hello conspiracy (Case Study)

1. Overview

This project investigates the structural behavior of Twitter networks associated with the 5G conspiracy narrative using graph analysis techniques. **By examining how misinformation spreads through network structures**, the study identifies distinctive patterns that differentiate coordinated misinformation clusters from normal social interactions.

---

2. Dataset Description

- **Source:** WICO Twitter Subgraph Dataset
- **Sampling Technique:** Random sampling of **25%** of all available subgraphs (via Python code(
- **Sample Size:**
  - **103 out of 412** subgraph folders for **5G Conspiracy**
  - 625 **out of** 2502 Subgraph folders for **Non 5G Conspiracy**

- Each folder contains:
  - `nodes.csv – metadata of users`
  - `edges.txt – relationships/interactions between users`

---

3. Tools & Technologies Used

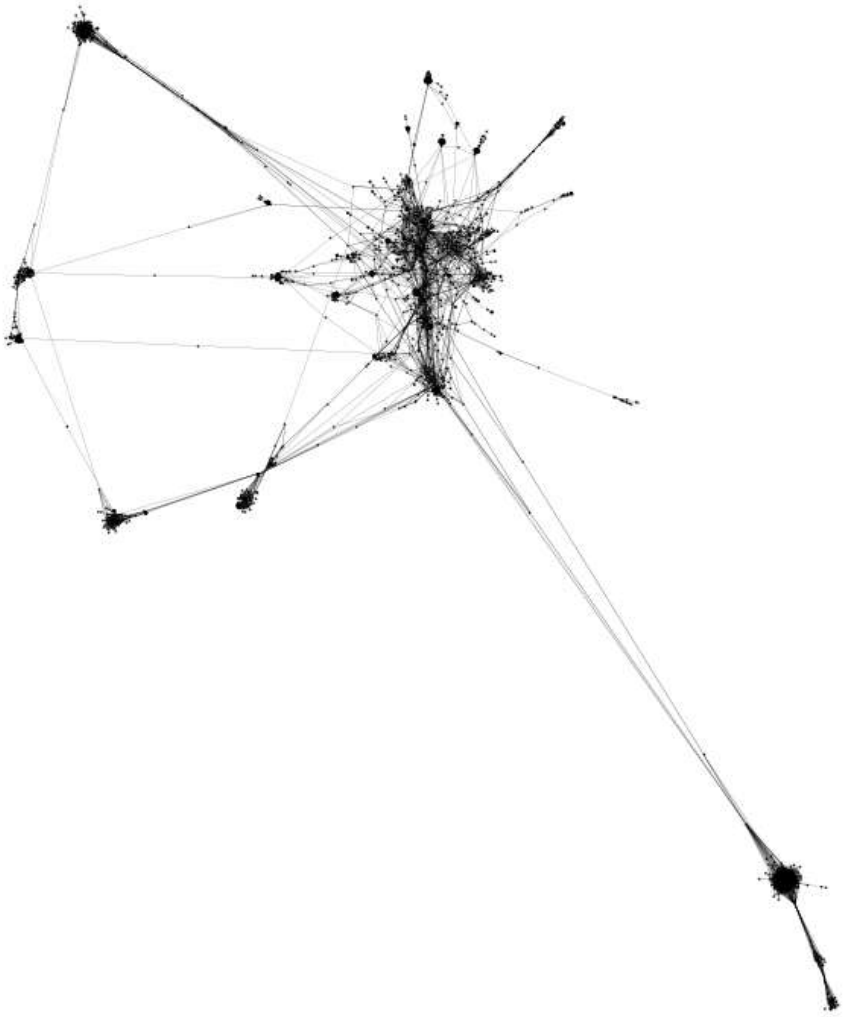
- **Python**
  - Custom script designed to:
    - Search folders recursively
    - Merge all node files into one unified **CSV**
    - Convert messy **TXT** edge files into clean **CSV** format
    - Handle different separators and malformed structures
    - Sample **25%** of graph folders automatically

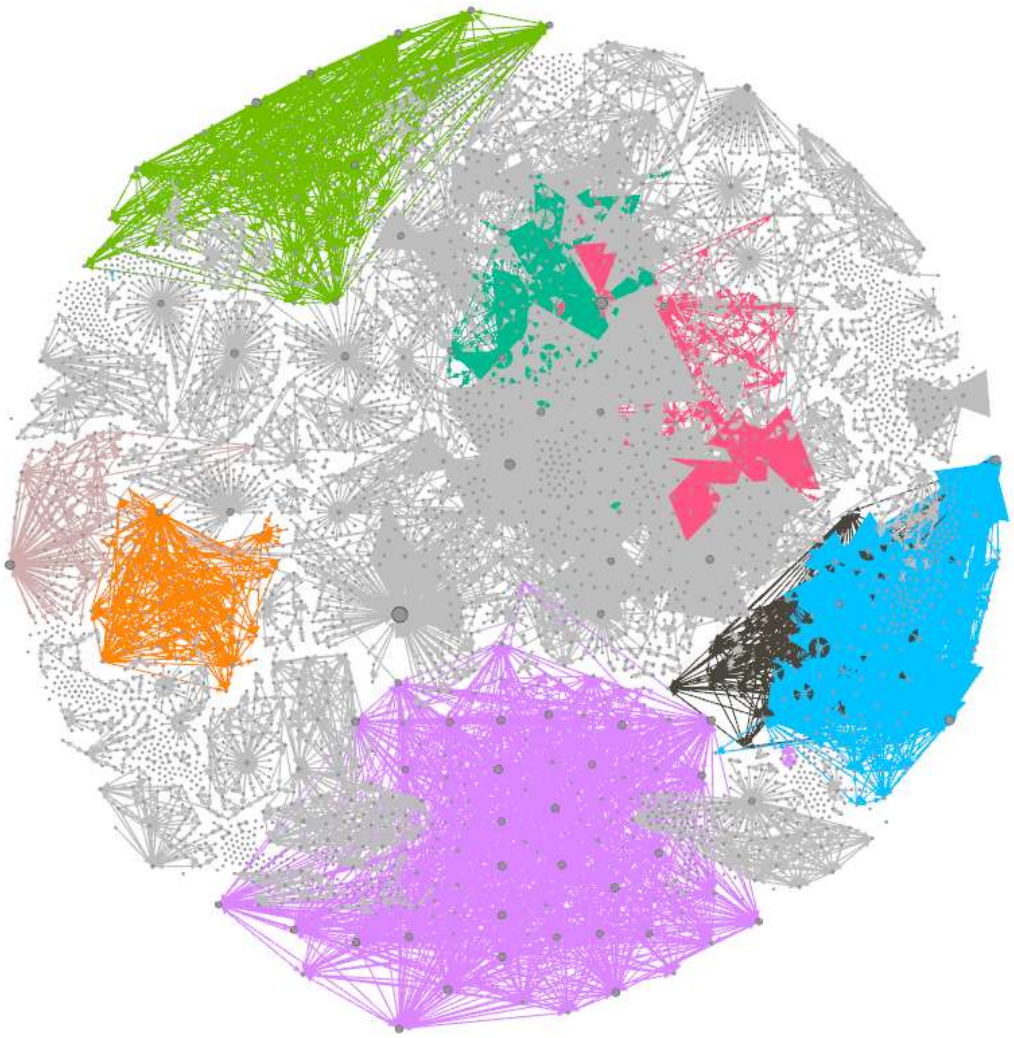
3.2 Graph Analysis & Visualization

- **Gephi**
  - Used for:
    - Graph construction
    - Layout visualization
    - Statistical analysis (degree, betweenness, eigenvector, clustering, modularity)
    - Community detection
    - Identifying structural anomalies

5G Conspiracy Data

I. Graphs :





Look how clearly the separate groups appear!

2. Network Overview

Metric	Value	Interpretation
Nodes	**3094**	Moderate-sized network
Edges	**11875**	Indicates high interaction volume
Average Degree	**2.927**	Each node connects to ~3 others on average
Graph Density	**0.001**	Very sparse, common in large social graphs
Network Diameter	**15**	Longest distance between two nodes
Connected Components	**971**	Highly fragmented; many disconnected groups

3. **Graph Interpretation :**

0. Overall :

The graph clearly **shows that the 5G conspiracy** network is composed of several **distinct communities (clusters)**, each represented by a different color. These communities are **highly interconnected internally**, but **weakly connected** to one another. This structure is typical of **echo chambers** and coordinated misinformation clusters.

---

1. High internal density within each cluster

Each colored group shows:

- Extremely dense internal connections
- Strong local interactions
- Frequent retweets and repetitive exchanges inside the same community

This indicates:

- **Coordinated behavior**
- **Message reinforcement**
- **Closed circles of communication**

Clusters like these often amplify the same misinformation within a restricted audience.

---

2. Very weak connections between communities

The noticeable gaps between clusters and the large amount of grey nodes represent:

- Very limited interaction between communities
- Messages do not naturally spread across the entire network
- Each group operates in a separate “bubble”

This is a clear sign of:

- **Echo chambers**
- **Insular communities**
- **Fragmented information flow**

A hallmark of misinformation networks.

---

3. Star-like patterns inside clusters

Within each community, you can see:

- One or a few **central nodes**
- Surrounded by many small leaf nodes

This "**star pattern**" is typically associated with:

- **Central accounts driving the narrative**
- **Followers amplifying the message**
- **Possible bot-assisted boosting behavior**

The network relies heavily on a few influencers or automated hubs.

---

4. Interpretation of the black-and-white skeleton graph

The simplified monochrome graph highlights:

- The central spine of highly influential nodes
- A hierarchical structure where a few nodes dominate the network
- Long, thin arms extending outward, representing weakly connected components

This reveals:

- **Reliance on a small set of core accounts**
- **Non-organic, centralized spread**
- **A structurally fragile but highly coordinated network**

---

4. Community Detection

Metric	Value
Modularity	**0.887**

Interpretation: Extremely high modularity indicates strong community boundaries. The network is composed of well-separated clusters. These clusters may represent isolated echo chambers promoting similar narratives.

---

5. Node - Level Centrality Metrics

Metric	Value
Avg. Clustering Coefficient	**0.199**
Avg. Path Length	**5.785**

Interpretation:

- **Low Clustering Coefficient** : Nodes are weakly interconnected within neighborhoods.
- **Moderate Path Length** : Indicates average spread across the network is not immediate but achievable.

---

6. Eigenvector Centrality (IMP)

- **Observation:**
    - Over 1800 nodes have **very low or zero** eigenvector centrality.
    - A **very small group** of nodes have high centrality scores.

Interpretation:

- This pattern is classic in disinformation campaigns. A handful of nodes (likely bots or key influencers) are central to the network, with the majority serving as passive amplifiers.

---

7. Security Implications

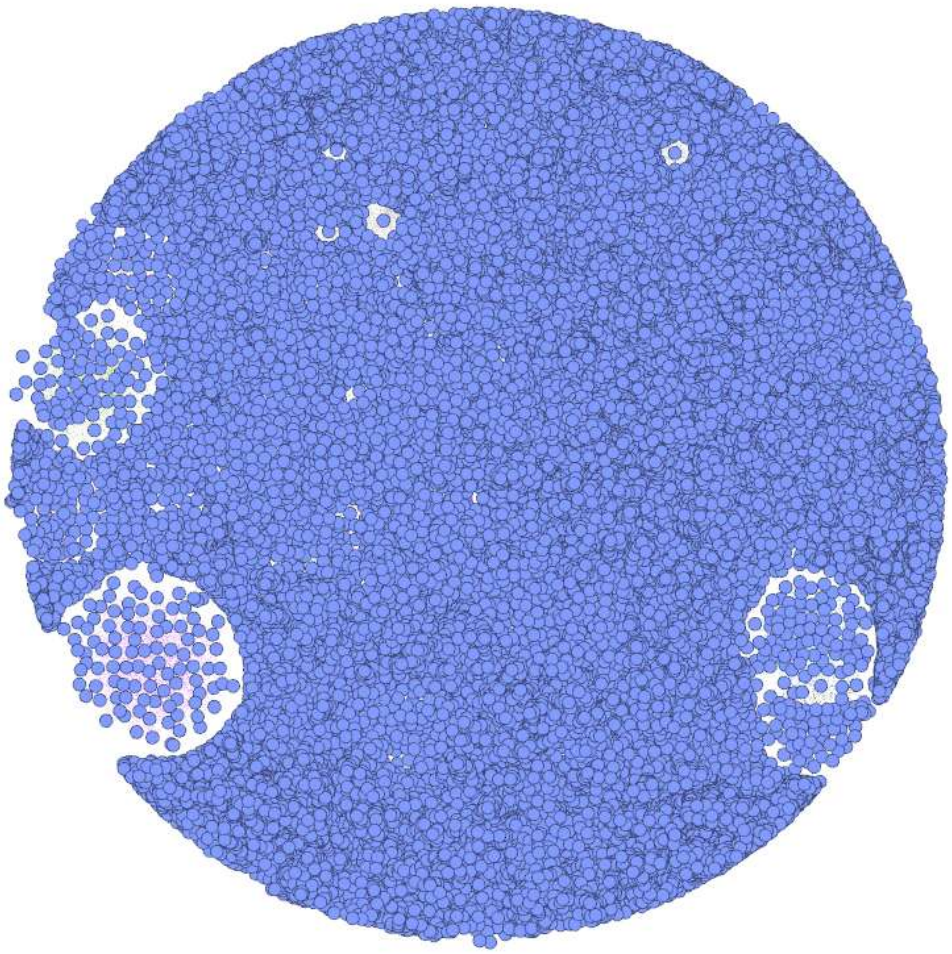
- **High Modularity + Low Density + Sparse Clustering:**
    - Suggests **intentional fragmentation** to evade detection.
    - Communities may act semi-independently, reinforcing specific claims.
- **Small number of central nodes:**
    - Likely **coordinated actors** or bots driving the narrative.
    - Can be focal points for mitigation or flagging.
- **Echo Chamber Traits:**
    - Interaction is repetitive within clusters; outside exposure is limited.

Final Conclusion

The 5G conspiracy network exhibits all key characteristics of coordinated misinformation campaigns: high modularity, fragmentation, sparse connection, and strong centralization around a few nodes. These findings support the hypothesis that this is a structured, semi-automated disinformation system, not organic public discourse.

5G Non Conspiracy Data

I.Graphs :



4. Network Overview

Metric	Value	Interpretation
Nodes	***37920***	Total number of accounts in the sampled non-conspiracy network
Edges	***87936***	Total number of interactions between accounts
Average Degree	**2.319**	Each node interacts with ~2-3 others on average
Average Weighted Degree	**2.335**	Accounts have similar interaction strength across weighted edges
Network Diameter	**48**	The longest shortest path between any two nodes : shows large spread potential

Graph Density	**0.000 (≈0)**	Extremely sparse network, typical for large-scale organic Twitter activity
Connected Components	**10,744**	Highly fragmented graph; most users are isolated or form tiny groups
Modularity	**0.933**	Very strong community structure; many well-separated clusters
Average Clustering Coefficient	**0.143**	Low : indicates weak triangle formation and minimal echo-chamber behavior
Average Path Length	**12.157**	On average, users are ~12 steps apart : typical for large organic networks

5. **Graph Interpretation :**

0. Overall :

This network shows **healthy, natural diffusion** of tweets:

- **Users interact in small groups**
  Not everyone is connected to everyone else.
- **Low density, low clustering coefficient**
  Indicates diversity and randomness in interactions.
- **No signs of coordinated manipulation**
  No tightly packed core, no abnormal hubs, no artificial amplification structure.

---

1. High-Level Shape of the Graph:

- The non-conspiracy graph shows a **huge, naturally formed social network** with thousands of scattered nodes.
  There is **no single dominant cluster**, but instead many small, loosely connected groups.

---

1. Key Visual Observations:

- **Massive spread of isolated nodes**
  Most nodes appear as small dots with very few edges. This indicates natural, everyday Twitter activity where users do not strongly engage with each other in a coordinated manner.

- **Several small clusters on the edges**
  You can see small **clouds** or pockets of nodes **(white circles)**. These represent groups of users interacting around the same tweet or topic - but still not tightly connected enough to form a coordinated pattern.

- **No dense echo chamber**
  Unlike conspiracy graphs, there is **no single highly interconnected ball of nodes**.
  This is typical of **organic Twitter communication**, where users engage sporadically.

- **The full-graph view (second image)**
  When reducing opacity, you can clearly see:

  - Several **star-shaped clusters**.
  - A number of **tree-like branches** .
  - No tight bot-like dense blocks.

Comparison Between the Non-Conspiracy Network and the 5G Conspiracy Network

Compared to the 5G conspiracy subgraph, the non-conspiracy network displays a very different structural behavior.

In the conspiracy network, even though modularity was high, the graph formed a smaller number of well-defined communities, often centered around a clear dense core of highly interconnected accounts. This structure produced higher clustering in certain regions and indicated more coordinated interaction patterns.

In contrast, the non conspiracy sample is far more fragmented, with over 10,000 disconnected components, extremely low density, and a long average path length within its largest connected component. This suggests that non conspiracy discussions are highly dispersed, consisting of many small, loosely connected groups or even individual accounts acting independently.

Overall:

- Conspiracy network → concentrated, tightly connected clusters (possible coordination)
- Non-conspiracy network → scattered, low-connectivity structure (organic, uncoordinated activity)