ONLINE SOCIAL NETWORK ANALYSIS

TEAM

Name	IIT ID
Venkata Manikanta Monic Kamisetty	A20446683
Rajesh Kumar Bandaru	A20446254

Overview

The network analysis is about three giant space organizations Elon Musk's **SpaceX**, Jeff Bezos's **Blue Origin** and Richard Branson's **Virgin Galactic.** The clash between these three people is famously known as 'Battle of Billionaires'. A 'Race to Space' challenge started after these three entered the Space Industry with their respective organizations. BBC News quoted "Jeff Bezos and Elon Musk don't simply want us to get us to space, instead they want us to stay there". This analysis is about which space organization is influencing the most.

Data Collection

Twitter is used as the social media platform to extract the data for analysis. The official pages of SpaceX (@SpaceX), Blue Origin (@blueorigin) and Virgin Galactic (@virgingalactic) are exercised to make the analysis. A Rival Network is built based on the users who three organizations are following (Friends: In Twitter Language).

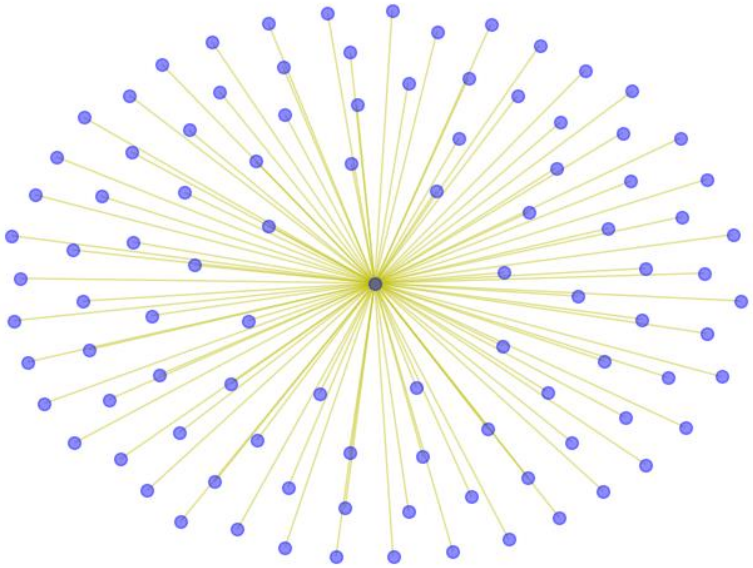
A Python Package **tweepy** is utilized to extract the data from the twitter API. Users information is obtained by using the **get_user** method. The ids of the friends are stored in a separate list for each of the three users using the **friends_ids** method. By extracting the common friends from the three users, a network is built.

Data Visualization

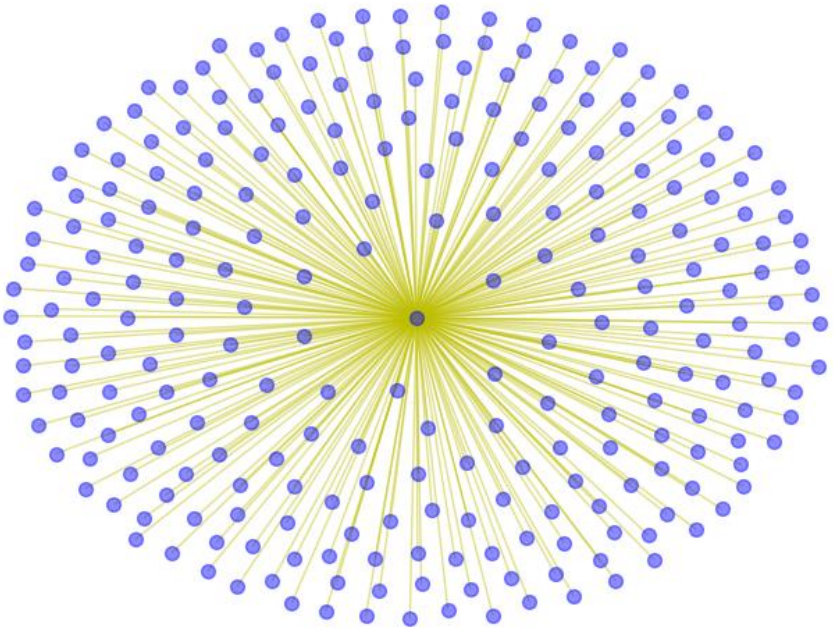
The collected data is used to visualize the data in the form of a network. A Python Package **networkx** is used to form a network from the collected data. Networkx uses the node list and edge list data to create a network.

Visualizations

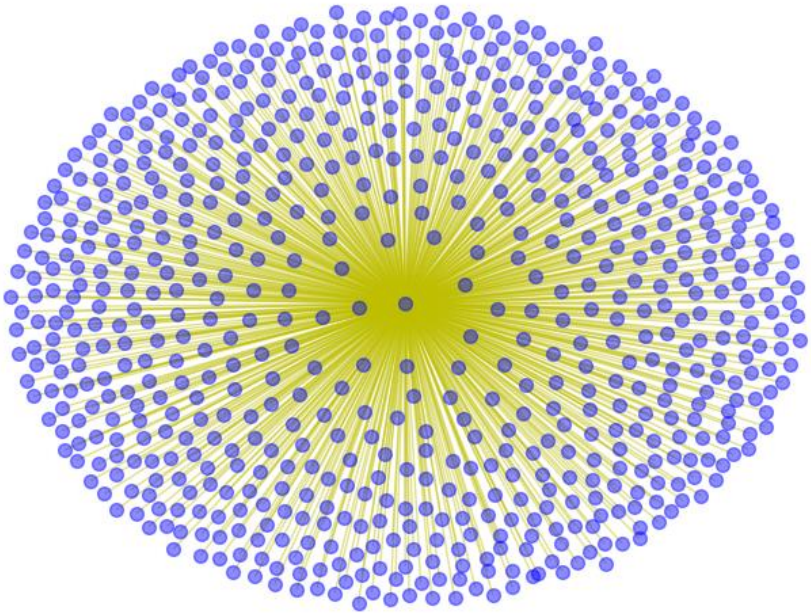
- Friends of SpaceX



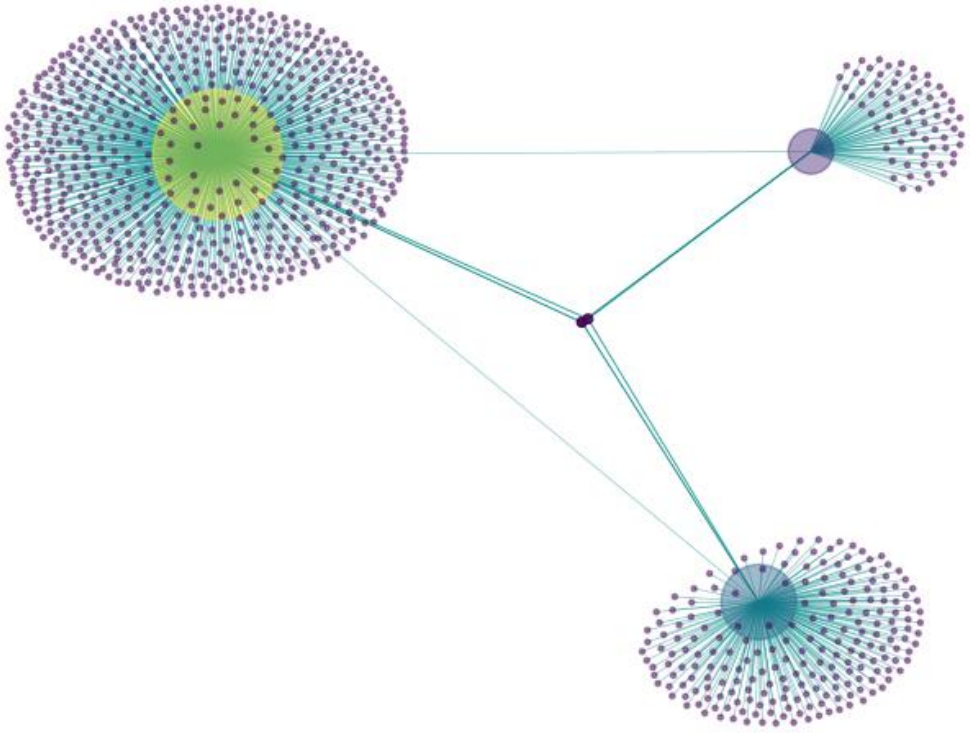
- *Friends of Blue Origin:*



- *Friends of Virgin Galactic:*



- *Graph with shared users:*

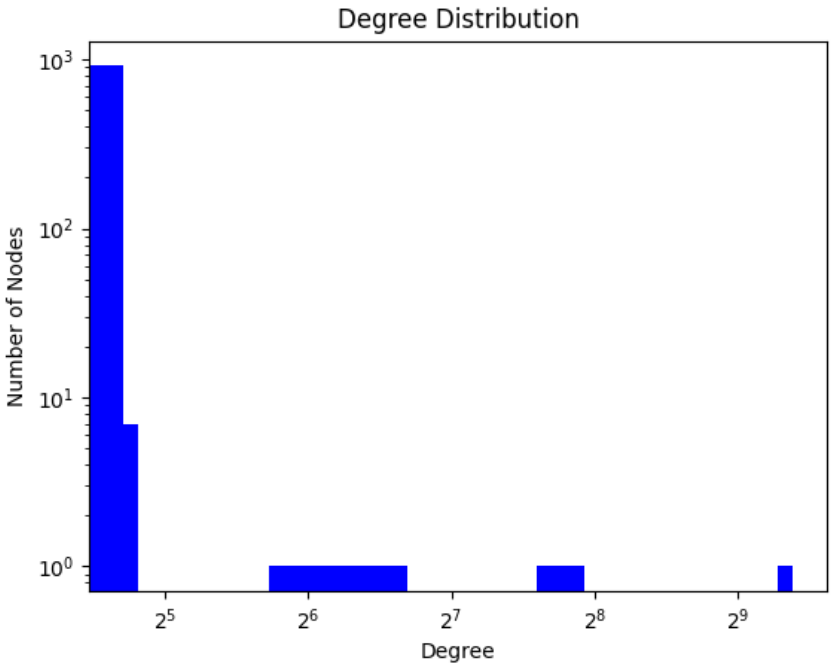


Network Measures Calculation

- *Degree Centrality*

This value is the score of importance based on number of links each node have. This measure is helpful to find more connected individuals, popular individuals or users who can quickly connect with the wider network.

Degree Distribution Graph



- *Closeness Centrality*

This score is based on the node 'closeness' to every other node in the network. This measures the calculates the shortest paths between the nodes, then gives a score to every node based on their sum of shortest paths. Closeness Centrality can help find the individuals who are best placed to influence the entire network (good broadcasters).

The closeness centrality measure is calculated for every node in the network and the node with the highest value is taken.

The highest closeness centrality value is **0.7678275290215588** for node **26208862**. The screen name link ed to the node is **'Virgin Galactic'** indicating it as a good broadcaster than anyone in the node.

- *Betweenness Centrality*

This score measures the number of times a node lies in the shortest path between other nodes thereby acknowledging the bridge nodes in the network. Betweenness Centrality can help identify the individuals who influence the system as a whole.

In the same way, betweenness centrality is calculated for every node and the node with the highest value is taken.

The highest betweenness centrality value is **0.9072176171852198** for node **26208862.** The result is 'Virgi n Galactic' being the most influential user in the entire network as it has the highest betweenness centra lity.

- *Page Rank*

This measure gives score to the nodes based on their connections and connection's connections uncovering nodes who influence beyond the direct connections to the complete network. Page Rank can be helpful for understanding authority and citations.

The highest page rank value is **0.30725528448682704** for node **26208862** which resembles 'Virgin Galact ic'.

Conclusion

Based on the closeness centrality and betweenness centrality values 'Virgin Galactic' is the most influential user as well as a good broadcaster compared to both SpaceX and Blue Origin which means the tweets made by Virgin Galactic reaches more users and creates an impact on people.

References

For Measures

https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/

https://mathinsight.org/degree_distribution

https://www.geeksforgeeks.org/network-centrality-measures-in-a-graph-using-networkx-python/

For networks

https://networkx.github.io/

https://www.cl.cam.ac.uk/~cm542/teaching/2011/stna-pdfs/stna-lecture11.pdf

For Data Collection

https://www.tweepy.org/

http://docs.tweepy.org/en/v3.5.0/getting_started.html