Sagar Patni.
1213217718

1) Explore the data set, convert to proper format and generate a network visualization

- The column names needed to be modified from "From, To" to "Source, Target" which allows gephi to detect the edges.

2)

a) Do you find communities (modularity)?

Resolution = 0.5

Number of communities = 2

Modularity: 0.07

On decreasing the resolution value I found two communities which are visible on color coding the graph, one of which are the nodes strongly connected to node 'Sharon'(1) and other is, nodes which are strongly connected to node 'Yihan'(2). The node 'Rohith' is connected to both 'Sharon' and 'Yihan' with weights 8.0 and 1.0 but as weight of connection 1 is larger it belongs to community 1.

b) What is the average degree centrality? What does it mean in this email network?

The degree centrality of a node is count of in degree and out degree. In the email network degree centrality the degree centrality shows the importance of the node. The average degree centrality shows that on an average how many emails does the person sent/receive.

3.68

c) What is the average betweenness centrality? What does it mean in this email network?

The betweenness centrality represents the degree of which nodes stand between each other. The higher betweenness centrality would have more control over the network, because more information will pass through that node.

The betweenness centrality for nodes 'Sharon' and 'Yihan' are 403 and 41 respectively while its 0 for all other nodes. Which means that all the other nodes are connected to either node 'Sharon', 'Yihan' and there are no connections in between them.

In the email network, it means that most of communication flows through nodes 'Sharon', 'Yihan'. If two nodes need to communicate they have their communication through nodes 'Sharon', 'Yihan'.

17.6

Sagar Patni.
1213217718

d) What is the average Eigenvector centrality? What does it mean in this email network?

0.215108

The eigenvector centrality is sum of centrality values of the node that it is connected to.
In email network, The eigen vector centrality can be larger if the node has many neighbors or it has important neighbors. In directed graph the centrality is given by number of nodes pointing towards node.
In the email network the nodes are either connected to 'Sharon' or 'Yihan' 0 centrality means uni directional email communication to nodes to 'Sharon' or 'Yihan'. Centrality 1 means the node 'Sharon' has path to every other node. The eigenvector centrality is number of nodes pointing towards node the average value can be ascribed to weighted frequency of incoming emails.

e)
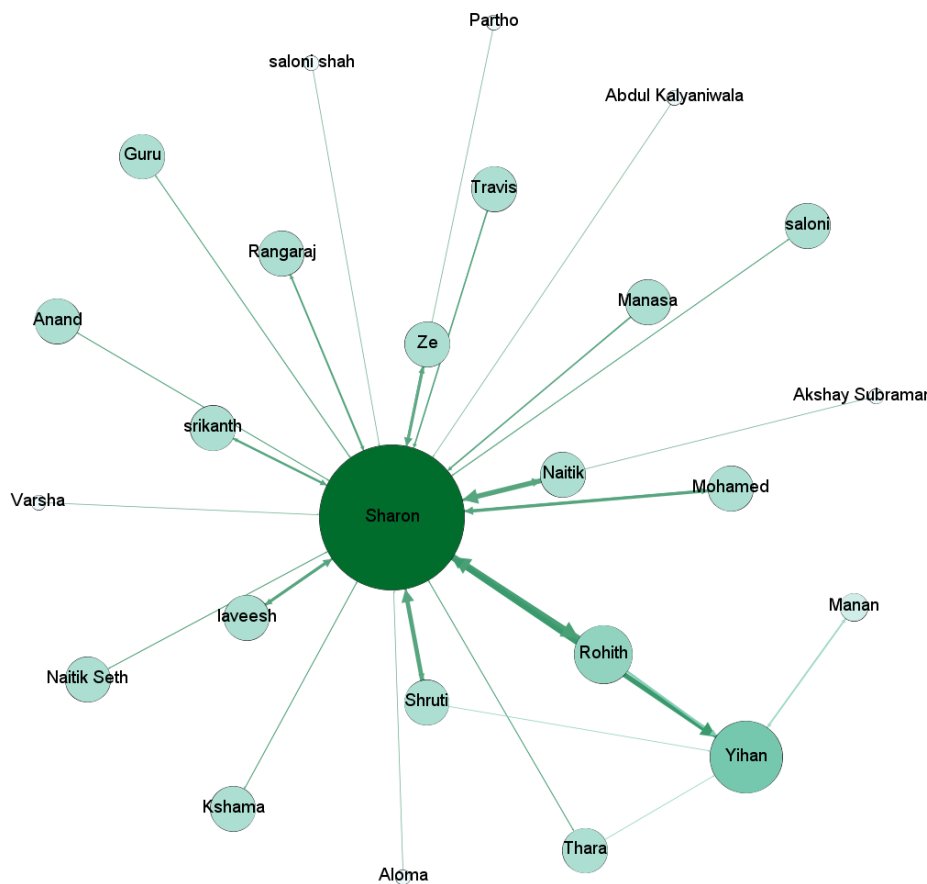
Sagar Patni.
1213217718

Layout: Fruchterman Reingold

It's a force directed layout. It clearly separates the communities, the weights of the edges are clearly visible with distance between nodes. The frequency of the email can be seen with the connection thickness between nodes.