Ass 1 Social Networks

Sara Nabil Tahawy 2205224

Comparing Malicious vs. Benign Twitter Subgraphs (WICO Dataset)

This report compares two Twitter subgraphs from the WICO Dataset: one from the 5G_Conspiracy_Graphs folder (a misinformation cluster) and one from the Non_Conspiracy_Graphs folder (a normal cluster). Both networks were analyzed using Gephi. The goal is to examine structural differences that reflect how misinformation spreads compared to normal interactions.

Methods and Visualization

I made the two Twitter graphs in Gephi and used the ForceAtlas2 layout to make the network clear. I checked metrics like average degree, density, clustering coefficient, modularity, betweenness, closeness centrality, connected components, diameter, and average path length. I colored the nodes by their community (modularity) and made the important nodes bigger depending on their degree or betweenness. This helped me see which accounts were more central.

Results

Metric	5G Conspiracy (Misinformation)	Non-Conspiracy (Normal)	Interpretation / Note
Graph type	Directed	Directed	Direction matters (who points to whom)
Nodes	34	51	Number of accounts
Edges	42	127	Number of directed connections
Average degree	1.235	2.490	Avg in+out degree per node
Density	0.037	0.050	Proportion of possible edges present
Avg. clustering coefficient	0.033	0.308	Local group cohesion
Modularity (Q)	0.685	0.396	Higher = clearer, more isolated communities
Number of communities	8	6	Detected communities by Louvain algorithm
Weakly connected components	7	1	Components ignoring edge direction
Strongly connected components	21	22	Components where each node can reach every other along directions
Diameter	4	7	Longest shortest-path between any two nodes
Average path length	1.764	2.981	Average shortest-path distance
Betweenness centrality (range)	0 → 47.5 (few high)	0 → 769.75 (wider)	Nodes with high betweenness bridge communities
Closeness centrality	low–medium	higher on average	How quickly a node can reach others

Interpretation and Analysis

In analyzing the two Twitter subgraphs from the WICO dataset, it becomes clear how misinformation spreads differently from normal interactions. One important concept that appears in the misinformation graph is the presence of echo chambers. An echo chamber is a small group of users who mostly interact with each other and rarely connect to the rest of the network. In the 5G-misinformation graph, this is visible through the high modularity value and the presence of many disconnected or weakly connected communities. Network analysis helps identify these echo chambers because they appear as clusters that have many internal links but almost no external links.

Graph visualizations also help a security analyst detect coordinated misinformation campaigns. When layouts like ForceAtlas2 are applied, suspicious patterns become visible—such as tight groups of accounts all pointing toward the same user, or clusters with very low clustering coefficient but highly centralized connections. These visual cues often indicate bot groups or coordinated attempts to push a narrative. The 5G-misinformation subgraph shows this clearly through its fragmented structure and its dependence on a few bridge nodes to connect the communities.

This important part of the analysis is community detection, which groups nodes into meaningful clusters. In security, this is essential because harmful or misleading content often spreads inside a specific community before reaching others. By identifying which communities exist, analysts can understand where misinformation originates and which nodes are acting as influencers inside those groups. It also helps highlight communities that are abnormally isolated or overly dense, which are common signs of botnets or organized manipulation.

The presence of many disconnected components in the misinformation graph (shown as purple or green groups in Gephi) also reveals something important: the misinformation tweet did not spread widely across the network. Instead, it remained trapped in several separate micro-clusters, each of which interacted mainly inside itself. This means the information did not travel far, and the spread relied heavily on a few central accounts to connect these isolated parts.

A key takeaway from the WICO subgraph study is that misinformation networks are structurally weaker and more fragmented than normal communication networks. Because of this, targeting the few high-centrality nodes—especially those with high betweenness—can significantly reduce the spread of harmful narratives. Understanding the network structure allows future misinformation campaigns to be detected earlier and disrupted more effectively, especially when combined with graph-based analysis and machine learning techniques.

Conclusion

The structural differences between the two graphs explain different spreading behaviors. Misinformation networks tend to form isolated echo chambers with few bridges, making them vulnerable to targeted interventions at high-betweenness nodes. Normal networks show denser, more clustered interaction patterns with distributed influence.

Thank You.