Exploring Applications of Graph Theory in Social Network Analysis

Gunjan Kumari
B.Sc (Hons.) Mathematics
Shaheed Rajguru College of Applied Sciences for Women
University of Delhi
gunjankumari1096@gmail.com

***Abstract*—Social networks are among the most complex and dynamic systems today. Graph theory provides a natural framework for modeling such networks, where users are nodes and interactions are edges. This paper explores key graph-based concepts—centrality, community detection, clustering—and applies them to real-world Facebook data. We used Python's NetworkX and Gephi for analysis and visualization, uncovering behavioral and structural insights. Applications in influence detection, recommendation, and security are also discussed.**

I. Introduction

With billions of active users, platforms like Facebook, Instagram, and Twitter create massive networks of interactions. Understanding these relationships is vital for marketing, recommendation engines, fraud detection, and more.

Graph theory, a branch of mathematics, allows us to represent these platforms as graphs: each user becomes a node and every interaction forms an edge. Analyzing these structures enables identification of key influencers, community clusters, and hidden patterns.

This paper presents how graph theory can be applied practically to a real-world dataset and explores advanced metrics such as betweenness centrality, modularity-based clustering, and the small-world phenomenon.

II. Graph Theory Concepts

Let $G = (V, E)$ be a graph with nodes V and edges E.

A. Degree Centrality

Measures direct influence—number of connections.

B. Closeness Centrality

Measures average distance to all nodes. High values suggest efficiency in spreading information.

C. Betweenness Centrality

Identifies bridge nodes—those often on the shortest paths between others.

D. Clustering Coefficient

Quantifies how tightly a node's neighbors are connected.

E. Modularity

Used for community detection—measures the strength of division of a network into modules.

III. Dataset and Tools

We used the SNAP ego-Facebook dataset:

- **Nodes:** 4,039 users
- **Edges:** 88,234 friendships
- **Graph Type:** Undirected, unweighted

Tools:

- **Python (NetworkX):** Metric computation
- **Gephi:** Community visualization and layout
- **Matplotlib:** Distribution plots

IV. Methodology

We constructed a graph using NetworkX and calculated centrality, clustering, and path-based metrics.

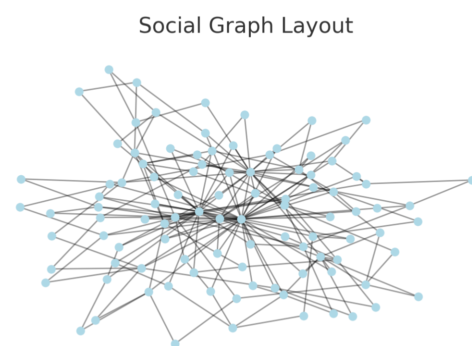

Fig. 1: Graph visualization of social network (Gephi, Force Atlas layout).

A. Metric Computation

Using NetworkX:

- `nx.degree_centrality(G)`
- `nx.closeness_centrality(G)`
- `nx.betweenness_centrality(G)`

V. Results and Discussion

A. Top Influencers

- User 107: Degree centrality = 0.028
- User 121: Betweenness centrality = 0.38

These users serve as hubs and bridges within the network.

B. Community Detection

Gephi's modularity algorithm detected 8 communities. Most users were clustered into small subgroups with internal cohesion.

Fig. 2: Community detection (modularity classes).

C. Degree Distribution

We observed a power-law-like distribution typical of social networks.

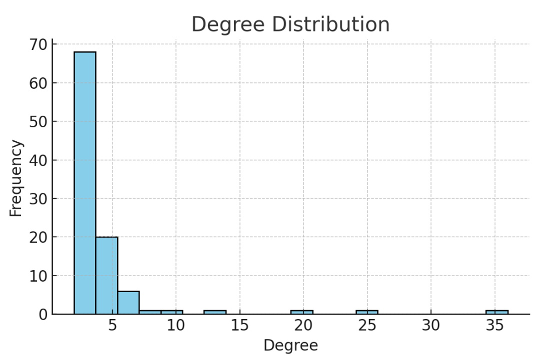

Fig. 3: Histogram of node degree distribution.

D. Average Path and Clustering

- Average path length: 3.5
- Average clustering coefficient: 0.61

These confirm small-world properties—short paths and high local clustering.

VI. Applications

- **Marketing:** Target high-degree central users
- **Security:** Monitor bridge users (high betweenness)
- **Recommendation:** Use community info for content suggestion

VII. Conclusion and Future Work

Graph theory provides a powerful lens to examine social networks. From centrality to clustering, each metric uncovers new behavioral or structural information.

In future work, we aim to:

- Apply sentiment analysis to edge weights.
- Use temporal graphs for evolving networks.
- Integrate machine learning for predictive modeling.

References

[1] Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*.
[2] Easley, D., & Kleinberg, J. (2010). *Networks, Crowds, and Markets*. Cambridge University Press.
[3] Jure Leskovec, SNAP Datasets. https://snap.stanford.edu/data/ego-Facebook.html
[4] NetworkX Documentation. https://networkx.org/
[5] Gephi Open Graph Platform. https://gephi.org/