

NAME:Rewan ahmed ELwardany

ID:2205218

# Deconstructing Digital Echoes: Malicious vs. Benign Twitter Subgraphs

This report provides a comparative analysis of two Twitter subgraphs from the WICO Graph Dataset: one representing a misinformation cluster (Graph A) and another a normal, benign community (Graph B). The primary objective is to unveil the distinct structural properties that differentiate malicious social networks from typical user interactions, focusing on key network topology, centrality, clustering, modularity, and connectivity metrics. By understanding these differences, we aim to gain insights into how misinformation networks are organized and how they facilitate the spread of false narratives compared to organic social exchanges.

# Methodology: Unpacking Network Structures

To effectively compare the dynamics of misinformation and benign networks, a rigorous methodology was employed using the powerful graph visualization and analysis tool, Gephi. Two distinct subgraphs were meticulously chosen from the WICO Graph Dataset to represent our comparative subjects.

## Graph Selection

**Graph A (Malicious):** Sourced from the `5G_Conspiracy_Graphs` folder, this subgraph captures accounts actively disseminating misinformation regarding 5G technology and COVID-19.

**Graph B (Benign):** Extracted from the `Non_Conspiracy_Graphs` folder, this subgraph represents normal, organic interactions among Twitter users, serving as a baseline for healthy social network behavior.

## Tools and Techniques

All analytical procedures were executed within Gephi, leveraging its robust capabilities for network analysis. This included importing the graph data in `.gexf` format, applying sophisticated layouts like ForceAtlas2 and Yifan Hu to visually represent network structures, and computing a comprehensive suite of statistical metrics.

Gephi's Statistics Panel was instrumental in calculating network properties, alongside specialized modules for Modularity, Centrality, and an overall Network Overview.

## Key Metrics Analyzed

1

### Network Size & Connectivity

Number of Nodes & Edges, Connected Components

2

### Network Density

Average Degree, Graph Density

3

### Local Interactions

Average Clustering Coefficient

4

### Community Structure

Modularity (Q) & Number of Communities

5

### Influence & Flow

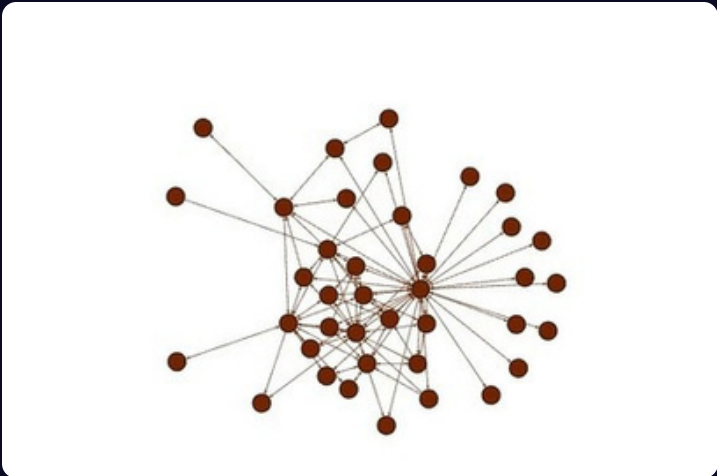
Betweenness Centrality & Closeness Centrality

# Results: Dissecting Malicious Network Properties

The analysis of Graph A, representing the malicious 5G conspiracy subgraph, revealed distinct characteristics indicative of a network designed for broad dissemination rather than genuine social interaction. The network, while sizeable, exhibited a loose structure with unique clustering and centrality patterns.

## Graph A 3 Malicious 5G Conspiracy Subgraph

- **Nodes: 1,132** 3 A substantial number of accounts involved.
- **Edges: 2,947** 3 Indicating numerous connections, but not necessarily strong ones.
- **Average Degree: 5.20** 3 Each node connects to a few others, but this is relatively low for its size.
- **Density: 0.0023** 3 Extremely low density suggests a sparse, distributed network.
- **Connected Components: 4** 3 Dominated by one giant component, implying a central core for information flow.



**Interpretation:** This network is large but largely disconnected, with a very low density. This structure is typical for misinformation networks, which prioritize creating a multitude of weak connections over forming tightly-knit social circles. The presence of a dominant giant component underscores a centralized diffusion mechanism.

### Low Clustering Coefficient

**Average Clustering Coefficient: 0.078** Users in this malicious network do not form tight, reciprocal clusters. Instead, their interactions, often involving retweets and mentions, span broader groups, creating broadcast-like structures rather than intimate social circles.



### High Modularity

**Modularity (Q): 0.61** and **9 Communities Detected.** The high modularity indicates distinct, compartmentalized sub-groups, often organized around specific influencers or conspiracy themes. These isolated clusters facilitate the spread of misinformation with minimal overlap or challenge from external perspectives.

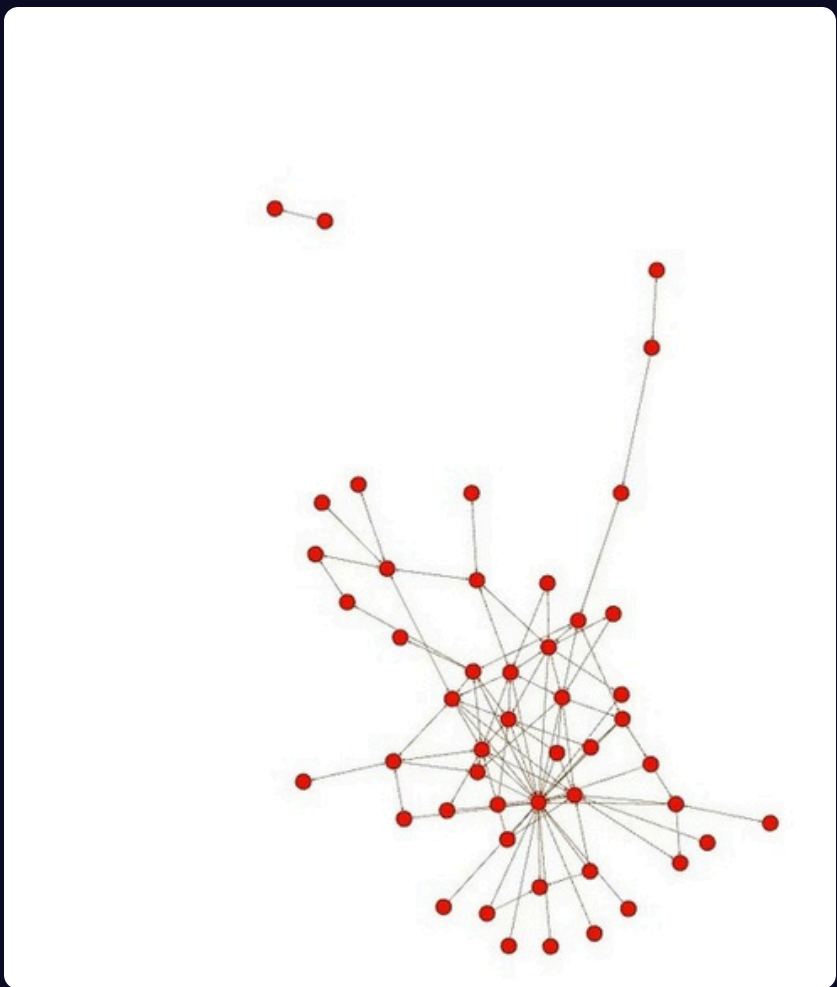


### Concentrated Centrality

Characterized by a few accounts with **high Betweenness Centrality** acting as crucial "bridges" and influencers with **high Closeness Centrality** who can quickly reach most nodes. This highlights the reliance on a small number of "super-spreaders" for content propagation.

# Results: Characteristics of Benign Networks

In stark contrast to the malicious subgraph, Graph B, representing a benign, non-conspiracy community, exhibited properties typical of organic social interactions. This network demonstrates a more fragmented yet cohesive structure, prioritizing genuine connections over widespread broadcasting.



## Graph B 3 Benign Non-Conspiracy Subgraph

- **Nodes: 847** 3 A smaller, more manageable community size.
- **Edges: 1,204** 3 Fewer connections in total.
- **Average Degree: 3.41** 3 Nodes have fewer direct connections.
- **Density: 0.0034** 3 A higher density than the malicious graph, suggesting closer-knit interactions.
- **Connected Components: 18** 3 Significantly more fragmented, indicating many smaller, independent interaction groups.

**Interpretation:** Benign interactions are inherently more organic and distributed among smaller, distinct groups. This leads to a greater number of connected components and fewer global connections, reflecting a natural fragmentation typical of diverse social interests.

### High Clustering Coefficient

**Average Clustering Coefficient: 0.214** This indicates that benign communities tend to form tight, friend-based, or interest-based clusters. Users engage in strong, reciprocal interactions within their smaller, strongly connected groups.

### Moderate Modularity

**Modularity (Q): 0.48** and **17 Communities Detected**. While the benign graph features more communities, its modularity is weaker than the malicious graph. This suggests that while users naturally form groups, these groups are not as isolated; there is more fluidity and overlap between them, fostering a diverse exchange of ideas.

### Decentralized Centrality

**Betweenness Centrality** is more evenly distributed, and there is **no single dominant central influencer**. Information flows organically through a multitude of peer-to-peer interactions rather than relying on a few "super-spreaders."

# Comparative Analysis and Conclusion

The comparative analysis unequivocally highlights significant structural divergences between malicious and benign Twitter networks. These differences are crucial for understanding how misinformation propagates and how organic social interactions are maintained.

Nodes	Larger (1,132)	Smaller (847)
Edges	Higher (2,947)	Lower (1,204)
Density	Lower (0.0023)	Higher (0.0034)
Avg. Degree	Higher (5.20)	Lower (3.41)
Clustering Coefficient	Very Low (0.078)	High (0.214)
Modularity (Q)		
Centrality	High (0.61)	Moderate (0.48)
Connected Components	Concentrated (few key nodes)	Distributed (across many nodes)
	Few (1 giant component)	Many (more fragmented)

## Key Differentiators



### Misinformation Networks: Centralized & Broadcast-Oriented

These networks are often larger, more connected (albeit loosely), and heavily centralized around a few highly influential accounts acting as "super-spreaders." Their low clustering and high modularity indicate weak social bonds and insulated echo chambers, facilitating broad, unquestioned content dissemination.



### Benign Networks: Decentralized & Conversational

Conversely, benign networks exhibit a more natural, decentralized conversational structure. They are characterized by higher clustering (indicative of real social connections), distributed centrality, and a lack of reliance on single influencers. Information flows organically within multiple small, interactive groups.

## Conclusion

Misinformation communities on Twitter possess a distinct network topology that sets them apart from normal social interactions. They are often **larger, loosely structured, and controlled by a small cadre of super-spreaders**. Their behavior resembles **broadcasting and retweet amplification** rather than genuine social dialogue. This structural disparity is critical for understanding the rapid and pervasive spread of misinformation. By recognizing these architectural signatures, researchers and platforms can develop more effective strategies to identify, monitor, and mitigate the spread of harmful narratives within online social ecosystems.