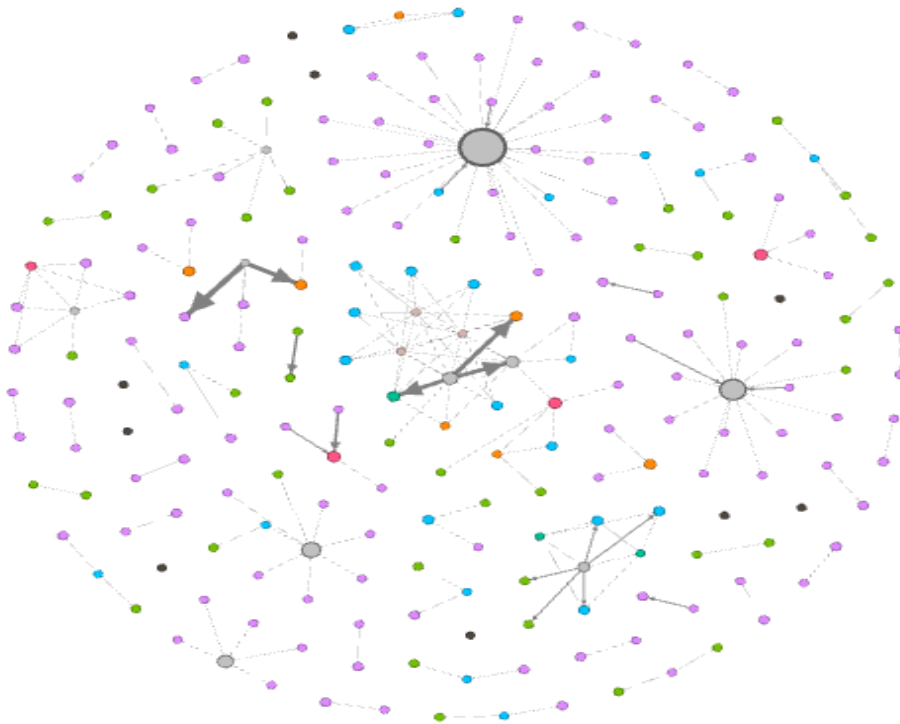
directed networks of annotations, hashtags, mentions, replies, retweets, URLs. These networks are extracted from the Tweets given in json file.
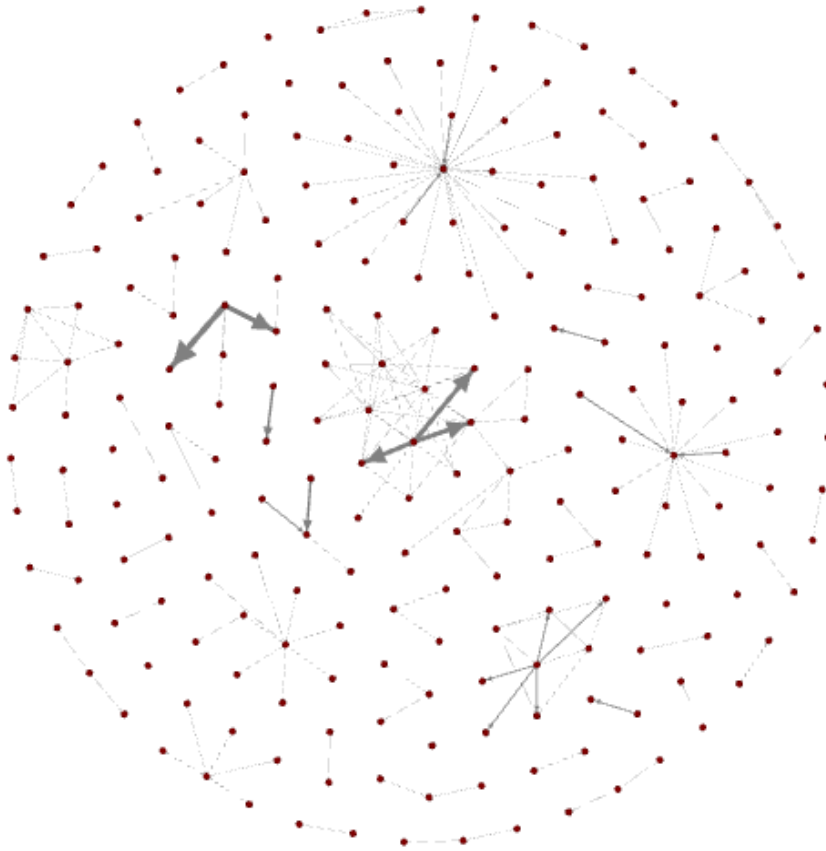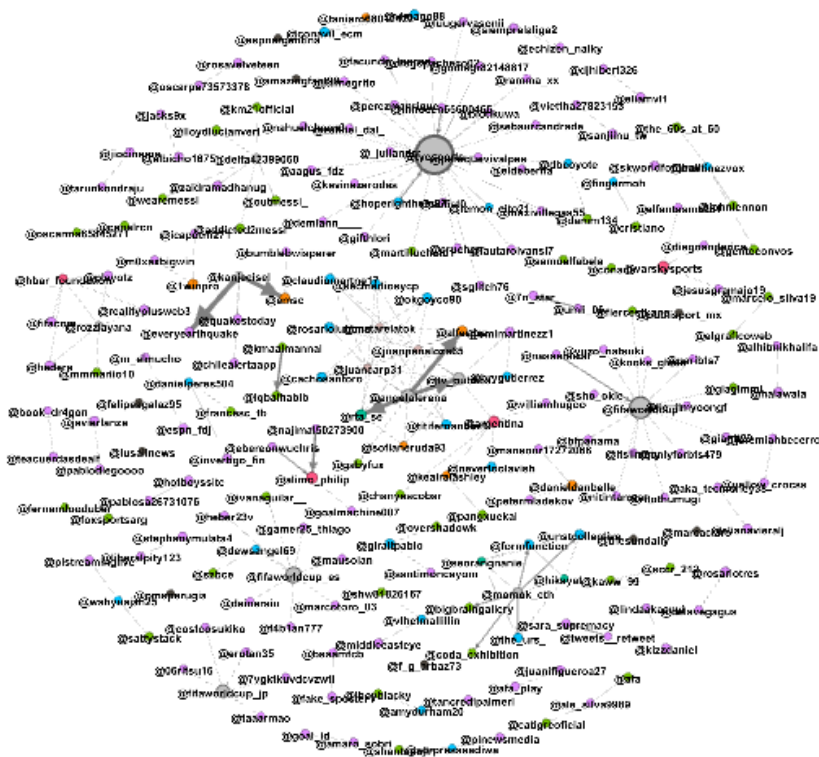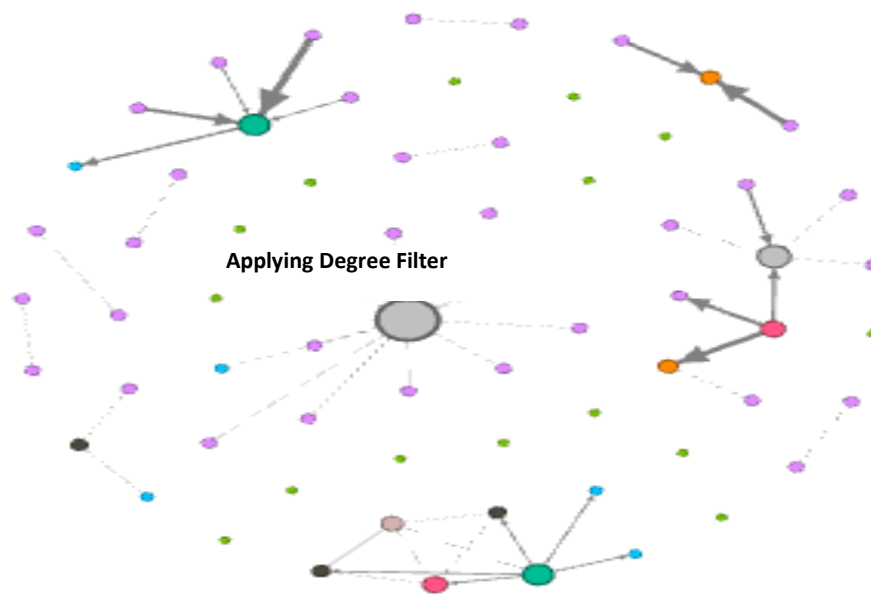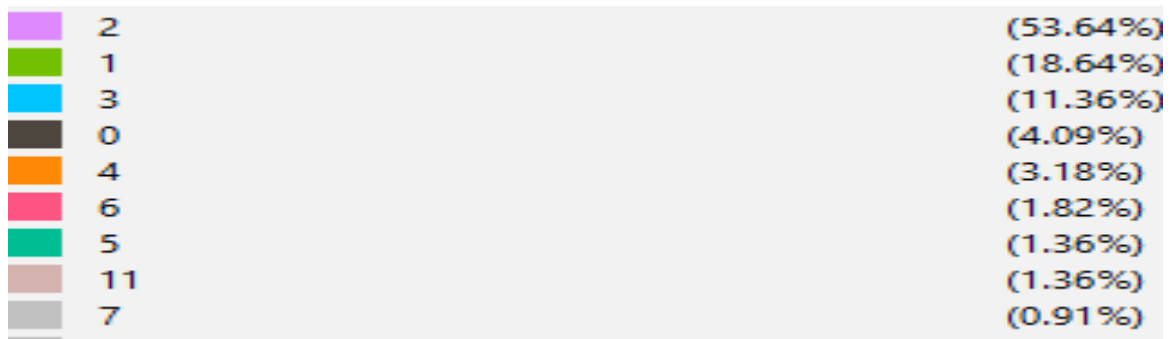

Hashing Network Diagram

Without Label Graph

With Labels Graph

2		(53.64%)
1		(18.64%)
3		(11.36%)
0		(4.09%)
4		(3.18%)
6		(1.82%)
5		(1.36%)
11		(1.36%)
7		(0.91%)

Applying Degree Filter

After Using User Network in Twitter Stream Importer

6. Macroscopic Analysis:

Our results for Directed Graph after applying different filters:

Tests	Values/ Results
Density	0.006

Degree Distribution (Average Degree)	1.391
Average Clustering Coefficient	0.035
Average Weighted Degree	1.632
Eigenvector Centrality Number of iterations	150
Modularity	0.919
Network Interpretation Diameter	3
Number of Weakly Connected Components	59
Number of Strongly Connected Components	220
Parameters Epsilon	0.001
Parameters Probability	0.85
Description Length	1467.858

In-Degree Distribution
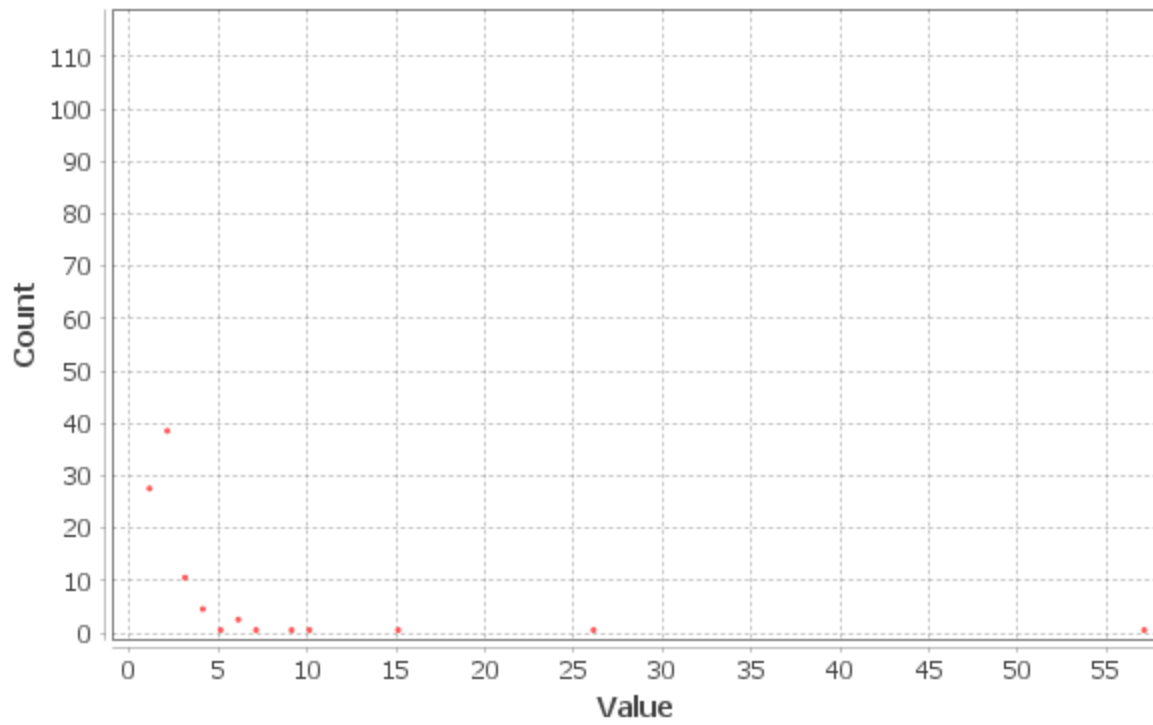
Degree Distribution
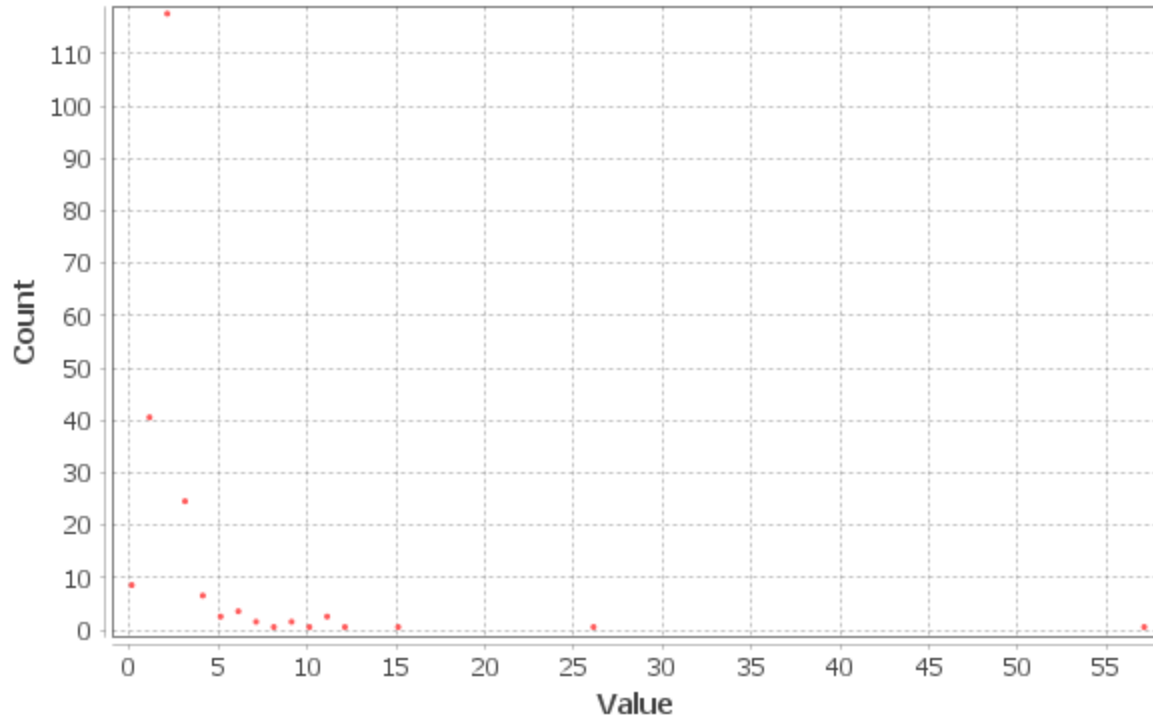
Out-Degree Distribution
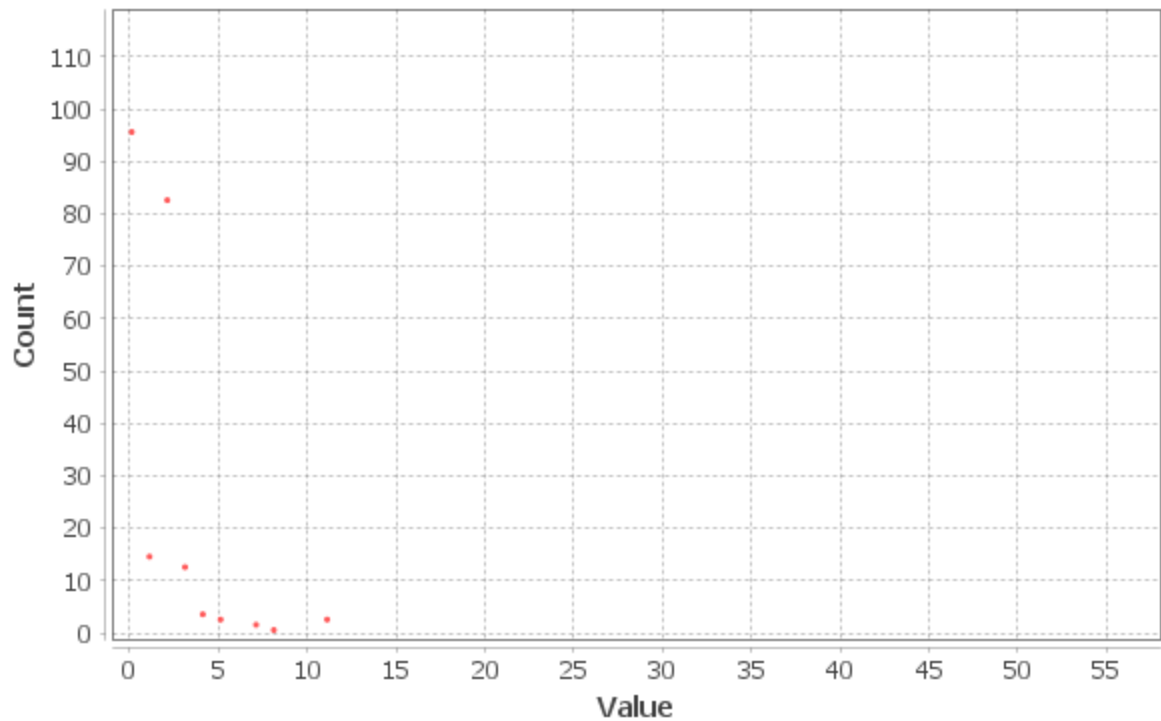
Clustering Coefficient Distribution

Out-Degree Distribution

Out-Degree Distribution

In-Degree Distribution

In-Degree Distribution

Degree Distribution
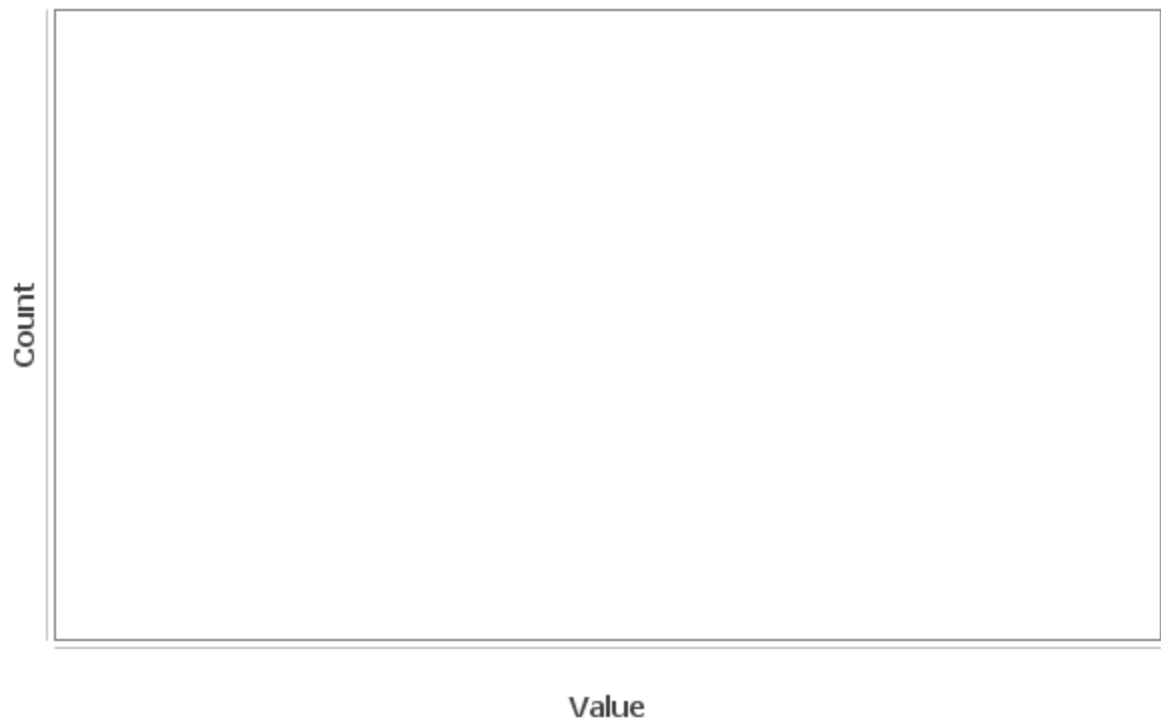
- Degree Distribution
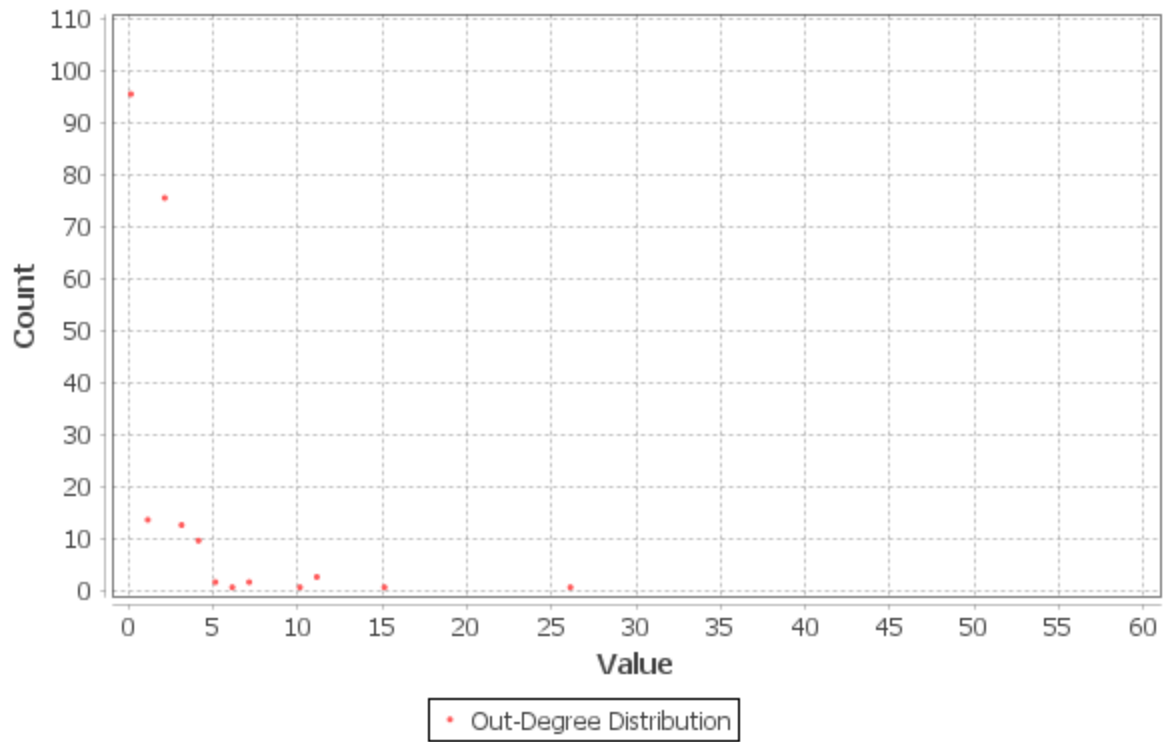
Eigenvector Centrality Distribution
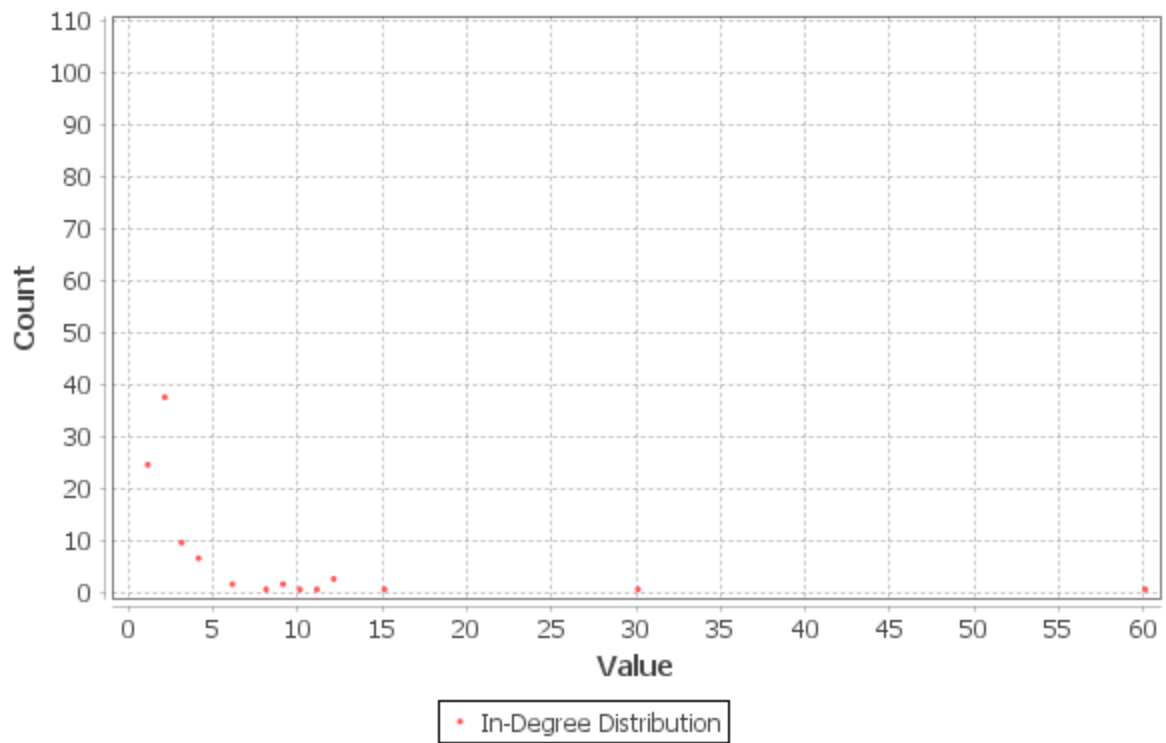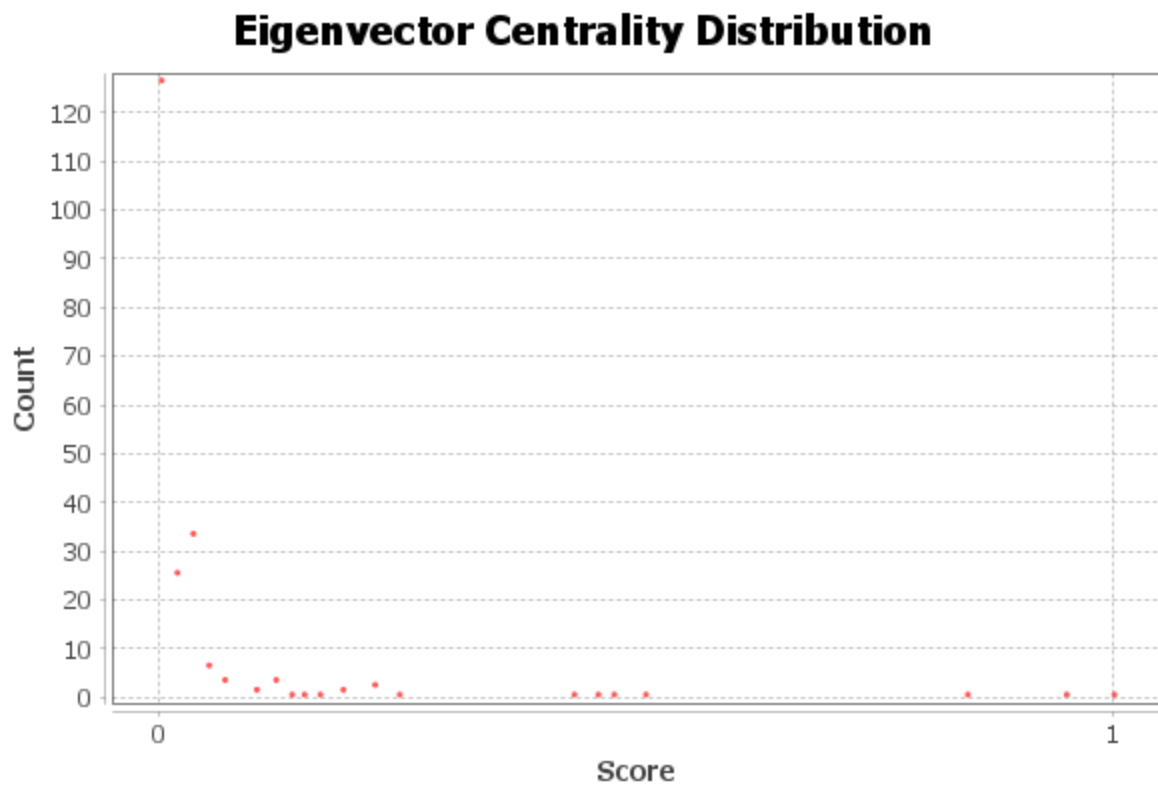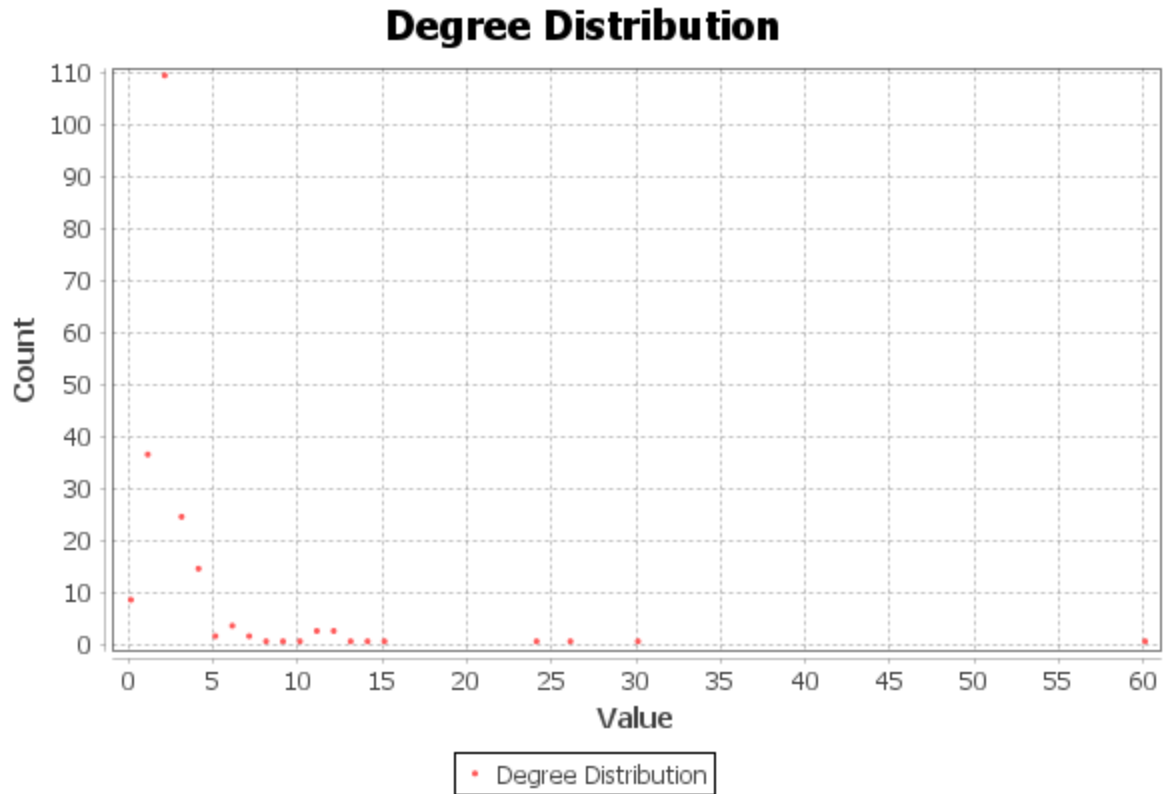
Graph Density Report

Parameters:

Network Interpretation: directed
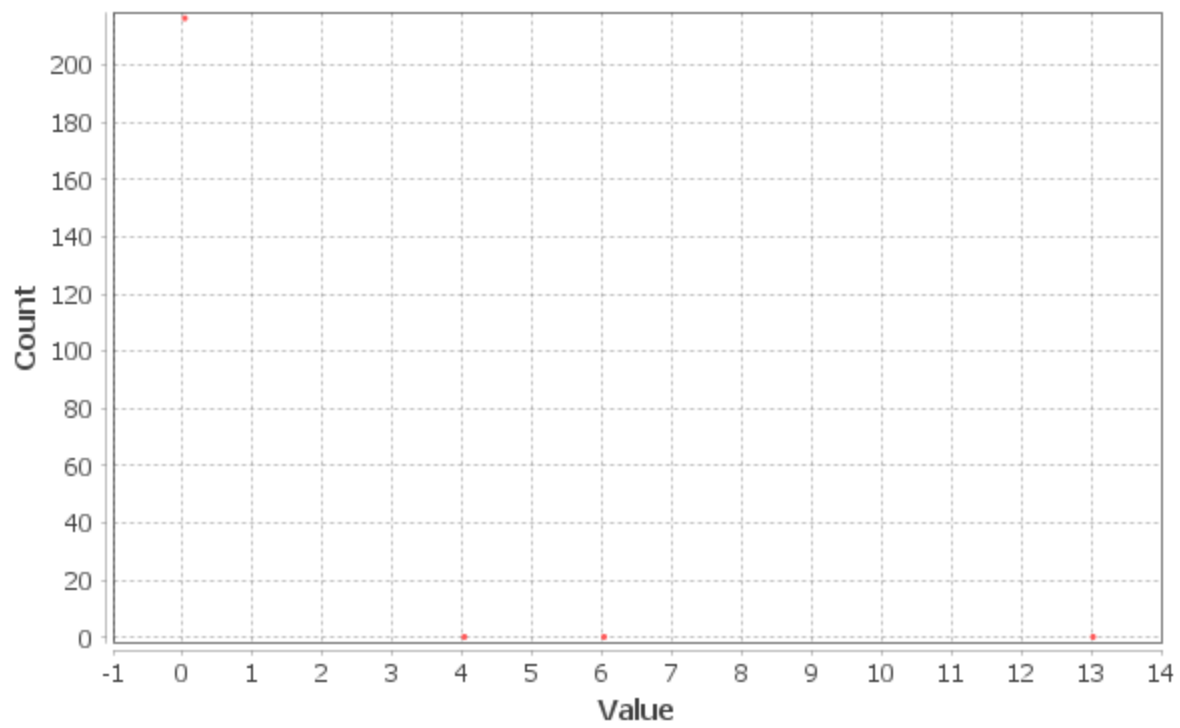
Results:

Density: 0.006

Size Distribution
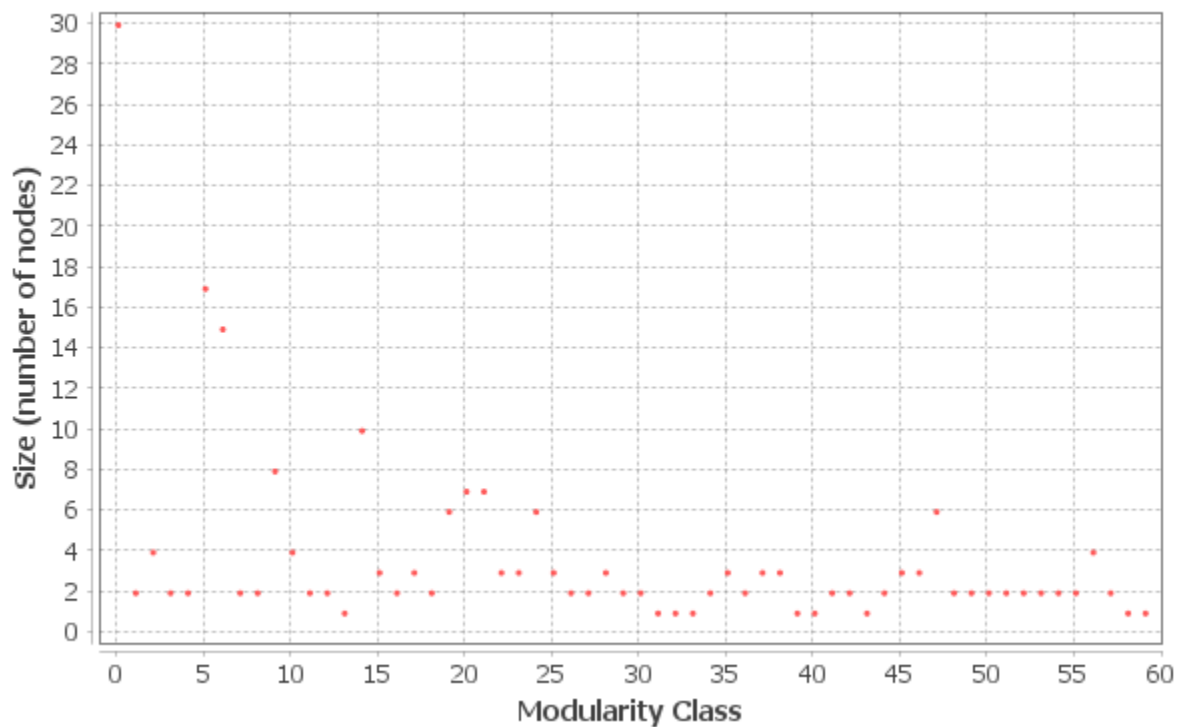
Eccentricity Distribution
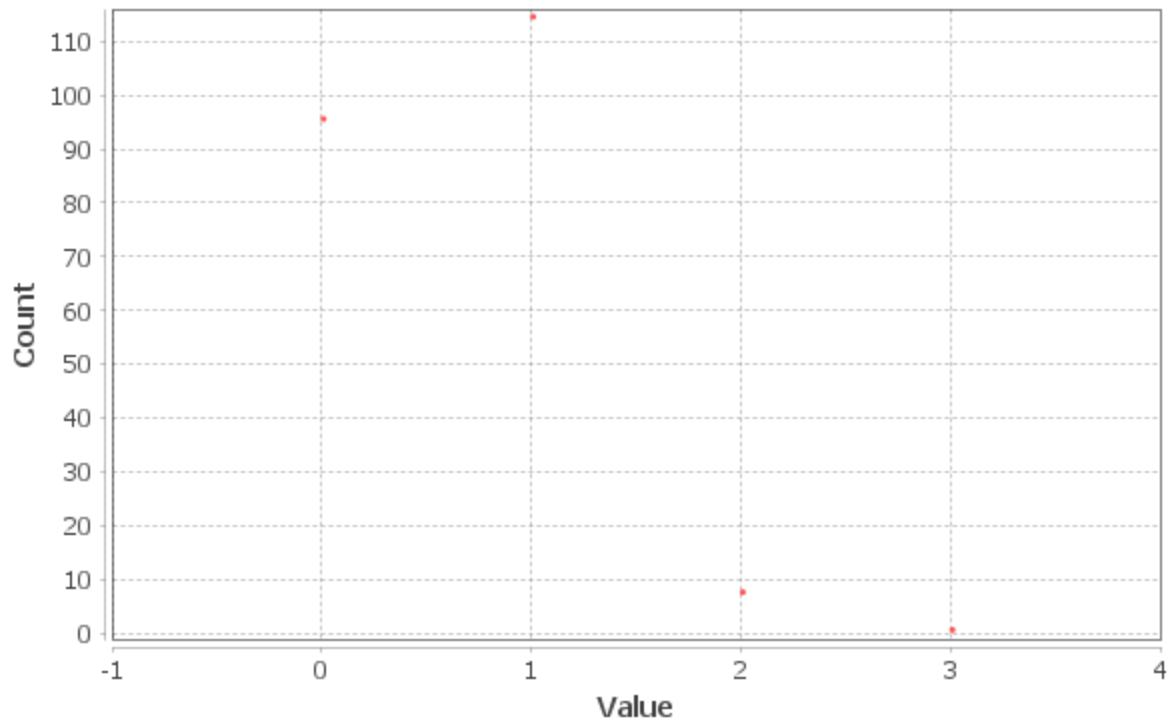
Closeness Centrality Distribution
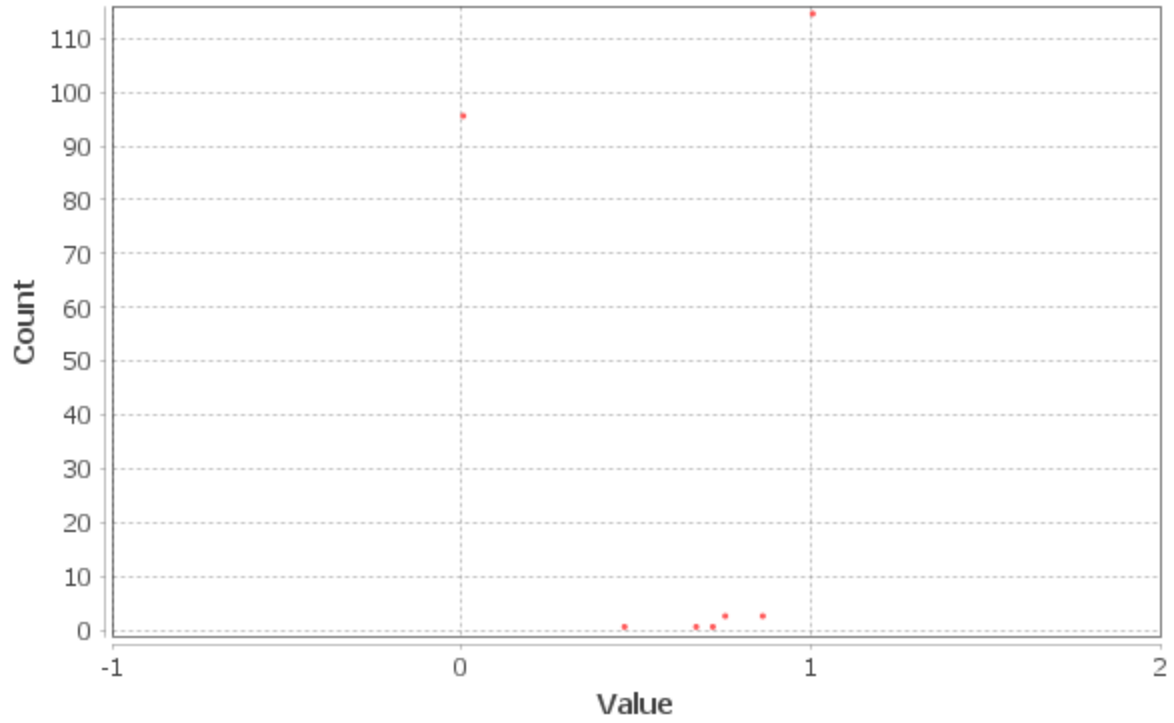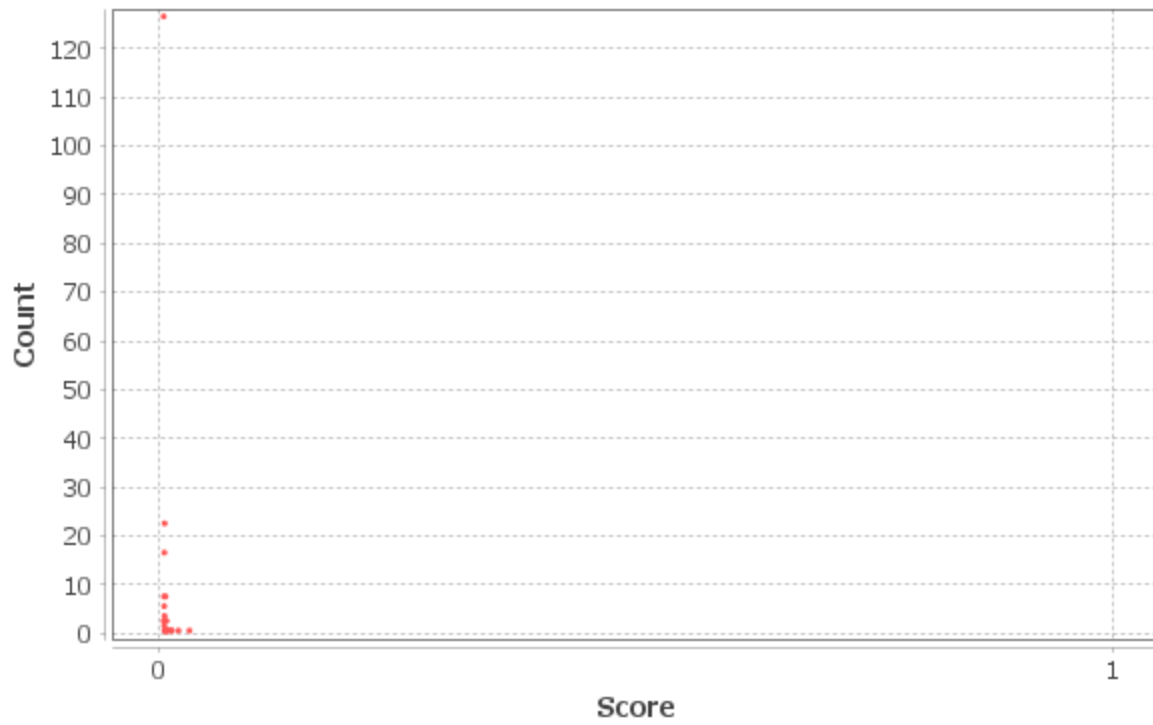
Betweenness Centrality Distribution

Harmonic Closeness Centrality Distribution
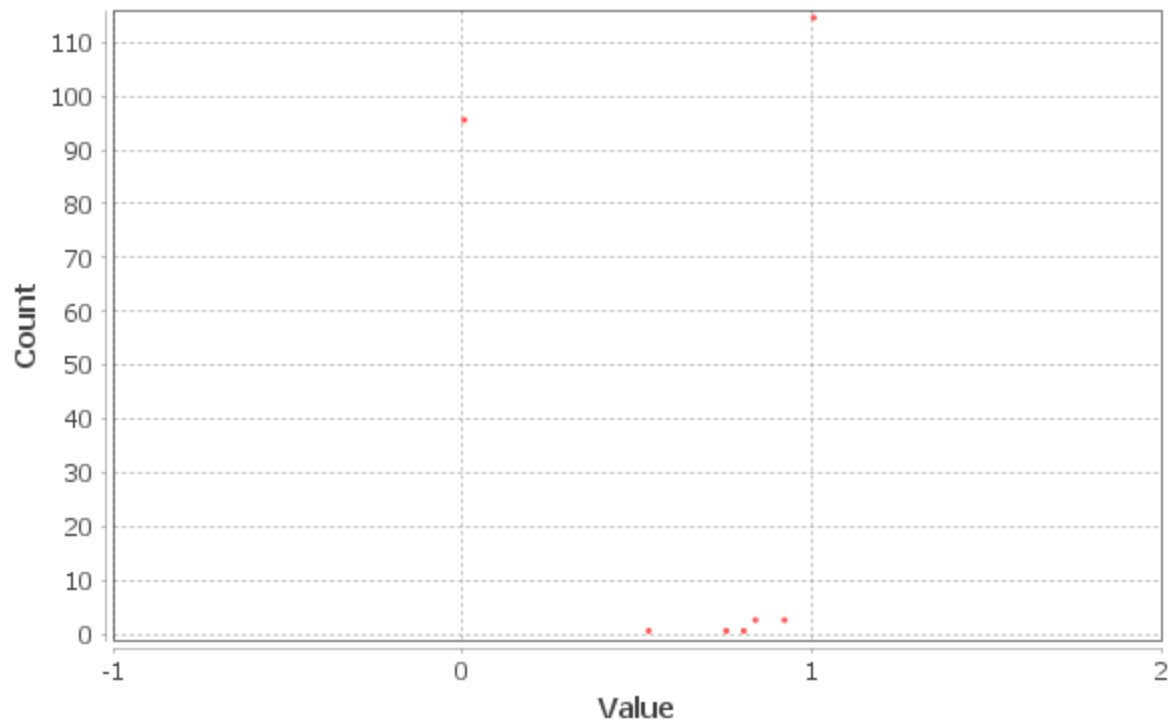
Size Distribution
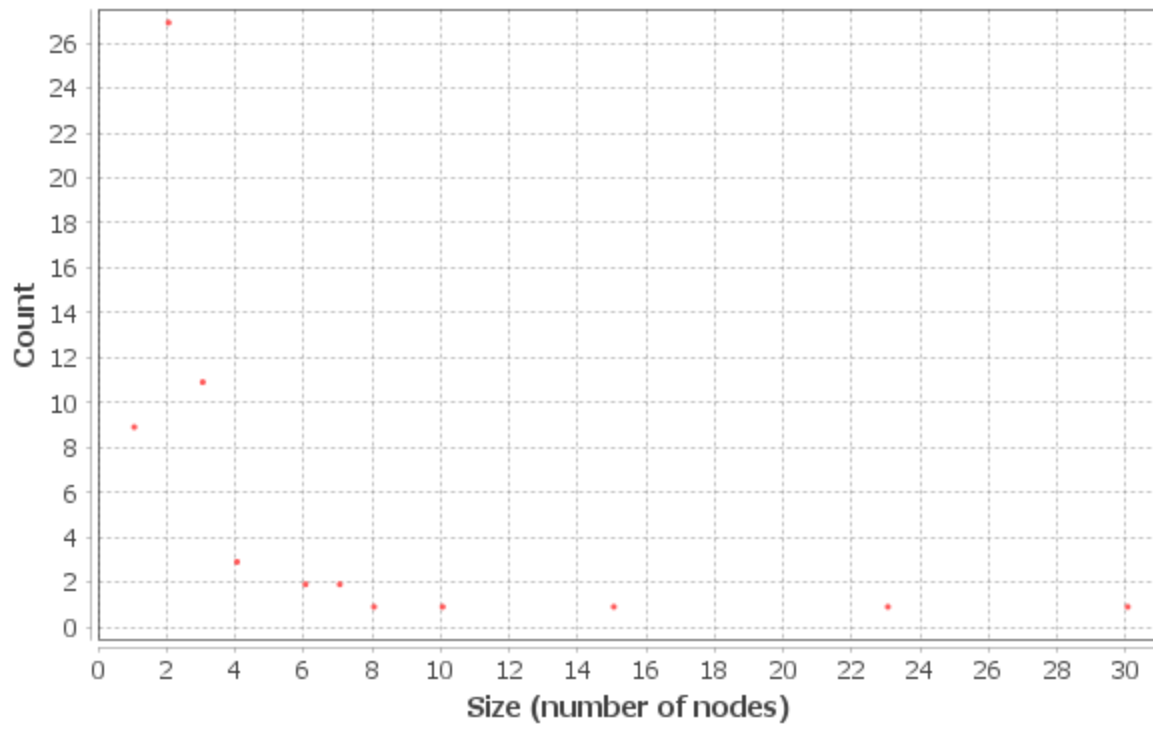
PageRank Distribution

Size Distribution

7. Eccentricity:

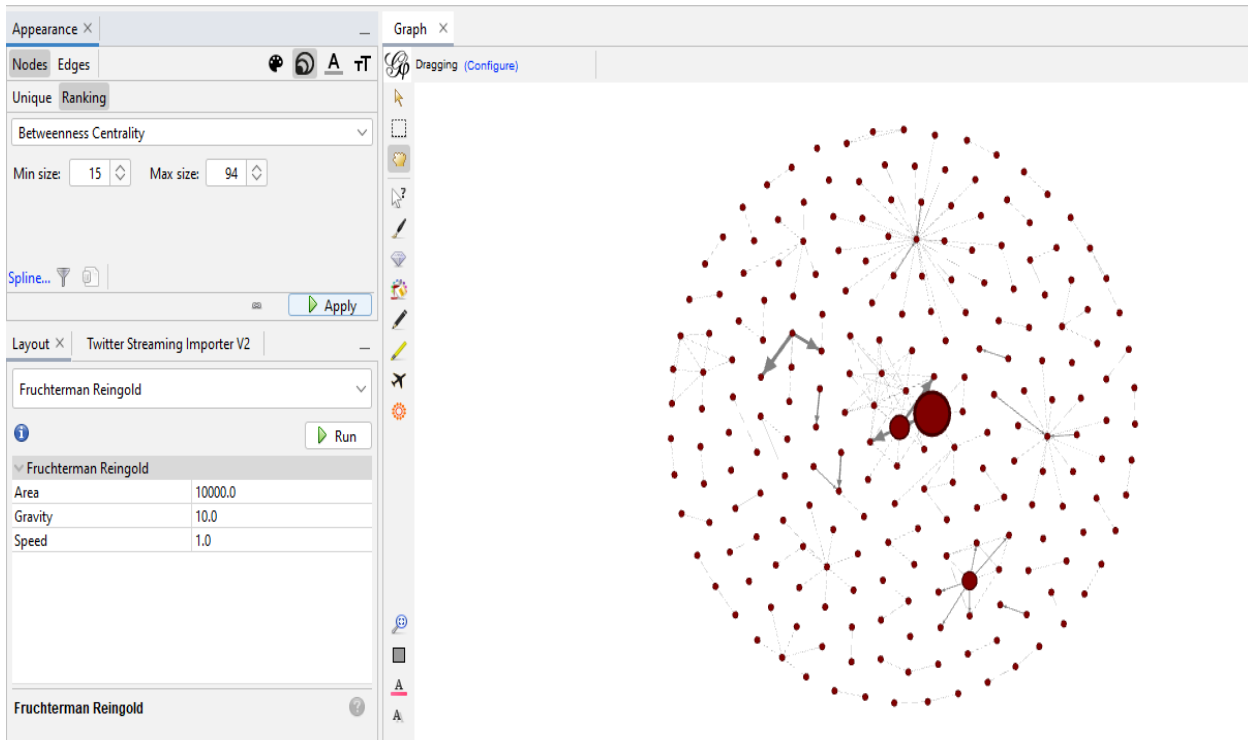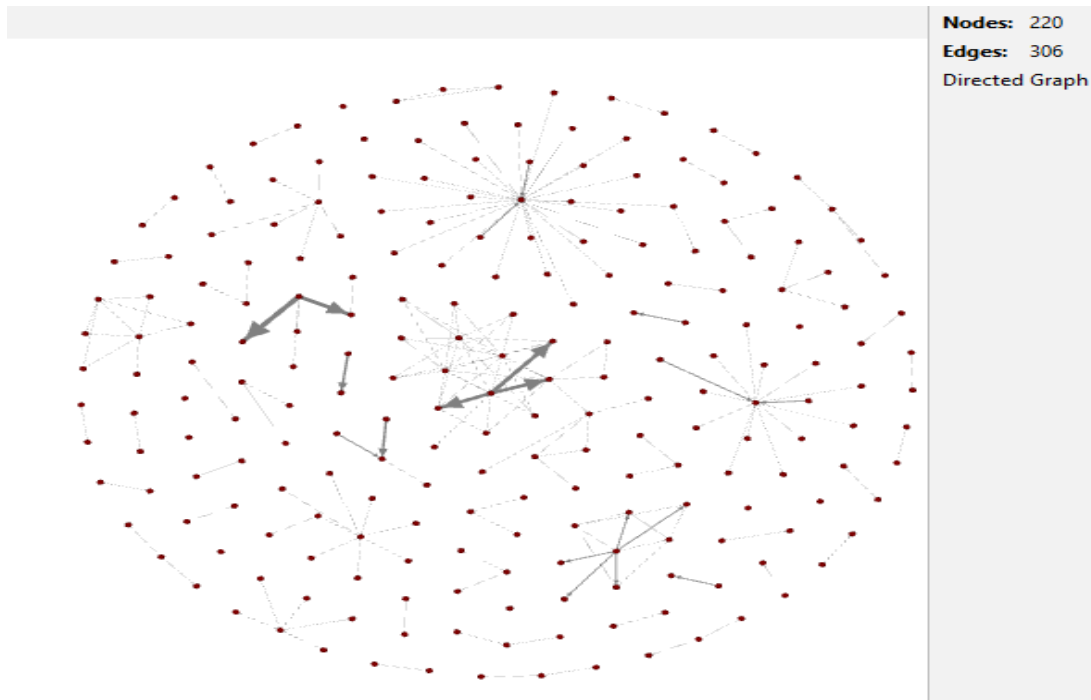| Eccentricity		
1.0		(52.27%)
0.0		(43.64%)
2.0		(3.64%)
3.0		(0.45%)
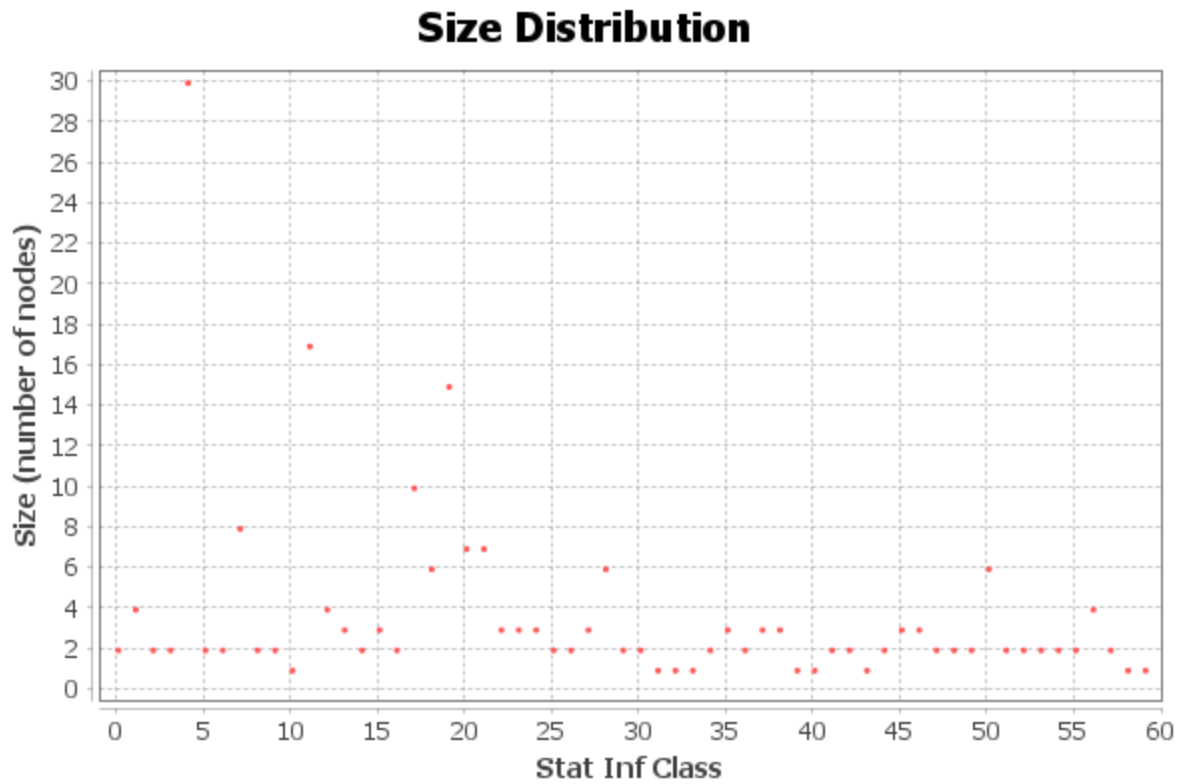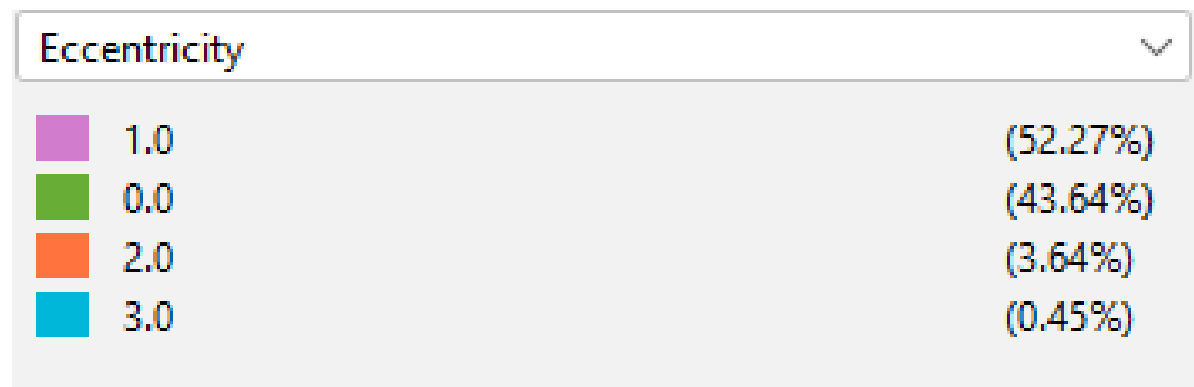
8. Betweenness Centrality

9. Directed Graph:

10. In-Degree and Fruchterman Reingold

Final View with Labels