Complex Network Project: Social Circles: Facebook

Varaga Haghoubians
University of Turin
June 2024

Abstract

This project is a comprehensive analysis of the Facebook social network graph. The study includes various network measures such as degree distribution, centrality metrics, clustering coefficients, community detection, and assortativity coefficient, providing insights into the structure and dynamics of social circles within the network.

1 Introduction

People connect through social networks, such as Facebook, which are intricate systems based on a variety of social criteria. Knowledge about these networks can help one understand how people behave in social situations, how information spreads, and how strong networks are. Using a variety of network metrics and methods, this study attempts to evaluate the Facebook social network graph.

2 Degree Distribution and Power-Law Fit

The degree distribution of the Facebook social network follows a power-law, indicative of a scale-free network. This is illustrated in Figure 1.

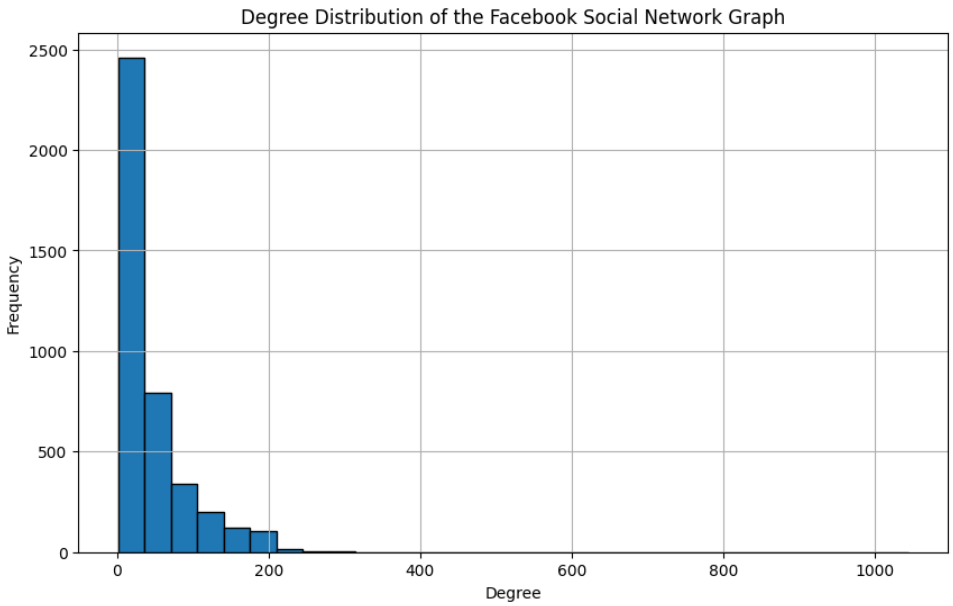

Figure 1: Degree Distribution of the Facebook Social Network Graph.

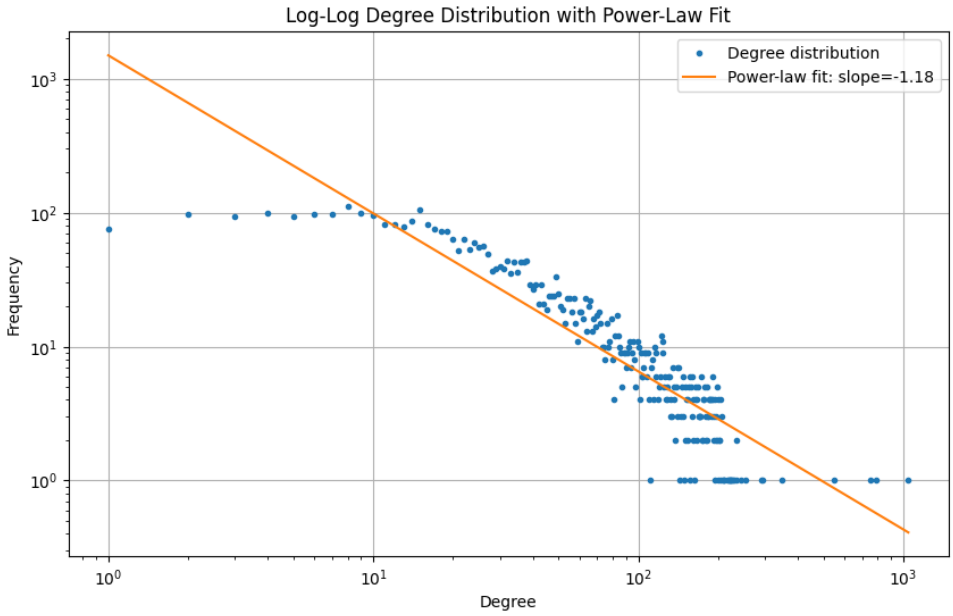

Figure 2: Log-Log Degree Distribution with Power-Law Fit.

Interpretation and Discussion:

- **Average Degree**: 43.69
- **Variance of Degree**: 2747.24
- **Standard Deviation of Degree**: 52.41
- **Power-Law Slope**: -1.18
- **R-squared of Fit**: 0.809

We plot log-log Scale is the best way for heavy tailed distributions The degree distribution is highly skewed, with most nodes having a low degree and a few nodes having a very high degree. The power-law slope of -1.18 and an R-squared value of 0.809 indicate a scale-free network, where few nodes act as hubs with many connections, while most nodes have fewer connections.

3 Distribution of Shortest Path Lengths

The distribution of shortest path lengths in the Facebook social network graph is shown in Figure 3.

Interpretation and Discussion:

- **Average Shortest Path Length**: 3.69

Most shortest paths are around 3 to 4 steps, suggesting that the network has a small-world property where most nodes can be reached from any other node with a relatively small number of hops.

4 Clustering Coefficient

The distribution of clustering coefficients is shown in Figure 4.

Interpretation and Discussion:

- **Average Clustering Coefficient**: 0.6055

The high average clustering coefficient of 0.6055 indicates strong local clustering, suggesting that friends of a person are likely to be friends with each other. This supports the existence of tight-knit communities within the network.

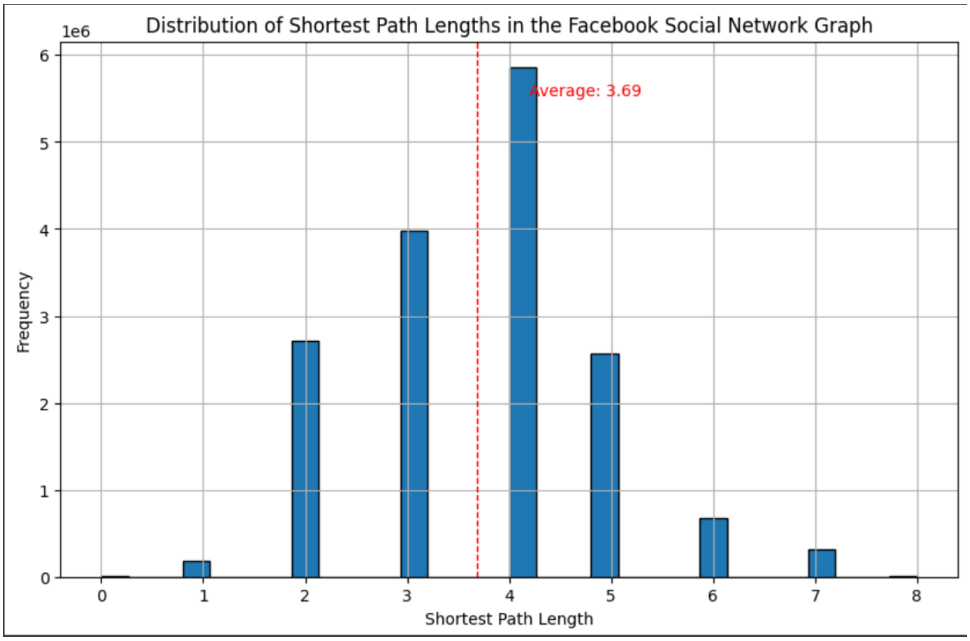

Figure 3: Distribution of Shortest Path Lengths in the Facebook Social Network Graph.

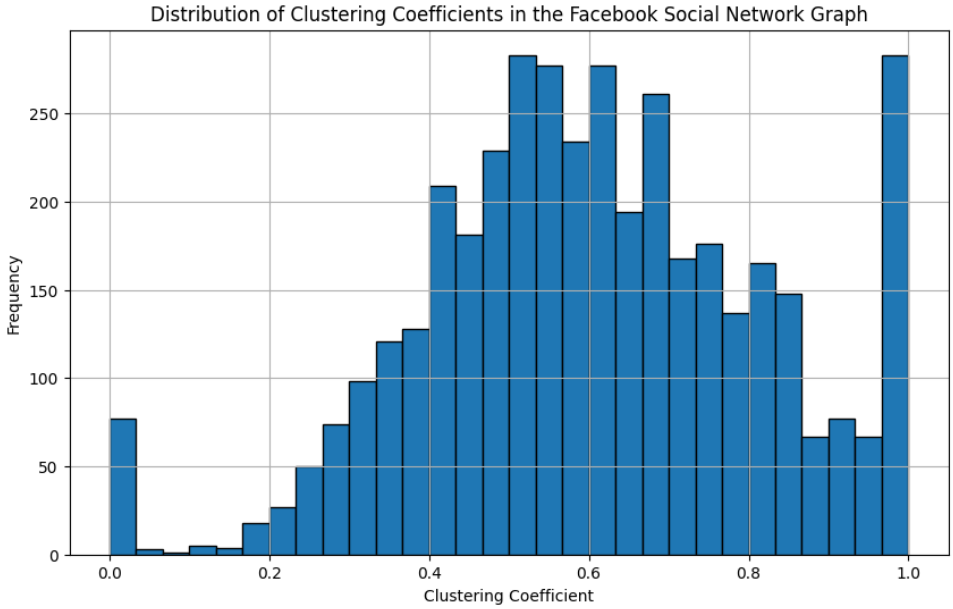

Figure 4: Distribution of Clustering Coefficients in the Facebook Social Network Graph.

5 Community Detection

The Label Propagation algorithm identified 44 distinct communities within the network, as visualized in Figure 5.

Interpretation and Discussion:

- **Number of Communities**: 44
- **Community Sizes**: [198, 36, 10, 8, 8, 34, 2, 215, 16, 3, 3, 1030, 6, 7, 3, 3, 753, 10, 2, 2, 469, 13, 9, 3, 49, 25, 2, 60, 547, 179, 10, 9, 8, 226, 19, 4, 3, 8, 6, 14, 12, 7, 6, 2]

The community detection results reveal a mix of large and small communities, reflecting the natural social circles within the network. The largest community consists of 1030 nodes, while several smaller communities have fewer than 10 nodes.

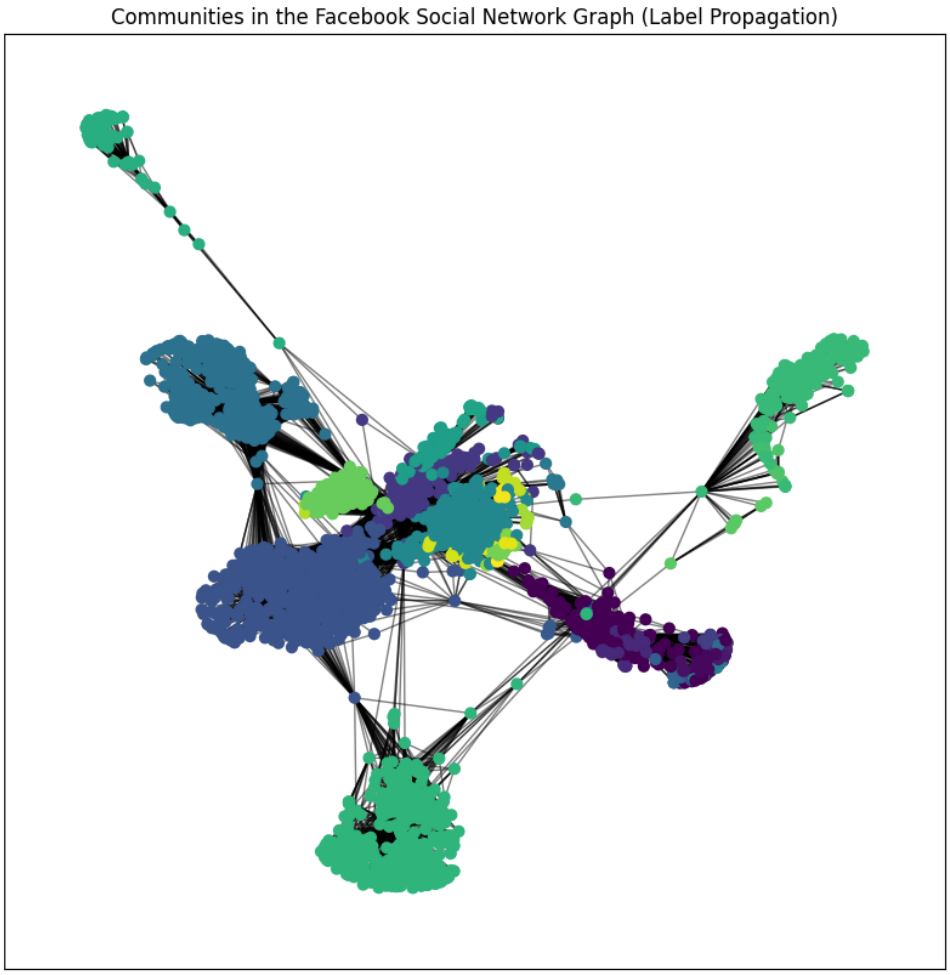

Figure 5: Communities in the Facebook Social Network Graph (Label Propagation).

6 Centrality Measures

The top 10 nodes by various centrality measures are as follows:

6.1 Degree Centrality

- Node 107: 0.2588
- Node 1684: 0.1961
- Node 1912: 0.1870
- Node 3437: 0.1355
- Node 0: 0.0859
- Node 2543: 0.0728
- Node 2347: 0.0721
- Node 1888: 0.0629
- Node 1800: 0.0607
- Node 1663: 0.0582

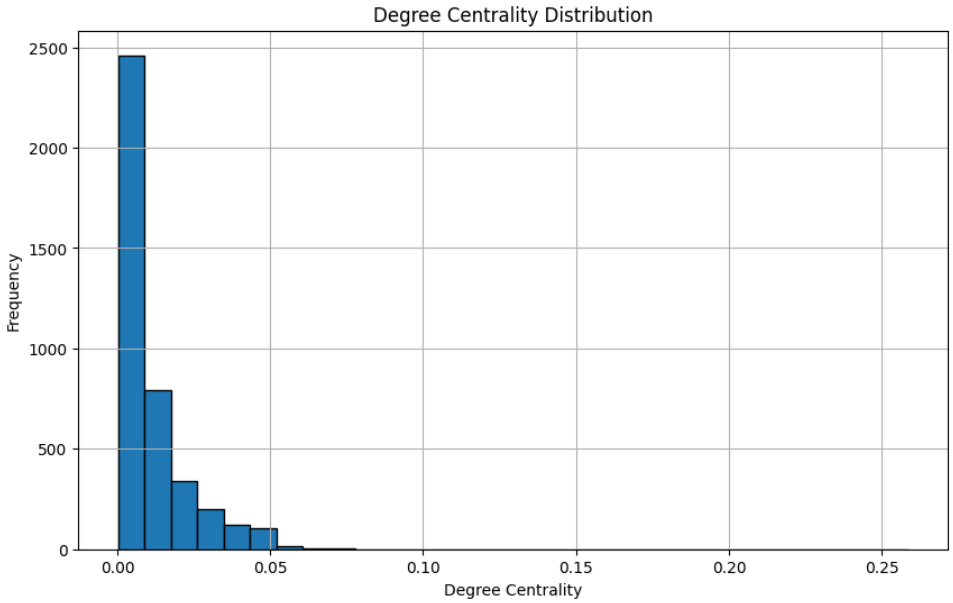

Figure 6: Degree Centrality Distribution.

6.2 Betweenness Centrality

- Node 107: 0.4805
- Node 1684: 0.3378
- Node 3437: 0.2361
- Node 1912: 0.2293
- Node 1085: 0.1490
- Node 0: 0.1463
- Node 698: 0.1153
- Node 567: 0.0963
- Node 58: 0.0844
- Node 428: 0.0643

6.3 Closeness Centrality

- Node 107: 0.4597
- Node 58: 0.3974
- Node 428: 0.3948
- Node 563: 0.3939
- Node 1684: 0.3936
- Node 171: 0.3705
- Node 348: 0.3699
- Node 483: 0.3698
- Node 414: 0.3695
- Node 376: 0.3666

6.4 Eigenvector Centrality

- Node 1912: 0.0954
- Node 2266: 0.0870
- Node 2206: 0.0861
- Node 2233: 0.0852
- Node 2464: 0.0843
- Node 2142: 0.0842
- Node 2218: 0.0842
- Node 2078: 0.0841
- Node 2123: 0.0837
- Node 1993: 0.0835

7 Homophily

The degree assortativity coefficient is 0.0636, indicating that the network is slightly assortative, meaning nodes tend to connect with other nodes of similar degree.

Interpretation and Discussion:

- **Positive Assortativity (0.0636)**:
  - Assortative mixing by degree indicates that nodes with high degrees are more likely to be connected to other high-degree nodes, and similarly for low-degree nodes.
  - In social networks, this can reflect homophily, where individuals tend to associate with others who have a similar level of connectivity.
- **Implications**:
  - **Community Structure**: Assortative mixing can lead to the formation of tightly-knit communities or clusters.
  - **Network Robustness**: Assortative networks are generally more robust to random failures but can be vulnerable to targeted attacks on highly connected nodes.
  - **Information Spread**: Information can spread quickly within clusters of highly connected nodes but may take longer to reach less connected parts of the network.

8 Conclusion

The Facebook social network graph demonstrates features common to social networks found in the real world, such as strong community structure, high connectedness, and mild assortativity. The dissemination of information, network resilience, and strategic targeting within the network are all significantly impacted by these attributes. Gaining an understanding of these network metrics can help you improve your marketing, network management, and communication tactics.