

Final Project Report:
An Exploration of the *Game of Thrones* Social Network
Elizabeth Combs (eac721)

New York University
GPH-GU-2505: Social Network Data Analysis with R
Professor Yang Feng
August 9, 2020

Introduction:

Network analysis applies to a multitude of disciplines from social networks, including social media platforms like Facebook and Twitter, to technological networks like airline routes. Network data can improve decision-making on investments, marketing, operations, and even biology. In addition to nodes and edges, which identify entities and the connections between them, supplemental information can also provide context to the network analysis. For example, Facebook tracks users and their friendship connections, but also has information on users' interests, locations, group-affiliations, and strengths of connections between individuals using mutual friendships. Facebook is able to use this information to improve user experience and drive advertising revenue through targeting (Kolaczyk and Csárdi, 2014, p. 7).

As the world globalizes, networks are becoming more important to understanding how entities are connected. In 1967, The Small World Problem identified on average six degrees of separation between any two individuals in the world (Milgram, 1967). Since then, the creation of the worldwide web and social media networks, this average has continued to drop in recent years. In February 2016, Facebook asserted that "each person...among the 1.59 billion people active on Facebook...is connected to every other person by an average of three and a half other people [3.57 intermediaries]." This figure continues to shrink further as more of the world's population wires up electronically. Only five years earlier, a similar study indicated that the connectedness on Facebook was 3.74 average intermediary connections with about half as many users on the site (Bhagat et. al, 2016).

Facebook is incredibly widespread and represents the quintessential social network as most people would understand that term today; however, social networks are present in all aspects of life, even fictional worlds. To show that these principles apply beyond the most obvious example of Facebook, this paper will explore principles of social networks in George R. R. Martin's 2000 novel, *A Storm of Swords*, the third from his series, *A Song of Ice and Fire*. The series achieved great fame via its HBO television adaptation, *Game of Thrones (GOT)*, airing from 2011 to 2019 (imdb.com, 2020) and breaking

viewership, award-nomination, and pirating records (Swatman, 2017). By using this world of famous characters, this paper aims to make the principles of social networks more accessible and entertaining to a general audience. It conducts an exploratory analysis of a fictional social network to (1) share a data overview (2) review an analysis of important nodes known as hubs (3) build visualizations known as network layouts (4) explore possible community structures within the network. This analysis shows how social networks can not only effectively summarize the connections between the *GOT* characters, but also add insight that goes beyond how many fans think about them.

Game of Thrones Background & Network Suitability

A Storm of Swords takes place in the fictional worlds of Westeros and Essos, where a group of noble families vie for wealth, power, and glory in a realm that resembles the European middle ages in the timeline of human civilization. The king governs the continent of Westeros from the capital city of King's Landing, and different noble houses govern different areas of the country on his behalf. In the north of Westeros, the Night's Watch man the Wall, which shields the realms of men from the Wildlings, barbarians who live north of the Wall. In the east, the continent of Essos exists largely isolated from Westeros apart from some commercial ties cultivated by merchants.

Political and social relationships are everything in this universe. Family ties are typically strong in a feudal world where political power is transferred by blood, and marriage is the best way to cement those ties. While geography is the main determinant of which characters know and interact with others, many characters traverse the world on their own individual stories and adventures, allowing the varied ensemble of main and secondary characters to dynamically build and alter their web of social connections. In this process, political connections and allegiances evolve as different characters acquire and lose power. The complex interplay of family, geography, power, wealth, and love makes the *GOT* a rich universe to explore with social network analysis.

Data Source and Structure

The data is sourced from a GitHub repository of sample social network datasets. The GOT dataset is retrieved from the [game-of-thrones repository](https://github.com/melaniewalsh/sample-social-network-datasets). The dataset is divided into two comma-separated values (CSV) files, representing nodes and edges of the social network, respectively (Melanie Walsh, 2018, <https://github.com/melaniewalsh/sample-social-network-datasets>).

Each node in the dataset corresponds to an individual character from the novel *A Storm of Swords*; there are 107 nodes in total. Edges between characters were computed based on the number of times two characters' names appeared within 15 words of one another in the book's text. It is important to note that "an edge between two characters ... means that they interact, speak of one another, or are mentioned together" rather than that they are directly "friends" (Beveridge and Shan, 2016). In total, there are 353 edges between characters. The edges assume mutual connection between the two characters rather than a source-to-target specific pairing, which indicates that this network is undirected. Additional steps would be required in the data generation process in order to build a directed network, which could be completed in future works. Figure 1 below gives an initial view of the social network.

In addition to the data from the website, the author also added information based on personal knowledge and that was cross-checked via an online source (gameofthrones.com, 2015). These variables allow for additional analysis based on characteristics of key characters, including gender, family name, geography, and political lean. These attributes allow for new analysis to be conducted based on a character's categorization. Later analyses and visualizations will utilize these attributes to explore additional aspects of the *GOT* social network.

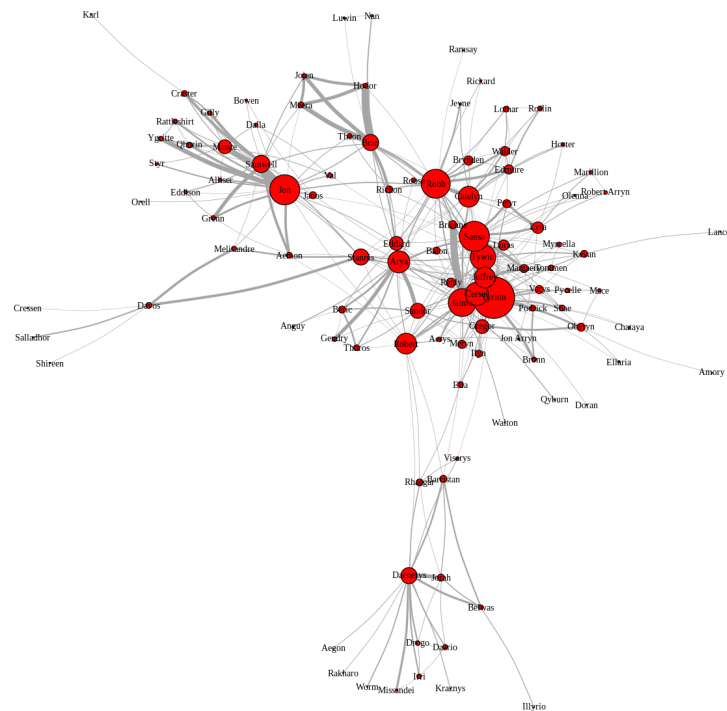


Figure 1. An initial view of the GOT social network where each node represents a character and node size corresponds to the degree and edge size corresponds to the weight.

Analysis & Results:

Graph Characteristics

The analysis begins by reviewing overall characteristics of the network graph. The average degree of separation of the GOT social network is 6.58. Notably, this network approximates the same closeness that was observed by Stanley Milgram in his famous experiment. There could be multiple explanations for the average degree being around the commonly referenced six degrees. First, interest in television programs comes from the way that they mirror, and even become, our own social networks. Jennifer Barnes, an assistant professor of psychology at the University of Oklahoma, told TIME magazine that watching scripted dramas on television has real-world benefits similar to friendships such as boosting

self-esteem and decreasing loneliness because “our brains aren’t really built to distinguish between whether a relationship is real or fictional” (Macmillan, 2017). Furthermore, *GOT* is set in an age preceding our own, so social networks would be smaller compared to today’s Facebook network, limiting the number of connections that could be made. The *GOT* social network is also well-spread due to geographic factors that play a large role in how the story plays out.

The degree distribution follows the power law distribution, making the *GOT* network a scale-free distribution. The median degree of the network is 4. Thus, the difference between the average degree and median degree also supports the skewness of the distribution shown in Figure 2. Even examining the degree distribution by gender, the power law persists with certain nodes having a high frequency compared to others.

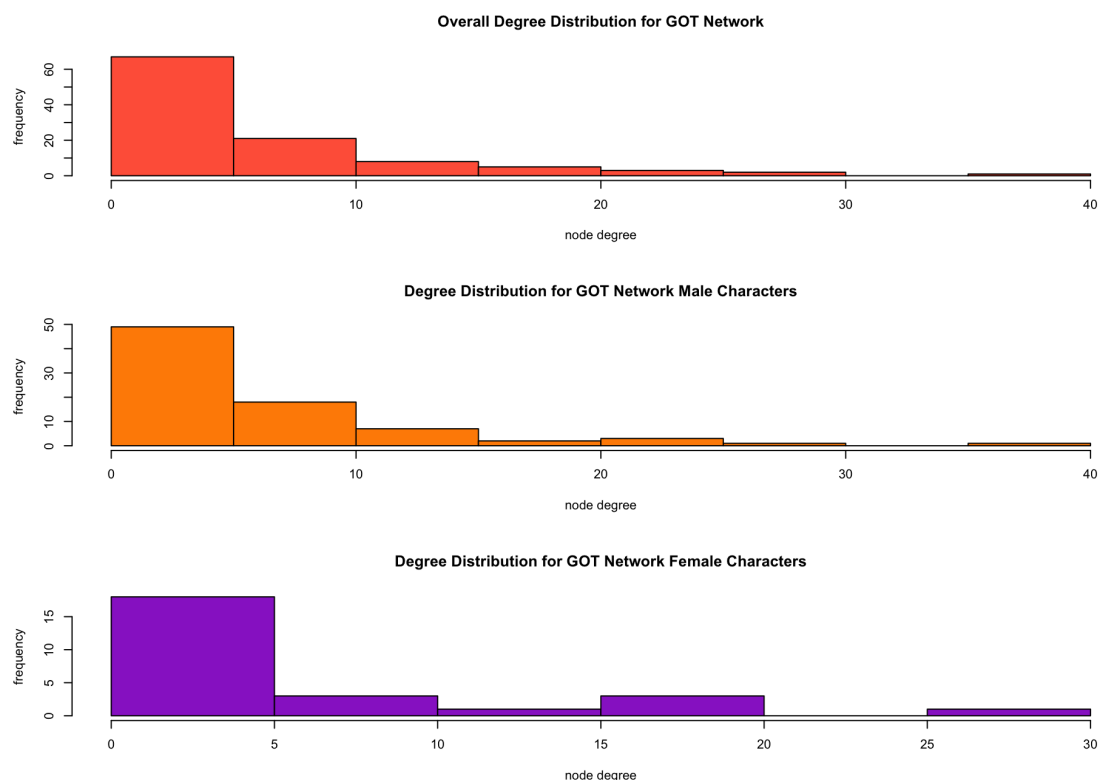


Figure 2. The degree distribution of the *GOT* social network is skewed at the overall and gender-specific levels.

Important Nodes:

A hub node has a large degree compared to most nodes within the network. In decreasing order of degree centrality, the top 15 characters by connections are Tyrion, Jon, Sansa, Robb, Jaime, Tywin, Cersei, Arya, Robert, Joffrey, Catelyn, Samwell, Bran, Stannis, and Daenerys. Tyrion has the most connections to other characters with 36 connections, while Daenerys has 14 connections. The gender distribution between the top 15 characters is 2:1, with 10 popular characters being male while 5 are female. Examining the entire network, this ratio decreased to 4:1, with only about 25% of characters being female (81 males, 26 females). The gender ratio in *GOT* favors males; however, females have a more important presence amongst the connected characters. Supporting characters that are less well connected tended to be male within this network. George R.R. Martin has been noted for writing “strong female characters with whom [we] can identify”, especially considering the dominance of men within the medieval-like setting (Delphine, 2011). Figure 3 shows the full social network as well as a subgraph of the network to illustrate this phenomenon.

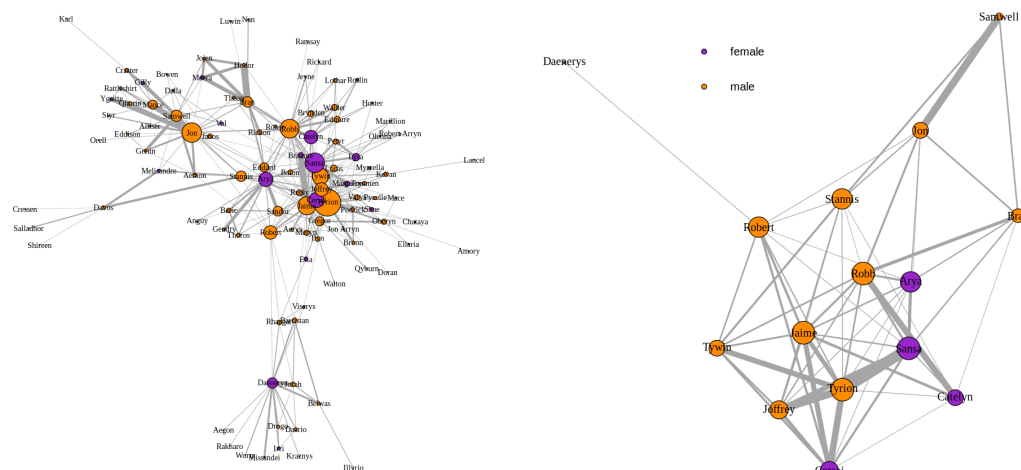


Figure 3. In each graph, nodes are colored based on the author-generated gender variable. On the left, the chart shows the full network with 81 males and 26 females. The right network is a subgraph of the first with weights and degree recomputed for only the top 15 characters by degree to zoom in and better understand this phenomenon.

Based on geodesic distance, closeness centrality of characters in the *GOT* network is similar to degree centrality. Tyrion, Sansa, and Robert have the highest normalized closeness centrality (0.512, 0.510, 0.500). However, Daenerys falls much lower in the ranking of closeness (rank: 42/107). Paths through Daenerys are not as direct as those of some of the other key characters; thus, this node has a higher geodesic distance and a lower closeness centrality (0.366). This might be attributed to Daenerys's geographic distance from the other characters, most of whom have never met or interacted with her. Most other characters speak of her only in the abstract, based on her historical claim to the throne and the history of her family, which was banished from Westeros long before the events of the book take place. In terms of betweenness, Jon, Robert, Tyrion hold the top spots (0.300, 0.209, 0.198). Jon's high ranking might be based on his prominence within the powerful Stark family as well as his being the principal character at the Wall and North of the Wall. The Wall and the area North of the Wall have incredible prominence in the book because of the existential threat of the Others / the White Walkers, but at the same time these events are viewed as peripheral to most characters elsewhere. This dynamic is often compared to current events and climate change in a popular fan theory (Segall, 2018).

This makes Jon a very prominent link for most of the characters to a central issue of the novel that is only experienced by a handful of characters. This also makes sense for Robert, the king who died in the first novel and whose legacy still impacts every character in the third novel, despite his physical absence. Tyrion is well known for his self-described talents, "I drink, and I know things," and most fans of the series would agree with his status as the character with the third-most betweenness. Since betweenness is indicative of passing information, Daenerys has the fourth highest betweenness measure (0.157). This also makes sense considering that her presence lurks in the background of every conversation about who should rightfully sit on the Iron Throne, her own claim having been temporarily

extinguished by the war that took place right before the events of the book. Even an initial network plot shows the interesting position Daenerys holds (Figure 1).

Deep diving on Daenerys who holds an interesting position in the *GOT* social network, her path to other main characters tends to be longer than other principal characters' similar paths to other main characters. For instance, from Daenerys to Robb, the path is highlighted in gold in Figure 4.

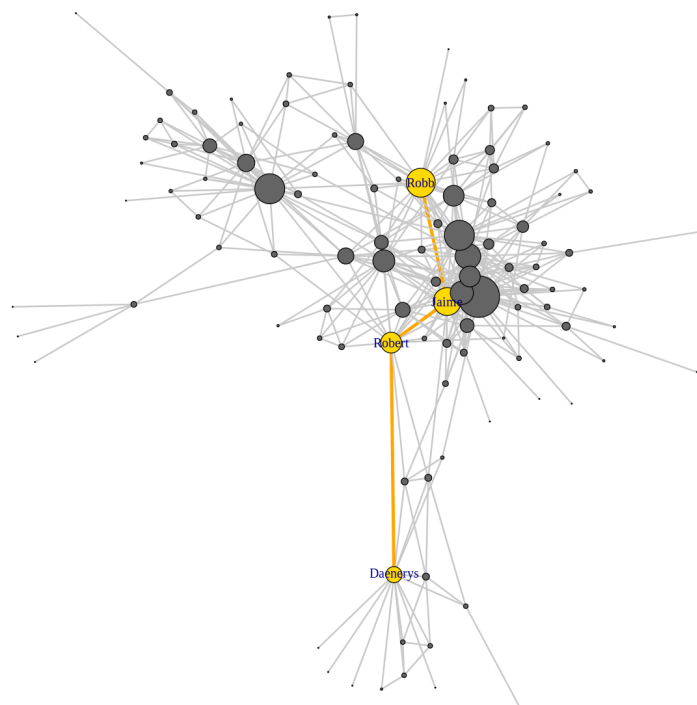


Figure 4. This view of the network highlights Daenerys' path to another major character, Robb. The most direct path to Robb is highlighted in yellow while all other nodes are depicted in gray. This path runs Daenerys-Robert-Jaime-Robb and is longer than the path between most other main characters.

The length of the path shown in Figure 4 makes a lot of sense considering that Daenerys is so geographically distant from the other central characters. As mentioned earlier, Daenerys exists mostly in the abstract to the rest of the characters in the book. While five potential kings in Westeros fight for the throne, Daenerys theoretically has the most direct claim on the throne because her father was king before

he was killed by Jaime Lannister, and subsequently Robert Baratheon took the throne for himself. Thus, it seems logical that Daenerys's most direct tie to the other characters is through Robert Baratheon, whom she appropriately calls "the usurper." Daenerys is not a person to the other characters, but she is a threat to their legitimacy based on her relationship to Robert. Note that Daenerys's path to Robb, who would proclaim himself King in the North, after Robert, runs through Jaime, the very man who killed her father Aerys, and who also has many direct interactions with Robb during the Wars of the Five Kings. In *A Storm of Swords*, Daenerys still hasn't left Essos, and so her main tie to Westeros is through Robert Baratheon, another character who no longer has a direct presence in Westeros (albeit because he is dead), and her relationship to other important principal characters is more in the abstract web of political ties rather than in the concrete realm of social interaction.

Visualizations & Discussion:

It is tempting first to try to use geography to make sense of our social network graph, as *GOT* takes place in a richly detailed fantasy world with a unique geography. Figure 5 depicts the network with this view, where nodes are colored by their geography.

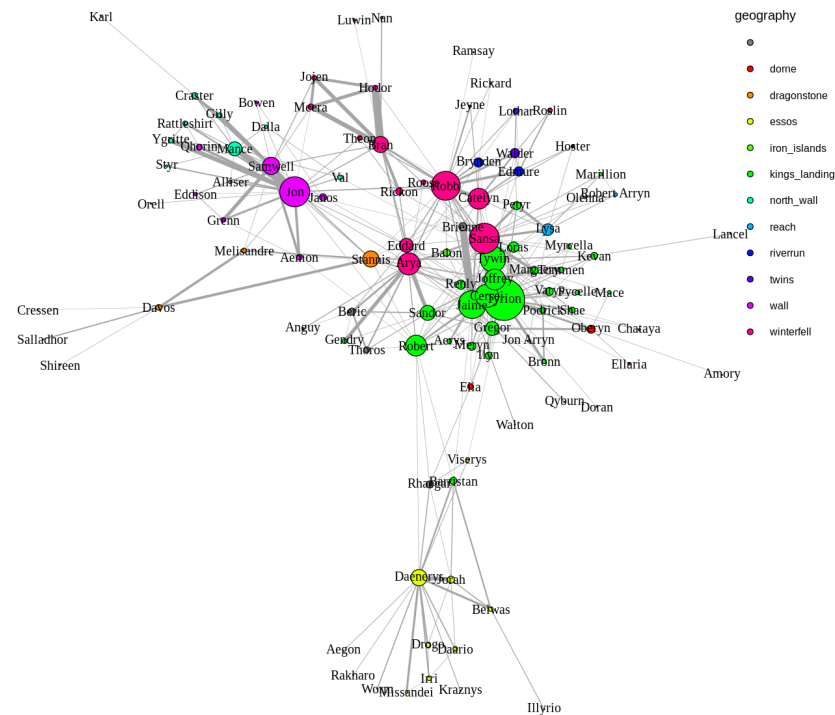


Figure 5. The network above shows the GOT characters where nodes have been colored by their author-generated geography. Since geography is one of the main factors of the storyline, it is clear that geography has grouped characters closely in the network. However, it is also apparent that not all characters are grouped only by geography.

The social network above in Figure 5 is overlaid with the author's own interpretation of the different characters' geographical locations in *A Storm of Swords*. The most obvious feature of our network is that the characters who are located in Essos are in their own section of the network, for the most part cordoned off from the rest except by ties that mainly run through the three Targaryens: Danaerys, Viserys, and Rhaegar; and Barristan Selmy who used to live in Westeros.

Similarly, the northern portions of Westeros are a distinct entity in our network map. The wildling characters who live north of the Wall are only connected to the rest of the characters through the members of the Night's Watch who live on the Wall. The physical wall imposes an interesting spatial

characteristic to our network map. In particular, the wildlings have strong connections to Jon Snow and Samwell Tarly, the only two members of the Night's Watch who both venture north of the Wall and also have strong ties to the noble houses south of the Wall, Jon especially being an illegitimate child of the Stark family which rules the North.

The third geographically distinct and separate portion of our network map is the host of characters that inhabit Dragonstone and support the claim of Stannis Baratheon to the Iron Throne. Stannis's family and lesser political aides are not even connected directly to him, living in a remote portion of our map and connected to the rest only through Davos, Hand of the King, (similar to a chief of staff) to Stannis. Melisandre, Stannis' other major advisor, is close to these characters as well but also involved in events near the Wall, allowing her to straddle the line between Dragonstone and the Wall. It is noteworthy that Stannis, as a major political figure with the most legitimate hereditary claim to the Iron Throne, is very near to the center of our map, despite his entourage being geographically distinct.

This distinctness of geography begins to break down as we analyze the role of the characters from Winterfell, especially the Stark family. Although many of the ancillary characters from Winterfell (housekeepers, maids, etc.) appear close to the other characters from the North, the Starks themselves are very close to the center of our map. Their great political stature in central events of the book, waging war against King Joffrey, means that they are not cordoned off into their own neat corner of our map.

Geographical centrality and political centrality overlap when we look at Kings' landing, the capital city of Westeros. There, characters from across the world gather to vie for power and influence. The geographical nuance that the author has tried to apply to different characters breaks down as their presence in the capital and interaction with central events pushes them together into a central grouping.

While some characters see geography fade away in importance as they interact with central events in the book, other characters are grouped together at the periphery despite their differing geography. Members of the Frey, Tully, and Bolton families find themselves in a distinct area of our map, even though they all live and mostly remain in very distinct areas of the physical world of Westeros. It is

clear that while geography is an important factor in analyzing our network map, our algorithm is thinking in a different plane than a typical *GOT* fan.

Community Detection:

Community detection offers another way to group characters more inherent to the data itself. The fast-greedy algorithm implements a hierarchical agglomeration algorithm to find communities in a large network to organize our characters into distinct membership groups as outlined in the map above (Clauset et. al, 2004). The algorithm created seven distinct communities, most of which would make immediate sense to any *GOT* fan (Figure 6).

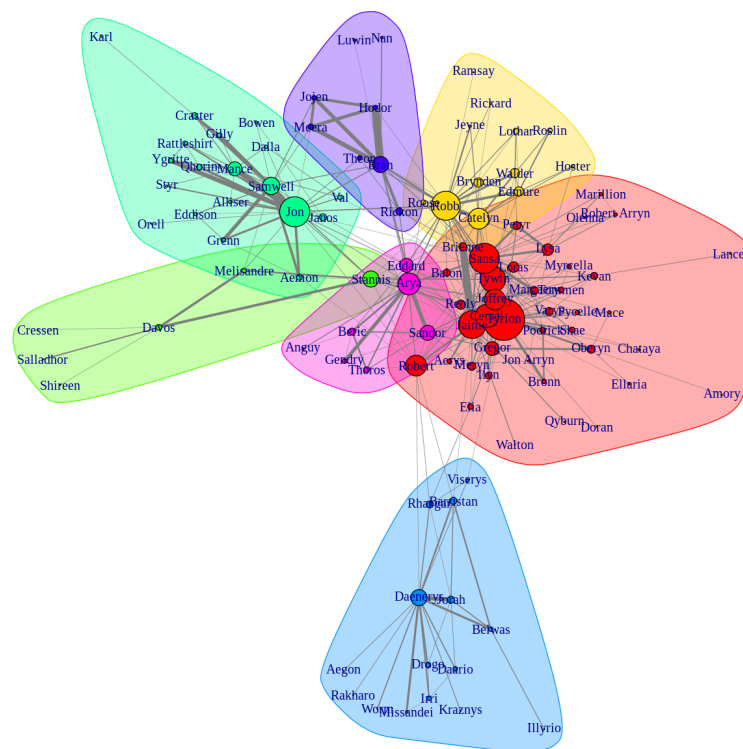


Figure 6. The network above is colored according to the groupings built by a fast-greedy community detection approach, which identified seven groupings within our dataset.

In particular, these are communities for the characters that inhabit Essos, those that inhabit the Wall and live north of the Wall, those that inhabit Dragonstone, and those that live in Kings' Landing. Another community is defined for the characters that are physically present in Winterfell during the events of the book: Bran, Rickon, Theon, and the ancillary characters of Winterfell, though not the remainder of the Starks who are physically elsewhere. This group was not nearly as distinct in our prior visualization by geography, where they straddled a middle ground between the general north, especially the Wall and north of the Wall and the Starks that are central to the book's plot.

Speaking of the Starks, the algorithm has separated Robb and Catelyn from Sansa, and those three from Arya and Eddard (Ned). Robb and Catelyn spend the whole book on the battlefield, waging war against King Joffrey's men, and do not interact with the other Starks. Robb and Catelyn are instead grouped with the Freys, Tullys, and Boltons, families whom they ally or negotiate with during the war. Thus we see that the algorithm has created a battlefield community, which our prior geographical interpretation made difficult to see.

The final group that our algorithm differentiates includes Ned, Arya, Gendry, Sandor, Beric, Anguy, and Thoros. All of these characters except Ned interact during Arya's attempted journey home to Winterfell. The final three in particular are members of the Brotherhood without Banners who encounter Arya, Gendry, and Sandor on the Kings' Road. These characters' brief interactions, driven by Arya's storyline, have overridden any other factors to have them grouped together by our algorithm. The inclusion of Ned is a bit puzzling, considering that he died in the first book in the series. Arya at this point is attempting to disguise herself as a common boy to avoid detection as the daughter of Ned Stark, whom the king hanged for treason. Her attempts to hide her identity, and her struggle to come to terms with her father's death, which she witnessed, might be the reason for Ned's inclusion in the group. Whatever the reason, it is interesting that this grouping is driven by Arya's personal story arc, which emphasizes her importance in the series despite her separation from most of the rest of the characters. It should be noted that Arya is the only character whose point of view is represented in each of the five novels released so

far in the *A Song of Ice and Fire* series, making it especially interesting that her personal arc is powerful enough on its own to create a separate community in our network.

Conclusion:

Both the community detection and geographical graphs hold important information with respect to key plot points in *GOT*. Overlaying both sets of categories on the same graph (Figure 7) allows for still further analysis of the social network of the *GOT* universe. As stated before, geography is still the prominent factor in the organization of our social network. Many of the communities defined by our algorithm are still very geographically focused, but some are so individual as to defy that characterization. An interesting insight is that, as in the real world, history in the *GOT* universe is defined both by great individuals and larger societal forces. For example, the community surrounding Arya is held together not by geography or common ties, but solely by the force of her individual plot, which culminates several books later in her defeat of the Night King. Meanwhile, the community at the Wall and north of the Wall represents the culmination of a long arc of history pitting the multitudes of wildlings against the Night's Watch, an important historical institution of Westeros society.

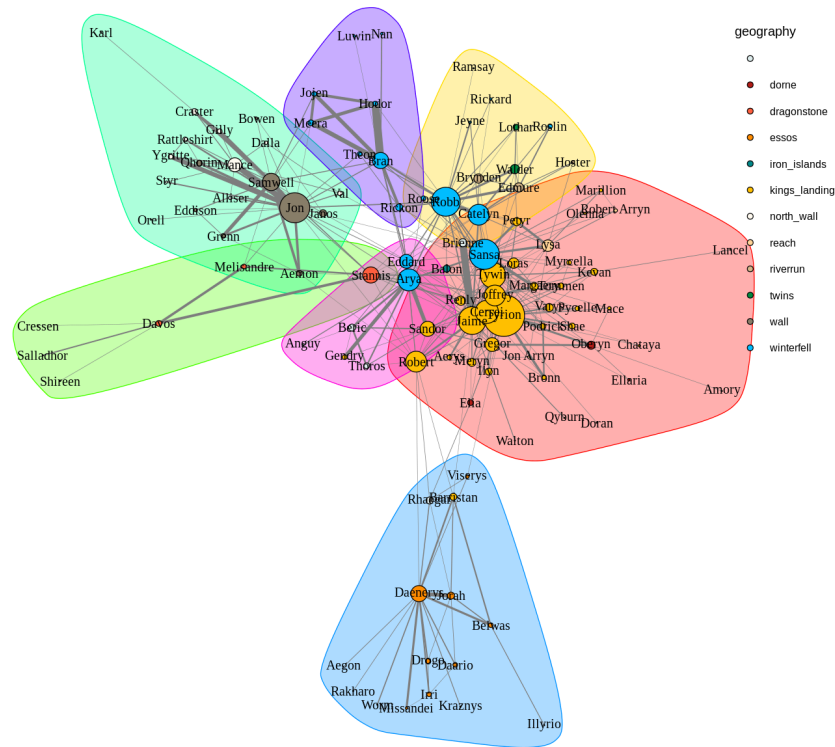


Figure 7. Results from the fast-greedy community detection are shown by the colored groupings while individual characters are colored based on their author-generated geography. This view allows for a comparison between the communities and the geographies that play a major role in the plot and building these groups; however, it is clear that additional information apart from geography is held in the connections between characters.

Across the narrow sea in Essos, the community surrounding Daenerys is almost entirely defined by its geographical proximity, but also by her family. Yet, if we wanted to extend our analysis to a more detailed level, we could develop even more interesting insights. While the author has provided more granularity on the geography of Westeros, Essos is treated as one location for this analysis. Yet, the continent of Essos is also filled with many competing factions who have complex relationships with each other, and whose relation to the greater universe is mainly through Daenerys.

Indeed, there are countless more ways to think about the *GOT* social network than are described in this analysis. Next steps might include the creation of a directed network graph which considers

relations as more complex than names appearing next to each other in the text. Such an analysis might try to use data that characterizes those relations as conversations, commands, mentions, or other useful categories in a way that might highlight power hierarchies or other social dynamics.

It would also be interesting to expand the analysis vertically to the other novels in the series. By conducting this analysis for each novel, we might construct a more powerful narrative of the evolution of the social network of the *GOT* universe. As mentioned earlier, the history of this universe is rich and complex, and by no means static. Characters rise, fall, and even die with captivating regularity. Consider as an example Tyrion Lannister, who is the central character in our social network of *A Storm of Swords*. A similar analysis conducted later in the series, when Tyrion has been exiled from Westeros and arrived in Essos to become an advisor to Daenerys, would yield a starkly different social network. The same might apply to Jon Snow when he finally leaves the Night's Watch to take part in Westeros politics.

The reason the author chose to apply this analysis to *GOT* was the universe's incredible depth and breadth of well-written, realistically motivated, and highly believable characters, who have become icons of popular culture today. Bringing the focus back to the real world, it would be fascinating to apply this analysis to the real world's "game of thrones". The ties that bind together the three branches of the US government could be one example. If some kind of analogous data set existed, it would be interesting to analyze the connections between higher and lower rankings members of Congress, as well as between federally elected Congressmen and the power brokers in their home constituencies, analogous to looking at the relations between the Starks and Lannisters, as well as between the noble houses and their bannermen. Social networks exist everywhere in our world, and the *GOT* universe provides an interesting case study in thinking about them.

References:

- Beveridge, A., & Shan, J. (2016, April). *Network of Thrones*. Maa.Org.
<https://www.maa.org/sites/default/files/pdf/Mathhorizons/NetworkofThrones%20%281%29.pdf>
- Bhagat, S., Burke, M., Diuk, C., & Filiz, I. O. (2017, January 14). *Three and a half degrees of separation*. Facebook Research. <https://research.fb.com/three-and-a-half-degrees-of-separation/>
- Clauset, A. (2004, August 9). *Finding community structure in very large networks*. ArXiv.Org.
<https://arxiv.org/abs/cond-mat/0408187>
- Delphine. (2017, June 3). *How Can I Be a Feminist And Love George R.R. Martin's Books?* WIRED.
<https://www.wired.com/2011/08/how-can-i-be-a-feminist-and-love-george-r-r-martins-books/>
- Game of Thrones (TV Series 2011–2019)*. (2011, April 17). IMDb.
<https://www.imdb.com/title/tt0944947/>
- Game of Thrones Viewer's Guide*. (2015). Home Box Office. <http://www.gameofthrones.com/game-of-thrones>
- Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical Analysis of Network Data with R (Use R!)* (2014th ed.). Springer.
- MacMillan, A. (2017, February 23). *Why It's Healthy to Cry Over TV Shows*. Time.
<https://time.com/4681115/crying-television-this-is-us/>
- Milgram, Stanley (May 1967). "The Small World Problem". *Psychology Today*. Ziff-Davis Publishing Company.
- Segall, M. (2018, October 23). *George RR Martin Confirms Popular Fan Theory About Game of Thrones's White Walkers*. Mental Floss. <https://www.mentalfloss.com/article/561583/george-rr-martin-confirms-game-of-thrones-fan-theory-white-walkers>
- Swatman, R. (2017, July 14). *Game of Thrones: 10 records deserving of the Iron Throne*. Guinness World Records. <https://www.guinnessworldrecords.com/news/2017/7/game-of-thrones-10-records-deserving-of-the-iron-throne-481693>
- Walsh, M. (2018, May). *melaniewalsh/sample-social-network-datasets*. GitHub.
<https://github.com/melaniewalsh/sample-social-network-datasets/tree/master/sample-datasets/game-of-thrones>

Appendix:

Example Data Nodes:

ID	Label	Gender	Family Name	Geography	Politics
Aemon	Grenn	5	Targaryen	Wall	Nightwatch
Aemon	Samwell	31		Wall	Nightwatch
Aerys	Jaime	18	Tarly	Wall	Nightwatch
Aerys	Robert	6	Targaryen	KingsLanding	Targaryen

Example Data Edges:

Source	Target	Weight
Aemon	Grenn	5
Aemon	Samwell	31
Aerys	Jaime	18
Aerys	Robert	6

Code Overview:

- Part 0: Data Retrieval
- Part 1: Data Exploration
- Part 2: Hub Nodes
- Part 3: Visualization
- Part 4: Community Detection
- Code available at https://github.com/lcombs/got_socialnetwork