Notes

Dr. Shirin Glander

July 6, 2017

In this post, I am exploring network analysis techniques in a family network of major characters from Game of Thrones.

What can network analysis tell us?

A network is a graph of interconnected nodes/vertices.

Network analysis can e.g. be used to explore relationships in social or professional networks.

These can give us a lot of information about the patterns of how people interact.

- Who is the most connected (i.e. influential or "important")?
- Are there clusters of tightly connected people?
- Are there a few key players that connect clusters of people?

These answers can give us a lot of information about the patterns of how people interact.

The Game of Thrones character network

The basis for this network is Kaggle's Game of Throne dataset (character-deaths.csv). Because most family relationships were missing in that dataset, I added the missing information in part by hand.

Because the books and the TV series differ slightly, I have introduced edges that are only supported or hinted at by the TV series and are not part of the original narrative in the books.

- Fruchterman-Reingold layout

As we can see, even with only a subset of characters from the Game of Thrones world, the network is already quite big.

What we can see right away is that there are only limited connections between houses and that the Greyjoys are the only house that has no ties to any of the others.

Centrality describes the number of edges that are in- or outgoing to/from nodes. High centrality networks have few nodes with many connections, low centrality networks have many nodes with similar numbers of edges.

Who are the most important characters?

We consider a character "important" if he has connections to many other characters. There are a few network properties, that tell us more about this. For this, I am considering the network as undirected to account for parent/child relationships as being mutual.

Node degree

Node degree or degree centrality describes how central a node is in the network (i.e. how many in- and outgoing edges it has or to how many other nodes it is directly connected via one edge).

"The degree of a vertex is its most basic structural property, the number of its adjacent edges." From the help pages of degree() We can calculate the number of out- or ingoing edges of each node, or - as I am doing here - the sum of both.

Closeness

The closeness of a node describes its distance to all other nodes. A node with highest closeness is more central and can spread information to many other nodes.

Betweenness centrality

Betweenness describes the number of shortest paths between nodes. Nodes with high betweenness centrality are on the path between many other nodes, i.e. they are people who are key connections or bridges between different groups of nodes. In a social network, these nodes would be very important because they are likely to pass on information to a wide reach of people.

Ned Stark is the character with highest betweenness. This makes sense, as he and his children (specifically Sansa and her arranged marriage to Tyrion) connect to other houses and are the central points from which the story unfolds. However, we have to keep in mind here, that my choice of who is important enough to include in the network (e.g. the Stark ancestors) and who not (e.g. the whole complicated mess that is the Targaryen and Frey family tree) makes this result somewhat biased.

Not surprisingly, we learn that House Stark (specifically Ned and Sansa) and House Lannister (especially Tyrion) are the most important family connections in Game of Thrones; they also connect many of the storylines and are central parts of the narrative.