ATHENS UNIVERSITY OF ECONOMICS & BUSINESS

DEPARTMENT OF MANAGEMENT, SCIENCE & TECHNOLOGY

MSc BUSINESS ANALYTICS

“Assignment for Course: Social Network Analysis”

Full Name: STAMATIOS SIDERIS

Register Number: f2822113

Athens,2022

Task 1

Your first task is to create an igraph graph1 using the network of the characters of 'A Song of Ice and Fire' by George R. R. Martin [1]. A .csv file with the list of edges of the network is available online.2 You should download the file and use columns Source, Target, and Weight to create an undirected weighted graph. For your convenience, you are free to make any transformations you think are appropriate to the file.

Answer:

We read the csv files of edges and nodes provided by github. The edges will give us the relationships between the characters and the nodes will give us the labels of the vertices. Using the igraph library, we create the graph setting the attribute directed equal to FALSE.

Task 2

Next, having created an igraph graph, you will explore its basic properties and write code to print:

- Number of vertices
- Number of edges
- Diameter of the graph
- Number of triangles
- The top-10 characters of the network as far as their degree is concerned
- The top-10 characters of the network as far as their weighted degree is concerned

Answer:

- The number of vertices is equal to 796.
- The number of edges is equal to 2823.
- The diameter of the graph is equal to 53.
- The number of triangles created is equal to 5655. We divided the number of counted triangles by 3 to get the number of unique triangles as each triangle consists of 3 nodes.
- The top-10 characters by degree are:

Tyrion-Lannister	122
Jon-Snow	114
Jaime-Lannister	101
Cersei-Lannister	97
Stannis-Baratheon	89
Arya-Stark	84
Sansa-Stark	75
Catelyn-Stark	75
Robb-Stark	74
Eddard-Stark	74

- The top-10 characters by weighted degree are:

Tyrion-Lannister	2873

Jon-Snow	2757
Cersei-Lannister	2232
Joffrey-Baratheon	1762
Eddard-Stark	1649
Daenerys-Targaryen	1608
Jaime-Lannister	1569
Sansa-Stark	1547
Bran-Stark	1508
Robert-Baratheon	1488

Task 3

After that, your task is to plot the network: You will first plot the entire network. Make sure you set the plot parameters appropriately to obtain an aesthetically pleasing result. For example, you can opt not to show the nodes' labels (vertex.label = NA) and set a custom value for parameters: edge.arrow.width, and vertex.size. Feel free to configure additional parameters that may improve your visualization results. Then, you will create a subgraph of the network, by discarding all vertices that have less than 10 connections in the network and plot the subgraph. In addition to the above plots, you are also asked to write code that calculates the edge density of the entire graph, as well as the subgraph, and provide an explanation on the obtained results (a few sentences in your report).

Answer:

We created the graph of the entire Network in **Figure 1**. The density of the graph in **Figure 1** is equal to 0.0089.

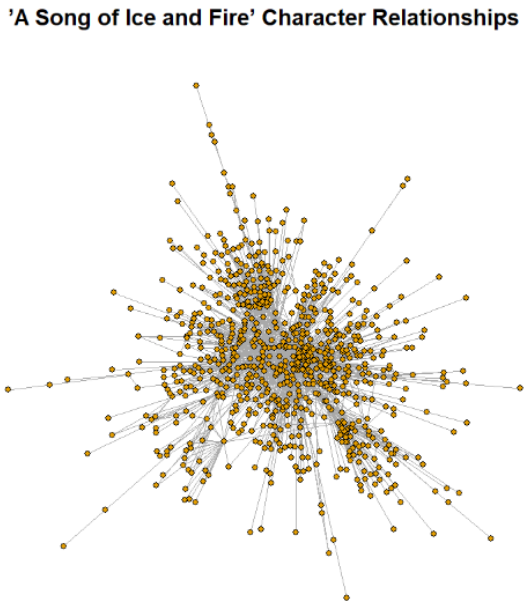

Figure 1 Plot of the entire Network

We created the subgraph of the Network discarding the vertices that have less than 10 connections, in **Figure 2**. The density of the subgraph in **Figure 2** is equal to 0.1258.

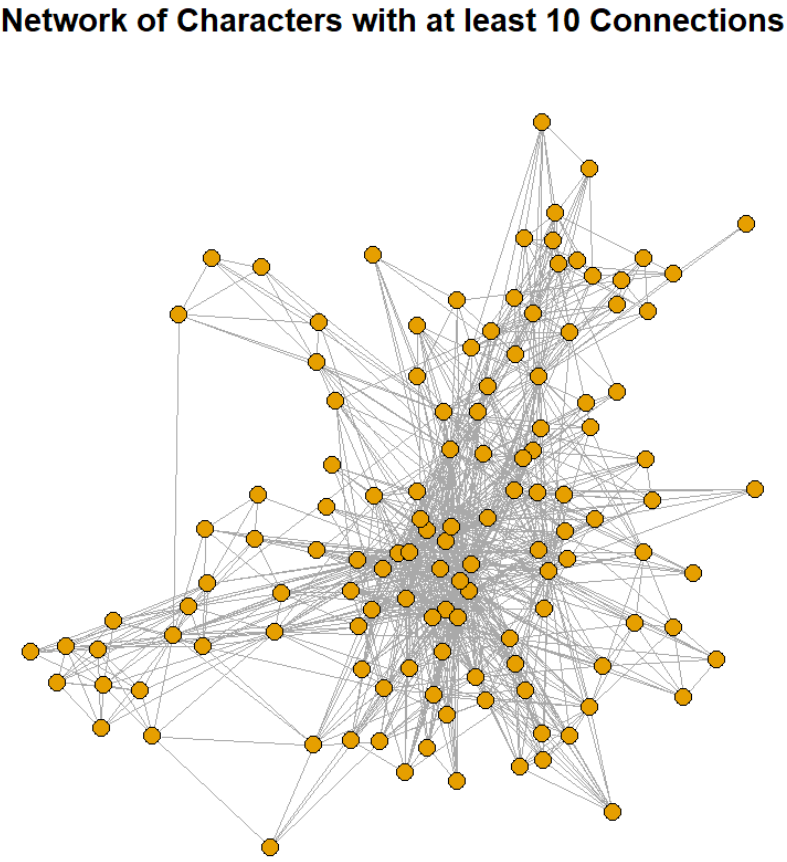

Figure 2 Plot of the Network for vertices with at least 10 connections

The density of a graph is the ratio of the number of edges and the number of possible edges. By excluding vertices with lower than 10 connections to create the subgraph, the ratio is increased compared to the entire network. This happens because the number of edges will decrease a little as the number of connections for the excluded vertices is low, compared to the decrease in the number of possible edges which is great as it represents the possible connections of the excluded vertex with the entire group of vertices.

Task 4

Next, you will write code to calculate and print the top-15 nodes according to the:

- closeness centrality
- betweenness centrality

In addition, you are asked to find out where the character Jon Snow is ranked according to the above two measures and provide an explanation (a few sentences) of the observations you make after examining your results.

Answer:

The top-15 nodes according to the closeness centrality are:

Jaime-Lannister	0.0001205982
Robert-Baratheon	0.0001162791
Stannis-Baratheon	0.0001146921
Theon-Greyjoy	0.0001146132
Jory-Cassel	0.0001141553
Tywin-Lannister	0.0001137656
Tyrion-Lannister	0.0001130071
Cersei-Lannister	0.0001129688
Brienne-of-Tarth	0.0001124480
Jon-Snow	0.0001118944
Joffrey-Baratheon	0.0001105094
Rodrik-Cassel	0.0001103631
Eddard-Stark	0.0001092180
Doran-Martell	0.0001088613
Robb-Stark	0.0001088495

The top-15 nodes according to the betweenness centrality are:

Jon-Snow	41698.94
Theon-Greyjoy	38904.51
Jaime-Lannister	36856.35
Daenerys-Targaryen	29728.50
Stannis-Baratheon	29325.18
Robert-Baratheon	29201.60
Tyrion-Lannister	28917.83
Cersei-Lannister	24409.67
Tywin-Lannister	20067.94
Robb-Stark	19870.45
Arya-Stark	19354.54
Barristan-Selmy	17769.29
Eddard-Stark	17555.36
Sansa-Stark	15913.44
Brienne-of-Tarth	15614.41

Jon Snow is ranked 1st by closeness centrality and 10th by betweenness centrality. It seems that Jon Snow is close to many of the characters and quite efficient on spreading information through the graph but not through the shortest available path for some of them while he has the greatest influence over the flow of information in the graph.

Task 5

In the final step of this project, you are asked to rank the characters of the network with regard to their PageRank value. You will write code to calculate the PageRank values and create a plot of the graph that uses these values to appropriately set the nodes' size so that the nodes that are ranked higher are more evident.

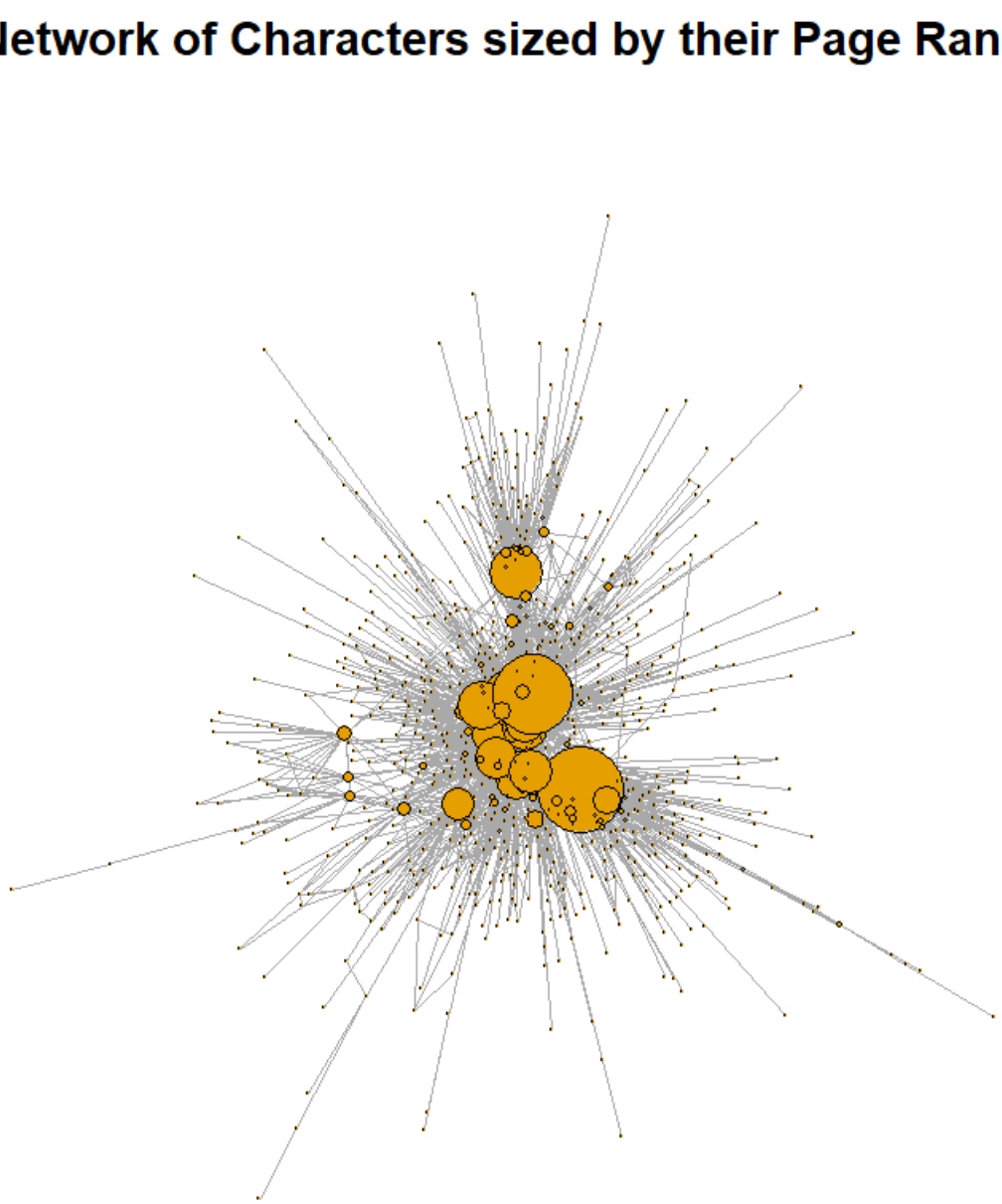

Figure 3 Plot of the Network with vertices sized by their page rank