9/6/2019

‘A Song of Ice and Fire’ network

Assignment 1

Social Network Analysis

Athens University of Economics and Business

Brikena Kokalari

Contents

1. Introduction

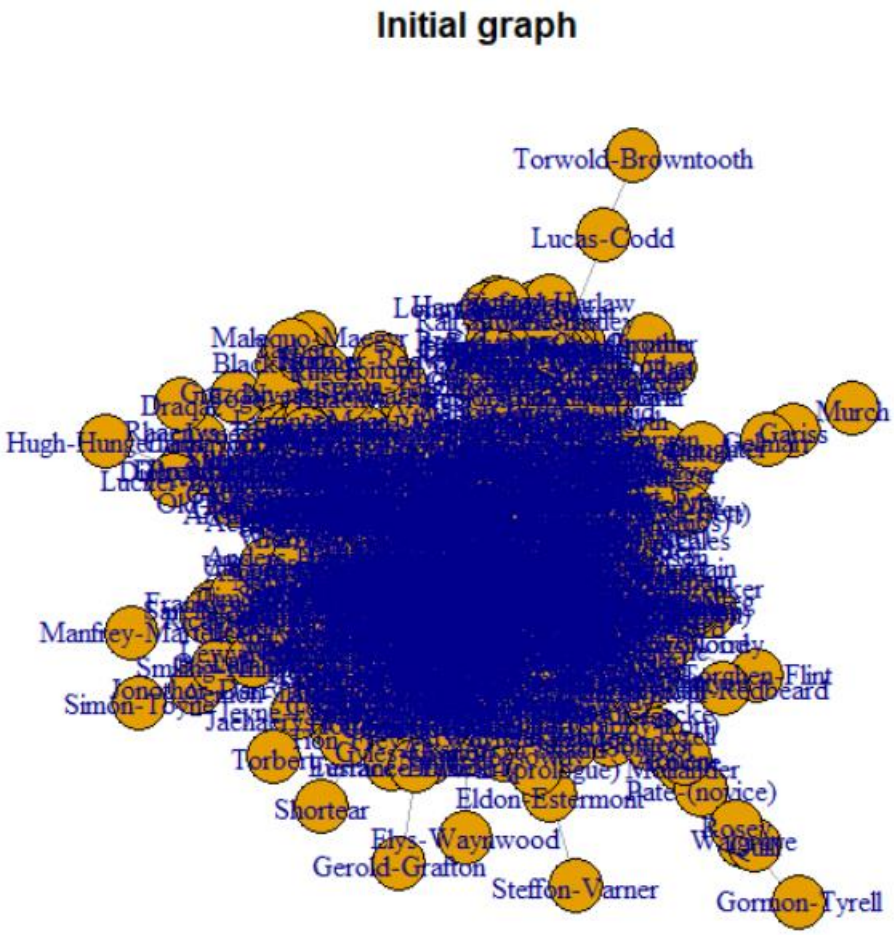
The aim of this assignment is to create an igraph graph using the network of the characters of 'A Song of Ice and Fire' by George R. R. Martin. We will use as a dataset a csv file with the list of edges of the network which are available online (https://github.com/mathbeveridge/asoiaf/blob/master/data/asoiaf-all-edges.csv).

Regarding the social network we are going to measure not only the nodes importance but also the nodes properties. In addition, we will create a subgraph of the network keeping only those vertices that have more than 10 connections (edges) and then visualize the result.

Last but not least, we will examine the rank of the characters of the network as far as their PageRank value is concerned. Taking into consideration the PageRank value, we will create a plot of the graph so as the nodes that are ranked higher are more evident.

2. Q1

For the creation of the graph, we used only the three columns Source, Target, and Weight to create an undirected weighted graph. Below we plotted the initial graph in R with the simple command *plot()* , which aesthetically is not very pleasant and pretty.



3. Q2

After the creation of the igraph graph, we examined the basic network properties, writing the code in R .

Network Properties	Outcome
Number of vertices	796
Number of edges	2823
Diameter of the graph	53
Number of triangles	16965

- The top-10 characters of the network as far as their degree is concerned

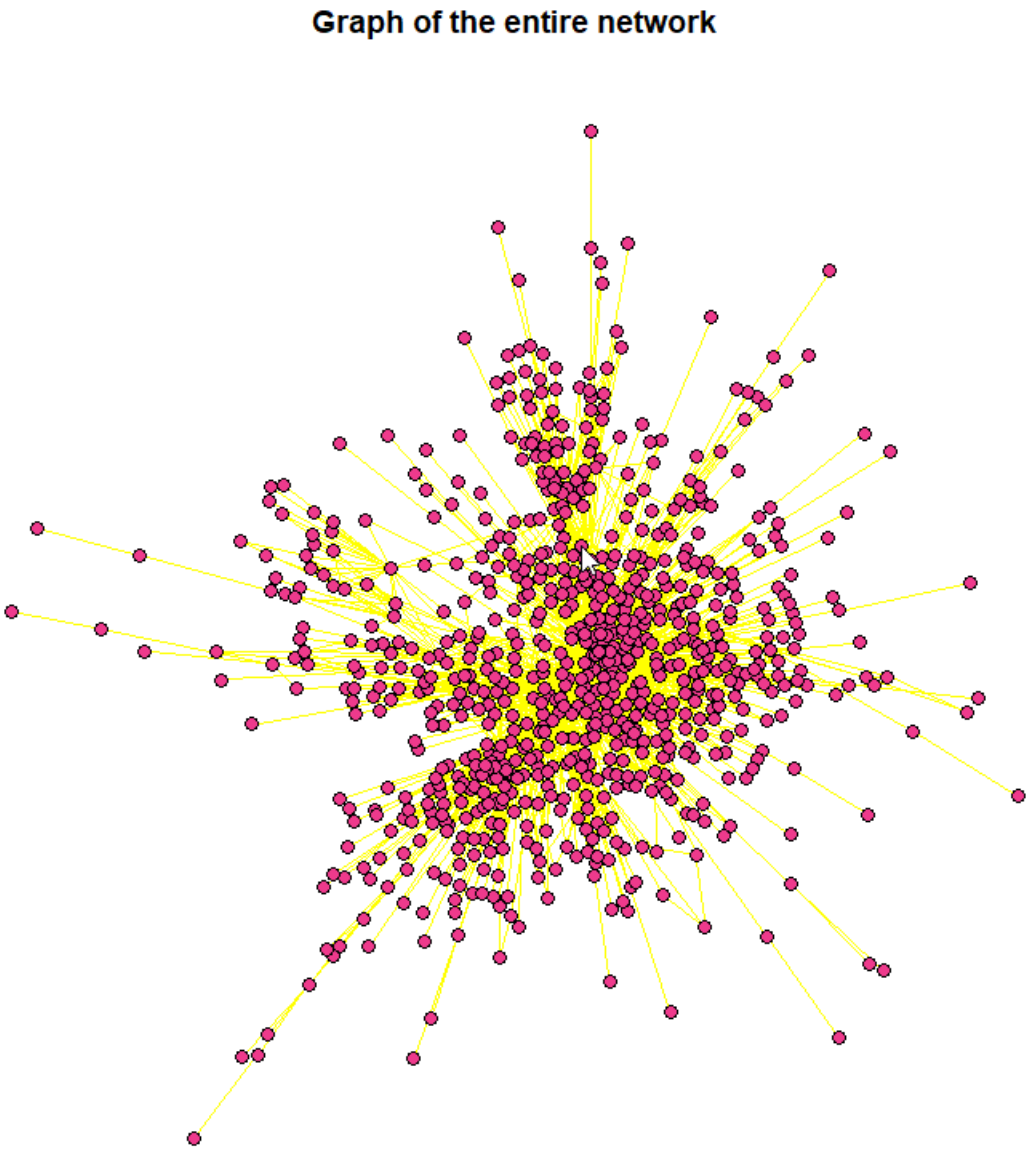
```
Tyrion-Lannister          Jon-Snow   Jaime-Lannister  Cersei-Lannister Stannis-Baratheon
             122               114               101                97               89
      Arya-Stark     Catelyn-Stark       Sansa-Stark      Eddard-Stark       Robb-Stark
              84               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

- The top-10 characters of the network as far as their weighted degree is concerned

```
  Tyrion-Lannister          Jon-Snow   Cersei-Lannister  Joffrey-Baratheon       Eddard-Stark
              2873              2757              2232               1762               1649
Daenerys-Targaryen     Jaime-Lannister       Sansa-Stark        Bran-Stark   Robert-Baratheon
              1608              1569              1547               1508               1488
```

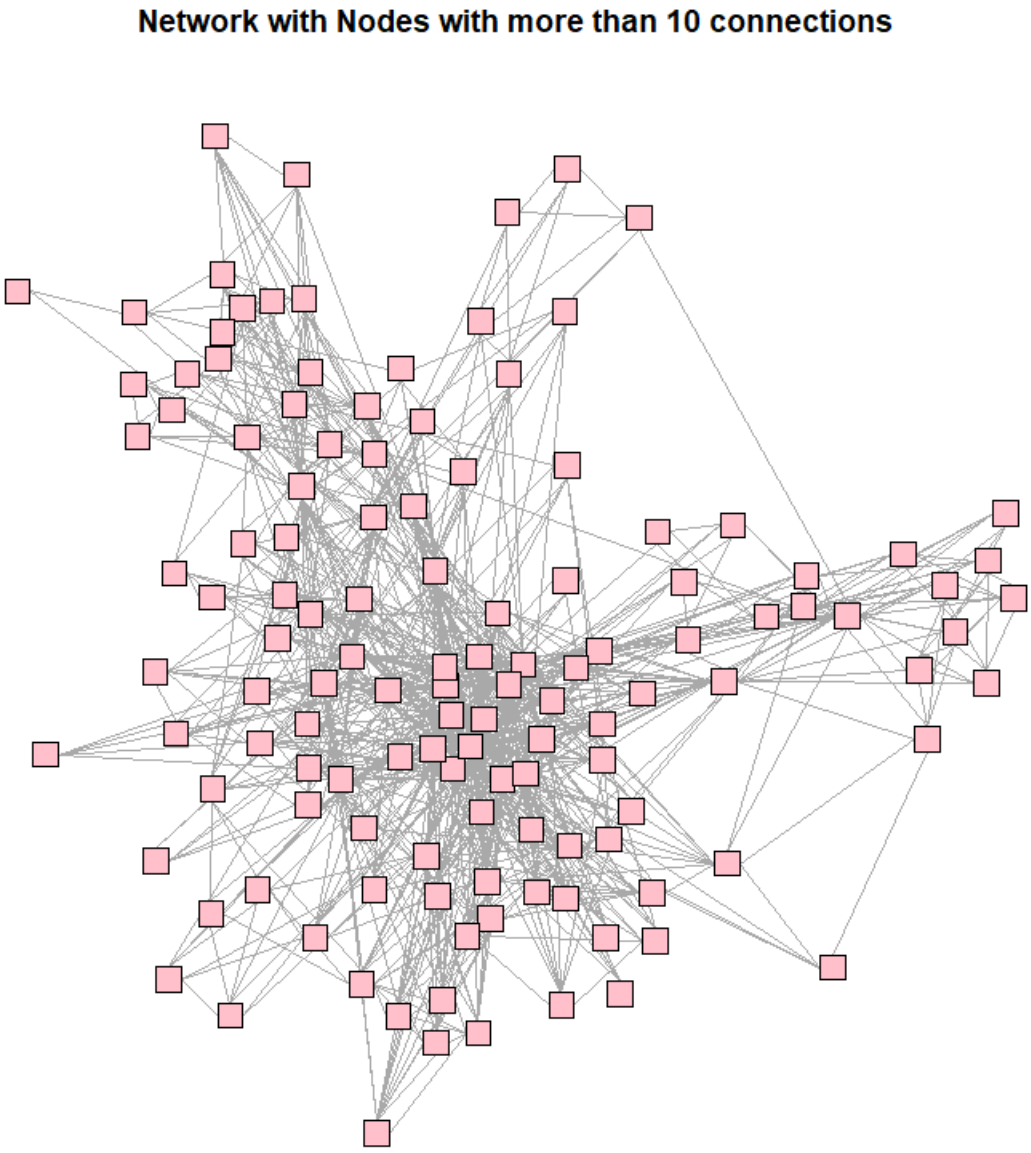
4. Q3

In this step we are going to plot the entire network, setting some parameters appropriately to have a more aesthetically pleasing result.



By eliminating al the verticals that have less than 10 connections, we can create a subgraph, smaller graph than the initial. To obtain such a graph, we just have to set the degree to be greater than 10 in our R code.

The resulting subgraph is the below.



The edge density of a graph is the ratio of the number of edges and the number of possible edges.
Firstly, we checked for loops in the graph. A loop (also called a self-loop or a "buckle") is an edge that connects a vertex to itself. The graph did not contain any loops.
For the initial graph the edge density is approximately 0.0089, while for the subgraph the edge density is approximately 0.1259.
For undirected simple graphs, the graph density is defined as:

$$D = \frac{2|E|}{|V|\left(|V|-1\right)}$$

Since the edge density is the proportion of present edges from all possible edges in the network, it makes sense that this ratio would have a higher value in a graph with less edges. In

our example, we are discarding all vertices that have less than 10 connections in the network, making the denominator of the fraction smaller and the ratio greater, since the denominator is calculated in square power.

5. Q4

In the next step we are going to calculate the top 15-nodes according to:

- closeness centrality
- betweenness centrality

In an undirected graph, closeness centrality (or closeness) of a node is a measure of centrality in a network, calculated as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the graph. Thus, the more central a node is, the closer it is to all other nodes. It's a measure of how long information takes to arrive (who hears news first?). Higher values mean less centrality.

1. We calculated the normalized centrality. Normalization is performed by multiplying the raw closeness by n-1, where n is the number of vertices in the graph.

Gormon-Tyrell	Walgrave	Rosey	Waif	Mord
0.02735907	0.02980207	0.03069024	0.03148889	0.03198938
Kindly-Man	Will-(prologue)	Quill	Haggon	Kerwin
0.03476930	0.03542781	0.03624510	0.03847643	0.03919925
Murch	Thistle	Dolf	Pyat-Pree	Malaquo-Maegyr
0.03986961	0.04004433	0.04017993	0.04103438	0.04124942

2. The raw closeness centrality without normalization is shown below.

Gormon-Tyrell	Walgrave	Rosey	Waif	Mord
3.441393e-05	3.748688e-05	3.860408e-05	3.960867e-05	4.023821e-05
Kindly-Man	Will-(prologue)	Quill	Haggon	Kerwin
4.373497e-05	4.456328e-05	4.559132e-05	4.839803e-05	4.930723e-05
Murch	Thistle	Dolf	Pyat-Pree	Malaquo-Maegyr
5.015045e-05	5.037022e-05	5.054079e-05	5.161557e-05	5.188606e-05

In both cases we obtain the same results but with different values of closeness due the formula of calculation.

Betweenness centrality is a measure of centrality in a graph based on shortest paths. For every pair of vertices in a undirected graph, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through (for unweighted graphs) or the sum of the weights of the edges (for weighted graphs) is minimized. The betweenness centrality for each vertex is the number of these shortest paths that pass through the vertex.

The betweenness centrality of top 15 nodes are shown below

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

According to the above two measures, the character Jon Snow is ranked first as far as betweenness centrality is concerned and among the characters with a very small value of closeness centrality.

- Jon-Snow Closeness Centrality: 0.0001106072
- Jon-Snow Betweenness Centrality: 41698.94

Almost 0 steps are required to access every other character from Jon Snow, which means that he can easily reach to the rest of network. The number of shortest paths between the characters that pass through Jon Snow is 41699. This means that Jon Snow would have more control over the network, because more information will pass through him and he have been involved between different groups of the network. The more central a character is, the closer it is to all other characters.

6. Q5

Finally, we are going to rank all characters of the network according to their PageRank value and then plot the graph uses these values to appropriately set the nodes' size so that the nodes that are ranked higher are more evident. The plot that arises from the above criteria is shown below.