

Information Visualisation Final Report

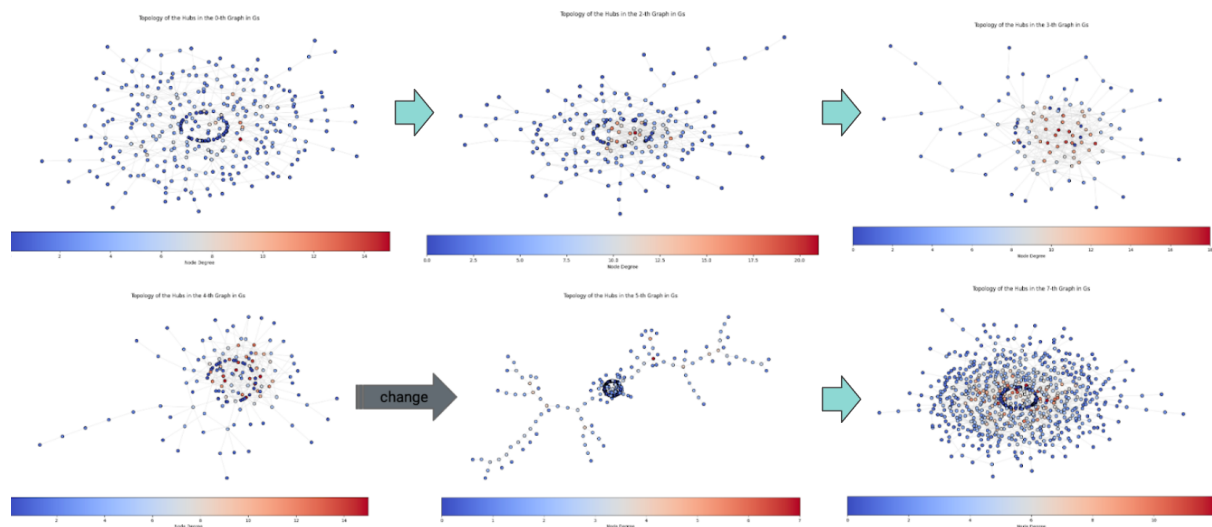
Peer to Peer network Visualisation - Yang Ma (YAM38)

Dataset

The dataset provides 9 snapshots of the P2P network from August 2002, captured on days 4, 5, 6, 8, 9, 24, 25, 30, and 31. Due to the irregular intervals between these snapshots, any ad-hoc analysis might not be fully representative. This dataset was sourced from the [SNAP Stanford](#) network dataset.

[Github Repo of this Project](https://github.com/emmhh/InfoVizFinalProject) - <https://github.com/emmhh/InfoVizFinalProject>

Figure 1 - Topology



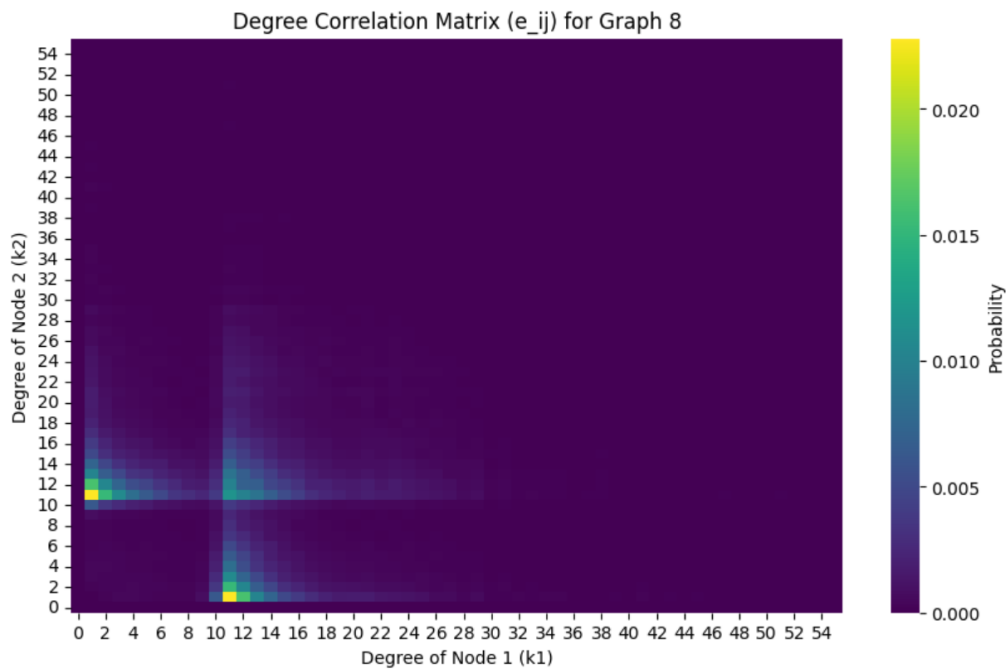
Legend Explained:

- This is the visualisation of the network topology over time. A few snapshots were not included since they are similar to the first three networks, they will be included in the final report if necessary.
- The node in red represents higher degree, nodes in blue represents lower degree. Each node in the graph represents a hub of the network with degree greater than 20. The indication of 0 in the graph represents the degree of 20, and vice versa.

Major Findings and significance:

- Graphs are dynamic throughout the month.
- Middle of the month graphs are less dense compared to the first and last graphs.
- From graph 4 to 5 (in third row) there's a sudden change in the topology and is indicated with the grey arrow in the visualisation.
- 15 days of data missing between graph 4 and graph 5; Explains the sudden change in network structure.
- It helps us to visualise the network and have a better understanding of the topology, helps us to settle the topics that needs to be studied.

Figure 2 - Degree Correlation



Legend Explained:

- The graph is a degree correlation matrix for Graph 8, showing the probability of a connection between nodes with degrees k_1 and k_2 .
- The x-axis represents the degree of Node k_1 , and the y-axis represents the degree of node k_2 .
- The color intensity corresponds to the probability of a connection, with darker colors indicating a lower probability and lighter colors (towards green) indicating a higher probability.

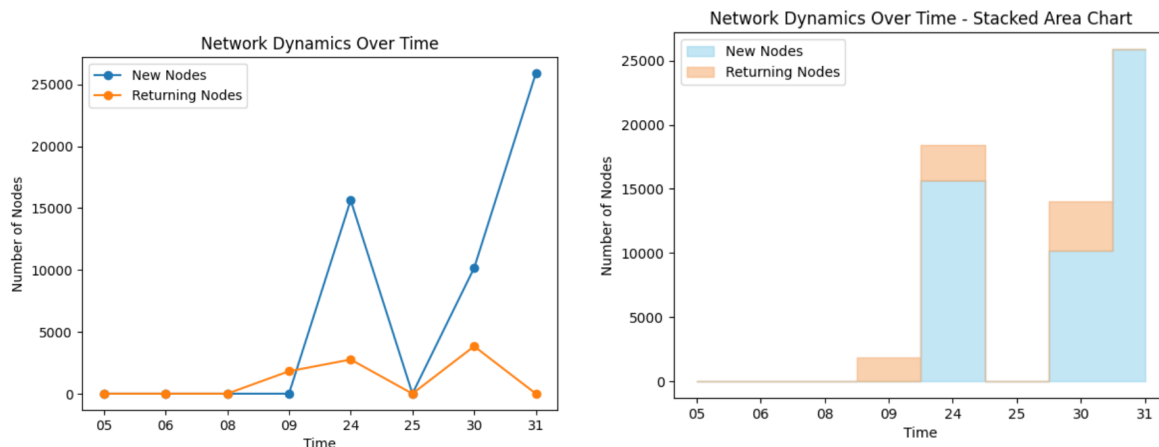
Major Findings:

- The matrix shows that higher probabilities are concentrated around nodes with lower degrees, as indicated by the lighter colors in the bottom left corner.
- There are fewer connections (lower probabilities) between nodes with higher degrees, as seen by the predominance of dark color in the upper and rightmost areas of the matrix.

Significance:

- This matrix helps to understand the tendency of nodes to connect with others that have a similar or different degree within the network.
- The matrix provides insights into the underlying structure of the network, which can be crucial for optimizing network performance or designing interventions.
- It also provides insights about analyzing the networks resilience to attacks or failures, for instance, a network where high-degree nodes are mostly connected to other high-degree nodes may be more robust against random failures.

Figure 3 - Network Dynamics



Network Dynamics Over Time (Figure on left)

Legend Explained:

- The x-axis represents time, marked by days of the month, while the y-axis represents the number of nodes.
- The blue line indicates the number of new nodes joining the network over time, and the orange line represents the number of nodes that were present before and returned to the network on a given day.

Findings and significance:

- There is a significant peak in new nodes on the 31st, suggesting a large influx of new users or a major event that caused many new nodes to join the network.
- The number of returning nodes remains relatively stable, with a slight increase towards the end of the period.
- Majority of the joining nodes are new nodes that weren't shown in the network.
- This graph is important for understanding the growth and retention dynamics within the network.
- Analyzing the trends of new versus returning nodes can help network administrators to evaluate the impact of interventions and the overall health of the network.

Stacked Area Plot (Figure on right) - Improve the figure with stacked area chart to increase readability and aesthetics.