Community detection and Throne Network analysis

Mritunjay kumar [1], Ashutosh kumar[2]

VIT UNIVERSITY CHENNAI CAMPUS

[1]*E-mail: mirtunjay.kumar2016@vitstudent.ac.in*

[2]*E-mail: ashutosh.kumar2016a@vitstudent.ac.in*

Abstract

The international hit HBO series *Game of Thrones*, adapted from George R. R. Martin's epic fantasy novel series *A Song of Ice and Fire*, features interweaving plotlines and scores of characters. With so many people to keep track of in this sprawling saga, it can be a challenge to fully understand the dynamics between them. To demystify this saga, we turn to *network science*, a new and evolving branch of applied graph theory that brings together traditions from many disciplines, including sociology, economics, physics, computer science, and mathematics. It has been applied broadly across the sciences, the social sciences, the humanities, and in industrial settings. In this paper we perform a network analysis of *Game of Thrones* to make sense of the intricate character relationships and their bearing on the future plot. We employed network analysis to systematically and quantitatively explore how the interactions between noble houses of the fantasy drama TV series Game of Thrones change as the show progresses. Our analysis discloses Degree Centrality, Betweenness Centrality and Closeness Centrality. We also try to find the communities from the data and the invisible threads that connected different houses and shows how tension across the houses, as measure via structural balance, changes over time. To boost the impact of our analysis, we further extended our analysis to explore how different network features correlate with viewer's engagement and appreciation of different episodes. This allowed us to derive a hierarchy of features that are associated with the audience response. All in all, our work show how network models may be able to capture social relations present in complex artificial worlds, thus providing a way to qualitatively model social interactions among fictional characters, hence allowing a minimal formal description of the unfolding of stories that can be instrumental in managing complex narratives.

Introduction

First, a quick introduction to *Game of Thrones*: Westeros and Essos, separated by the Narrow Sea, are homes of several noble houses. The narrative starts at a time of peace, with all the houses unified under the rule of King Robert Baratheon, who holds the Iron Throne. Early on, King Robert dies in a hunting accident, and the young, cruel Prince Joffrey ascends the throne, backed by his mother's house, Lannister. However, the prince's legitimacy, and even his identity, are seriously questioned across the kingdom. As a result, war breaks out, with multiple claimants Helen to the Iron Throne. Driven by cause or circumstance, characters from the many noble families launch into arduous and intertwined journeys. Among these houses are the

honorable Stark family (Eddard, Catelyn, Robb, Sansa, Arya, Bran, and Jon Snow), the pompous Lannisters(Tywin, Jaime, Cersei, Tyrion, and Joffrey), the slighted Baratheons (led by Robert's brother Stannis) and the exiled Daenerys, the last of the once powerful House Targaryen. We are doing net wok analysis to find most influencer person different size of communities ,ranking of different communities on the basis of social and economic background.

Literature Review

Network modelling is a powerful tool to represent interactions between different agents (usually called nodes), which can represent either physical or abstract entities [1, 2]. Once, a network has been constructed, it can be used to formally explore the different features observed, and hence to quantitatively assess the modes of interaction among the agents. Social sciences have a long history of using networks to explore human interactions [3], with the classical experiment of Milgram, paving the way to popular concepts like social hubs and six degrees of separation between individuals. While many theoretical approaches focused on studying the structures of networks with only one type of interaction (e.g., network of friends), in a few cases authors considered more rich models, where each interaction is associated with one of more interaction types. One of such theories is behind the concept of structural balance [4–7].

In this theory, the nodes represent individuals that can form positive or negative relationships with the others. Using a theoretical approach, it is also possible to show that the system will converge toward two ideal states: either all the individual display a positive attitude or the individuals forms two camps displaying a negative attitude towards each other [8]. Network analysis has been extensively applied in the field of social sciences [9], but its application to the study of dynamical changes of a social network over time are more limited, partly due to the complexity of controlling confounding factors such as uncontrollable external event. We therefore sought to explore the potential of network modelling to explore the isolated fictional world where the popular American fantasy drama TV series Game of Thrones takes place. The series has been adapted from George Martin's novel A Song of Ice and Fire [10] and has multiple interconnected plot lines and we decided to focus on the story arc unfolding on the fictional continent" Westeros" where noble houses of the "Seven Kingdoms" fight for the "Iron Throne" or independence. In employing network analytics to explore the structures and dynamics of the popular TV show, we assessed whether network theory could be used to quantify the changes associated with the evolution of the story and to gain insight into the reaction of the viewers to such changes.

- Comparison with other Paper:

Balance of thrones: a network study on Game of Thrones by Dianbo Liu and Luca Albergante They employed network analysis to systematically and quantitatively explore how the interactions between noble houses of the fantasy drama TV series Game of Thrones change as the show progresses All in all, their work show how network models may be able to capture social relations present in complex artificial worlds, thus providing a way to qualitatively model social interactions among fictional characters, hence allowing a minimal formal description of the unfolding of stories that can be instrumental in managing complex narratives.
Network of Thrones by Andrew Beveridge and Jie Shan: In this paper they perform a network analysis of Game of Thrones to make sense of the intricate character relationships and their bearing on the future plot

Proposed Work

Our first task is to turn the *Game of Thrones* world into a social network. Our network has sets of vertices *V* and edges *E*. The 107 vertices represent the characters, including ladies and lords, guards and mercenaries, councilmen and consorts, villagers and savages. The vertices are joined by 353 integer-weighted edges, in which higher weights correspond to stronger relationships between those characters. We generated the edges using *A Storm of Swords* the third book in the series. We opted for this volume because the main narrative has matured, with the characters scattered geographically and enmeshed in their own social circles. We parsed the ebook incrementing the edge weight between two characters when their names (or nicknames) appeared within 15 words of one another. Afterward, we performed some manual validation and cleaning. Note that an edge between two characters doesn't necessarily mean that they are friends—it simply means that they interact, speak of one another, or are mentioned together. The complex structure of our network reflects the interweaving plotlines of the story. Notably, we observe two characteristics found in many real-world networks. First, the network contains multiple denser subnetworks, held together by a sparser global web of edges. Second, it is organized around a subset of highly influential people, both locally and globally.

We now describe how to quantify these observations using the analytical tools of network science.

1) **Degree Centrality**: Degree centrality is defined as the number of links incident upon a node (i.e., the number of ties that a node has). If the network is directed (meaning that ties have direction), then two separate measures of degree centrality are defined, namely, indegree and outdegree.

2) **Betweenness Centrality**: In graph theory, betweenness centrality is a measure of centrality in a graph based on shortest paths. For every pair of vertices in a connected 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.

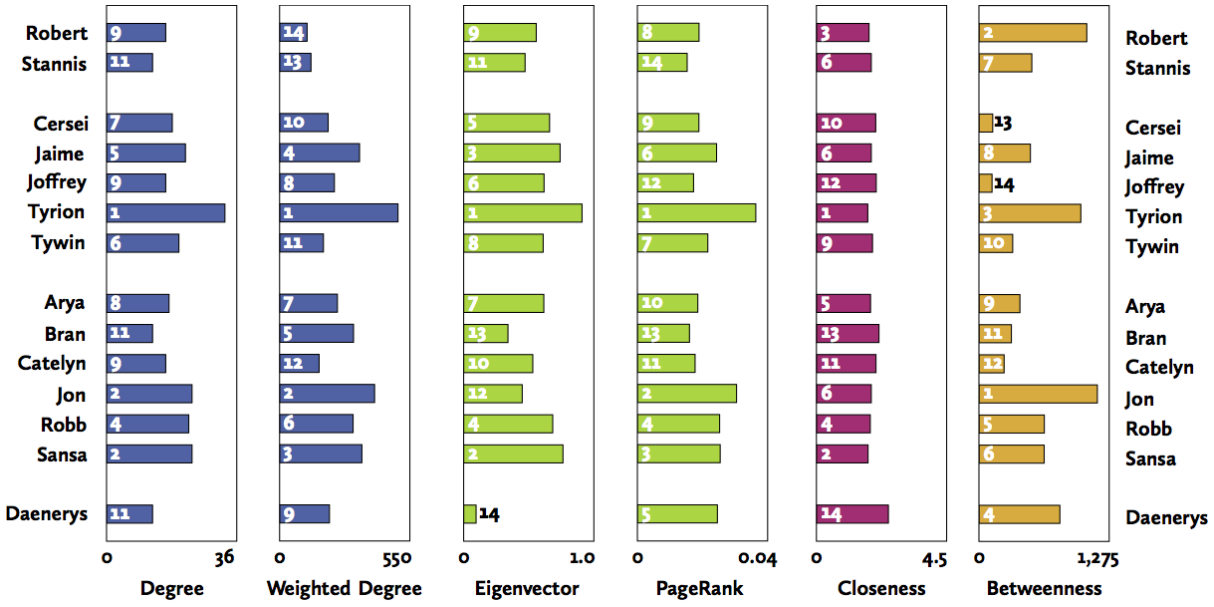
3) **Closeness Centrality**: In a connected graph, closeness centrality (or closeness) of a node is a measure of centrality in a network, calculated as the reciprocal of the sum of the length of the shortest paths between the node and all other nodes in the graph.

4) Community Detection : We wanted to find a natural division of the network into coherent communities, meaning that there are many links within communities and few links between communities. We optimized a measure called modularity, which captures this idea pretty well. This mathematical optimization splits the network into seven communities, as shown by the colors above. Note that we didn't tell the algorithm how many communities there were, or what their sizes were. These numbers were determined by the structure of the network itself. The resulting

communities make sense to anyone familiar with the saga: they are heavily influenced by the geographic location of the characters**.**

5.Character Importance

Next, we investigated character importance using six distinct centrality measures. Each of these measures addresses a different way to measure influence. The results and rankings are shown in the following table.



There is no single "right" centrality measure for a network. Each measure gives complementary information, and taking them in concert can be quite revealing. Here are an intuitive explanations for these measures.

- **Degree Centrality**: the number of other characters that you interact with.
- **Weighted Degree Centrality**: the number of interactions you participate in.
- **Eigenvector Centrality**: weighted degree centrality with a feedback boost for interacting with other important characters. You get full credit for the importance of your neighbors.
- **PageRank Centrality**: weighted degree centrality with a feedback boost for interacting with other important characters. The importance of your neighbors is split among its neighbors.
- **Closeness Centrality**: the average distance to all other characters (measured by number of links you must traverse). For this measure only, **smaller** numbers are better.
- **Betweenness Centrality**: how often you lie on shortest paths between two other characters, making you a broker of information.

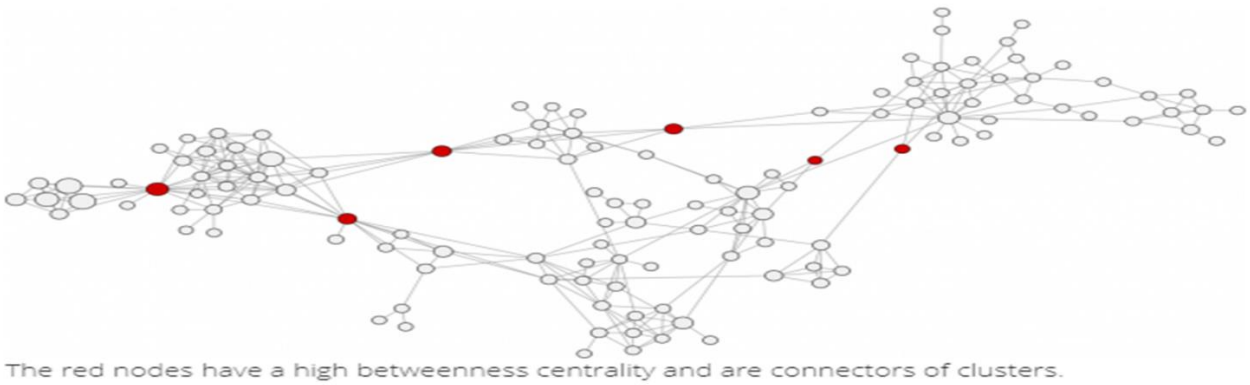
Results and Discussion

1.Diameter of the network (the longest shortest path)

"path"	"len"
[{"name":"Steffon-Varner"},{"name":"Eldon-Estermont"},{"name":"Bryce-Caron"},{"name":"Renly-Baratheon"},{"name":"Sansa-Stark"},{"name":"Arya-Stark"},{"name":"Gendry"},{"name":"Cutjack"},{"name":"Tarber"}]	8
[{"name":"Murch"},{"name":"Gariss"},{"name":"Aggar"},{"name":"Ramsay-Snow"},{"name":"Bran-Stark"},{"name":"Arya-Stark"},{"name":"Gendry"},{"name":"Cutjack"},{"name":"Kurz"}]	8
[{"name":"Steffon-Varner"},{"name":"Eldon-Estermont"},{"name":"Bryce-Caron"},{"name":"Renly-Baratheon"},{"name":"Sansa-Stark"},{"name":"Arya-Stark"},{"name":"Gendry"},{"name":"Cutjack"},{"name":"Kurz"}]	8
[{"name":"Murch"},{"name":"Gariss"},{"name":"Aggar"},{"name":"Ramsay-Snow"},{"name":"Robb-Stark"},{"name":"Renly-Baratheon"},{"name":"Bryce-Caron"},{"name":"Eldon-Estermont"},{"name":"Steffon-Varner"}]	8
[{"name":"Murch"},{"name":"Gariss"},{"name":"Aggar"},{"name":"Ramsay-Snow"},{"name":"Bran-Stark"},{"name":"Arya-Stark"},{"name":"Gendry"},{"name":"Cutjack"},{"name":"Tarber"}]	8

2.Betweenness centrality and Pivotal node

Betweenness centrality identifies nodes that are strategically positioned in the network, meaning that information will often travel through that person. Such an intermediary position gives that person power and influence. Betweenness centrality is a raw count of the number of short paths that go through a given node. For example, if a node is located on a bottleneck between two large communities, then it will have high betweenness. A node is said to be pivotal if it lies on all shortest paths between two other nodes in the network. We can find all pivotal nodes in the network.



The red nodes have a high betweenness centrality and are connectors of clusters.

Most Influential Person in Book

c.name	centrality
"Eddard-Stark"	4638.534951255039
"Robert-Baratheon"	3682.391035767813
"Tyrion-Lannister"	3272.6060155260348
"Jon-Snow"	2952.057281565677
"Catelyn-Stark"	2604.755646755593
"Daenerys-Targaryen"	1484.2780232288708
"Robb-Stark"	1255.6896562838228
"Drogo"	1115.0946392450378
"Bran-Stark"	960.0319135675136
"Sansa-Stark"	639.0769144474225

Pivotal Node in Graph

a.name	b.name	PivotalNode	pathLength	pathCount
"Varly"	"Cayn"	"Eddard-Stark"	2	1
"Tregar"	"Hallis-Mollen"	"Eddard-Stark"	3	5
"Tobho-Mott"	"Jyck"	"Eddard-Stark"	3	1
"Tobho-Mott"	"Jyck"	"Tyrion-Lannister"	3	1
"Porther"	"Rast"	"Eddard-Stark"	3	1
"Porther"	"Rast"	"Jon-Snow"	3	1
"Joss"	"Mance-Rayder"	"Eddard-Stark"	4	9
"Joss"	"Mance-Rayder"	"Benjen-Stark"	4	9
"Nestor-Royce"	"Jonos-Bracken"	"Catelyn-Stark"	3	1
"Nestor-Royce"	"Jonos-Bracken"	"Tytos-Blackwood"	3	1

Betweeness centrality vs Biggest Communicators

c.name	centrality	totalInteractions	charactersInteractedWith
"Eddard-Stark"	4638.534951255039	1284.0	["Daryn-Hornwood", "Beric-Dondarrion", "Balon-Greyjoy", "Wyl-(guard)", "Wylla"]
"Robert-Baratheon"	3682.391035767813	941.0	["Myrcella-Baratheon", "Lancel-Lannister", "Jorah-Mormont", "Meryn-Trant", "Jeyne-Poole"]
"Tyrion-Lannister"	3272.6060155260348	650.0	["Sansa-Stark", "Theon-Greyjoy", "Timett", "Tommen-Baratheon", "Varys"]
"Jon-Snow"	2952.057281565677	784.0	["Matthar", "Jory-Cassel", "Rodrik-Cassel", "Dywen", "Chett"]
"Catelyn-Stark"	2604.755646755593	520.0	["Robert-Arryn", "Rickon-Stark", "Wendel-Manderly", "Nestor-Royce", "Mychel-Redfort"]
"Daenerys-Targaryen"	1484.2780232288708	443.0	["Rakharo", "Rhaegar-Targaryen", "Aegon-I-Targaryen", "Varys", "Quaro"]
"Robb-Stark"	1255.6896562838228	516.0	["Osha", "Renly-Baratheon", "Hullen", "Wendel-Manderly", "Yoren"]
"Drogo"	1115.0946392450378	256.0	["Jommo", "Quaro", "Robert-Baratheon", "Rakharo", "Doreah"]
"Bran-Stark"	960.0319135675136	531.0	["Renly-Baratheon", "Petyr-Baelish", "Yoren", "Joffrey-Baratheon", "Hali"]
"Sansa-Stark"	639.0769144474225	545.0	["Rhaegar-Targaryen", "Janos-Slynt", "Jaime-Lannister", "Barristan-Selmy", "Tyrion-Lannister"]

3.PageRank

This is another version of weighted degree centrality with a feedback loop. This time, you only get your “fair share” of your neighbours’ importance i.e. your neighbour’s importance is split between their neighbours’, proportional to the number of interactions with that neighbour.

Intuitively, PageRank captures how effectively you are taking advantage of your network contacts. In our context, PageRank centrality nicely captures narrative tension. Indeed, major developments occur when two important characters interact.

c.name	c.book1PageRank	c.book1BetweennessCentrality	totalInteractions	charactersInteractedWith
"Eddard-Stark"	8.189869499999999	4638.534951255039	1284.0	["Daryn-Hornwood", "Beric-Dondarrion", "Balon-Greyjoy", "Wyl-(guard)", "Wylla"]
"Tyrion-Lannister"	5.948751000000001	3272.6060155260348	650.0	["Sansa-Stark", "Theon-Greyjoy", "Timett", "Tommen-Baratheon", "Varys"]
"Catelyn-Stark"	5.435648500000001	2604.755646755593	520.0	["Robert-Arryn", "Rickon-Stark", "Wendel-Manderly", "Nestor-Royce", "Mychel-Redfort"]
"Robert-Baratheon"	5.334583499999999	3682.391035767813	941.0	["Myrcella-Baratheon", "Lancel-Lannister", "Jorah-Mormont", "Meryn-Trant", "Jeyne-Poole"]
"Jon-Snow"	4.836755499999999	2952.057281565677	784.0	["Matthar", "Jory-Cassel", "Rodrik-Cassel", "Dywen", "Chett"]
"Robb-Stark"	3.8610744999999995	1255.6896562838228	516.0	["Osha", "Renly-Baratheon", "Hullen", "Wendel-Manderly", "Yoren"]
"Bran-Stark"	3.5735025	960.0319135675136	531.0	["Renly-Baratheon", "Petyr-Baelish", "Yoren", "Joffrey-Baratheon", "Hali"]
"Sansa-Stark"	3.5706210000000005	639.0769144474225	545.0	["Rhaegar-Targaryen", "Janos-Slynt", "Jaime-Lannister", "Barristan-Selmy", "Tyrion-Lannister"]
"Jaime-Lannister"	3.1322759999999996	550.724381183904	241.0	["Harys-Swyft", "Galbart-Glover", "Gregor-Clegane", "Addam-Marbrand", "Rickard-Karstark"]
"Cersei-	3.0479814999999992	454.8086339495463	424.0	["Benjen-Stark", "Arys-Oakheart", "Mordane", "Tommen-Baratheon",

Figure : page rank of community

Intra community page rank

c.name	c.communityPageRank
"Tyrion-Lannister"	4.767412500000001
"Podrick-Payne"	0.7141875
"Lemore"	0.33353200000000005
"Pia"	0.2775
"Osmund-Kettleblack"	0.2775
"Haldon"	0.24562500000000004
"Sweets"	0.24083950000000004
"Jon-Connington"	0.23406500000000002
"Tybero-Istarion"	0.21375000000000002
"Penny"	0.21375000000000002

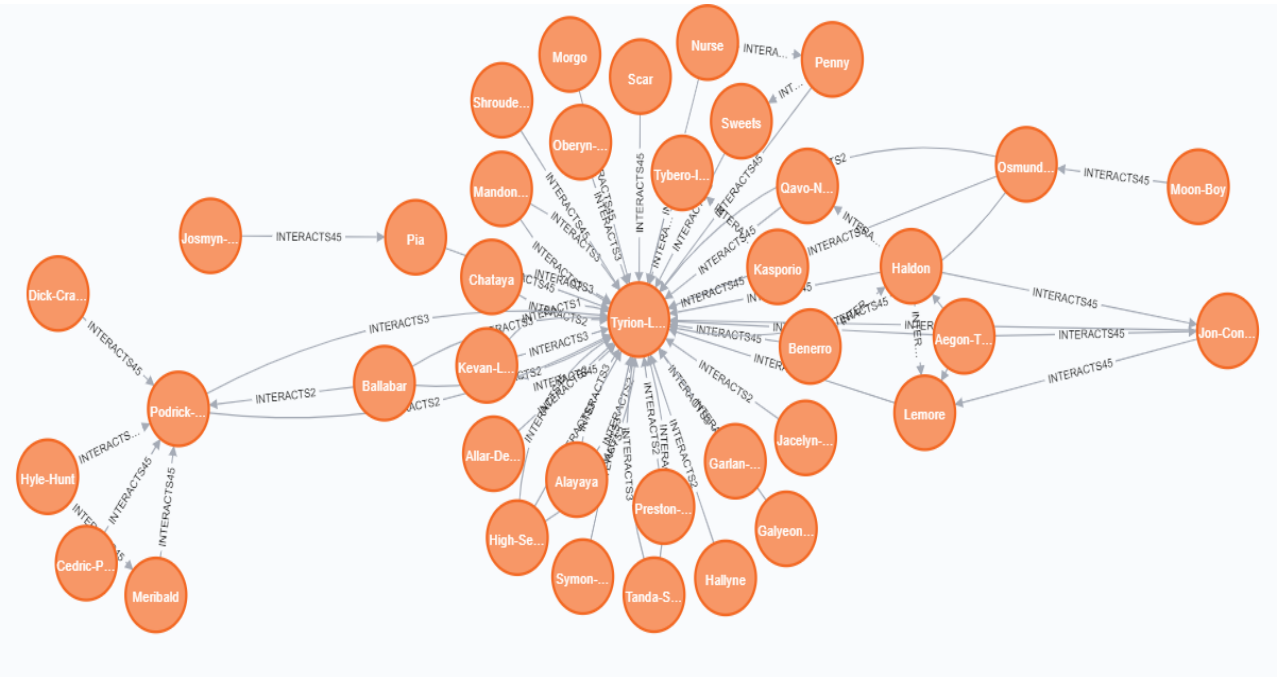
4.Community Detection

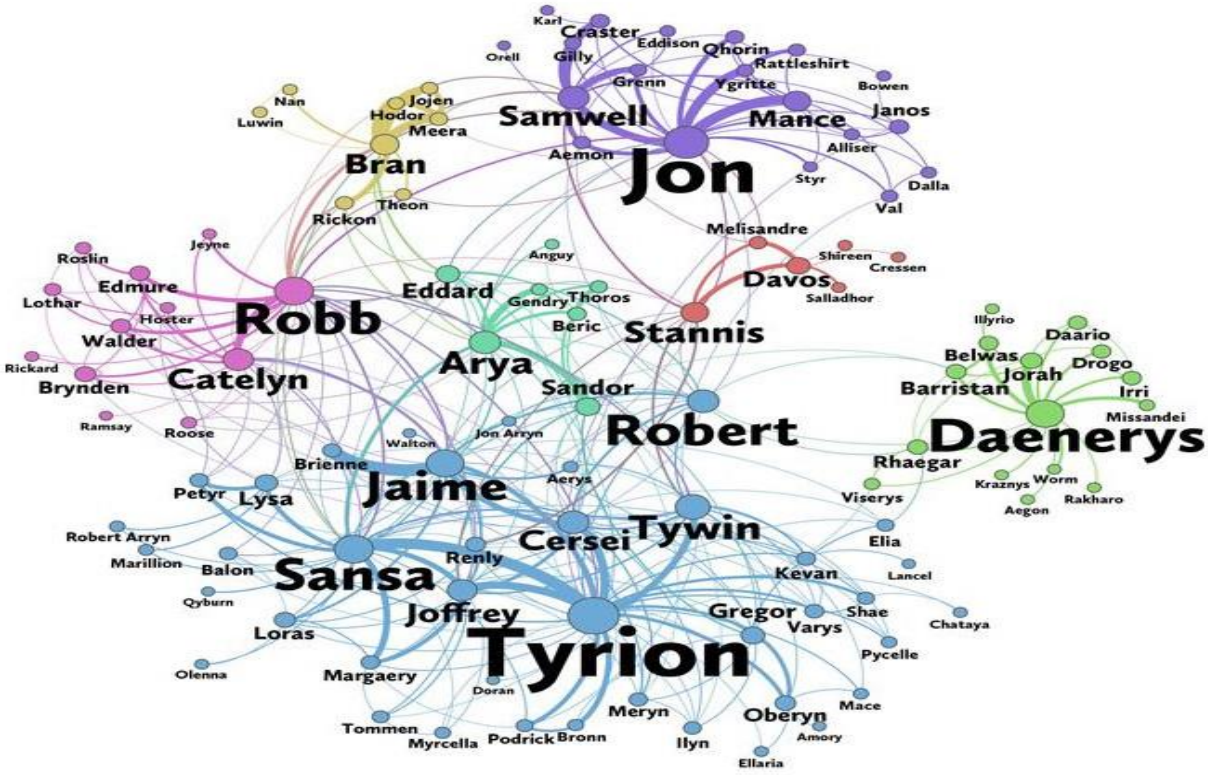
We can detect communities in our data by running an algorithm which traverses the graph structure to find highly connected subgraphs with fewer connections other subgraphs

c.community	count
136	111
137	40
331	38
140	22
128	18
738	17
133	16

Influential Person in each of the community:

cluster	count	collect(c.name)[..10]
136	111	["Stannis-Baratheon", "Sansa-Stark", "Samwell-Tarly", "Tommen-Baratheon", "Jaime-Lannister", "Selyse-Florent", "Petyr-Baelish", "Shireen-Baratheon", "Satin", "Melisandre"]
137	40	["Tyrion-Lannister", "Osmund-Kettleblack", "Tanda-Stokeworth", "Podrick-Payne", "Oberyn-Martell", "Jon-Connington", "Marq-Piper", "Pia", "Penny", "Symon-Silver-Tongue"]
331	38	["Theon-Greyjoy", "Wex-Pyke", "Walder-Frey-(son-of-Merrett)", "Ramsay-Snow", "Lorren", "Walder-Frey-(son-of-Jammos)", "Sour-Alyn", "Skinner", "Hagen", "Mikken"]
140	22	["Jon-Snow", "Pypar", "Randyll-Tarly", "Rast", "Hareth-(Moles-Town)", "Eddison-Tollett", "Jarl", "Hop-Robin", "Gorne", "Clydas"]
128	18	["Sandor-Clegane", "Robert-Arryn", "Joffrey-Baratheon", "Jory-Cassel", "Mordane", "Horas-Redwyne", "Polliver", "Arya-Stark", "Myranda-Royce", "Mya-Stone"]
738	17	["Victarion-Greyjoy", "Wulfe", "Rodrik-Harlaw", "Rodrik-Sparr", "Ralf-Stonehouse", "Moqorro", "Hotho-Harlaw", "Humfrey-Hewett", "Nute", "Quellon-Greyjoy"]
133	16	["Robb-Stark", "Rodrik-Cassel", "Rickon-Stark", "Lothar-Frey", "Leobald-Tallhart", "Jeyne-Westerling", "Balon-Greyjoy", "Jon-Umber-(Smalljon)", "Dacey-Mormont", "Raynald-Westerling"]
79	14	["Jeor-Mormont", "Grenn", "Aemon-Targaryen-(Maester-Aemon)", "Dywen", "Halder", "Hobb", "Bowen-Marsh", "Donal-Noye", "Alliser-Thorne", "Benjen-Stark"]
122	11	["Nan", "Rickard-Karstark", "Catelyn-Stark", "Maege-Mormont", "Edmure-Tully", "Jon-Umber-(Greatjon)", "Galbart-Glover", "Harren-Hoare", "Brynden-Tully", "Hallis-Mollen"]
82	11	["Daenerys-Targaryen", "Jhogo", "Drogo", "Hero", "Daario-Naharis", "Belwas", "Groleo", "Ben-Plumm", "Arstan", "Cleon"]
751	11	["Quentyn-Martell", "Tattered-Prince", "Meris", "Gerris-Drinkwater", "Kedry", "Cletus-Yronwood", "Archibald-Yronwood", "Caggo", "Denzo-Dhan", "Dick-Straw"]





Game of Thrones - Kings' Allies

So who is the most important character in "A Storm of Swords" In our network, three characters stand out consistently.

- **Tyrion Lannister**. Acting as the Hand of the King, Tyrion is thrust into the centre of the political machinations of the capitol city. He comes out on top in 5 of the 6 centrality measures. This suggests that he is the true protagonist of the book.
- **Jon Snow**. Jon Snow is the second most important character in the network. Indeed, he holds a unique position, with connections to highborn Lords, the Night's Watch militia, and the savage wildlings beyond the Wall.
- **Sansa Stark**. Sansa's high ranking might come as a surprise, since she is a de facto captive in King's Landing. However, other players are aware of her value as a Stark heir and they repeatedly use her as a pawn in their plays for power. If she can develop her cunning, then she can capitalize on her network position to dramatic effect.

Meanwhile, two characters stand out by over performing in certain centrality measures: **Daenerys Targaryen** and **Robert Baratheon**. They provide a clear counter-point to one another, and return our attention to the Iron Throne itself. Robert's memory unifies the crumbling network of the recent past, while Daenerys will surely upend the current network when she returns to Westeros in pursuit of the Throne.

Dataset

Character Interaction Networks for George R. R. Martin's "A Song of Ice and Fire" saga

These networks were created by connecting two characters whenever their names (or nicknames) appeared within 15 words of one another in one of the books in "A Song of Ice and Fire." The edge weight corresponds to the number of interactions.

Two characters appearing together in the same location

Two characters in conversation

One character talking about another character

One character listening to a third character talk about a second character

A third character talking about two other characters

References

[1] Newman, M. E. The structure and function of complex networks. SIAM review 45, 167–256 (2003).

[2] Newman, M., Barabasi, A.-L. & Watts, D. J. The structure and dynamics of networks (Princeton University Press, 2011).

[3] Barabasi, A.-L. Linked: How everything is connected to everything else and what it means. Plume Ed. (2002).

[4]Heider, F. Attitudes and cognitive organization. The J. psychology 21, 107–112 (1946).

[5]. Cartwright, D. & Harary, F. Structural balance: a generalization of heider's theory. Psychol. review 63, 277 (1956).

[6]. Harary, F. & Kabell, J. A. A simple algorithm to detect balance in signed graphs. Math. Soc. Sci. 1, 131–136 (1980)

[7]. Kunegis, J. et al. Spectral analysis of signed graphs for clustering, prediction and visualization. In Proceedings of the 2010 SIAM International Conference on Data Mining, 559–570 (SIAM, 2010).

[8]. Martin, G. R. A Song of Ice and Fire (Bantam, 1997).

[9]. Facchetti, G., Iacono, G. & Altafini, C. Computing global structural balance in large-scale signed social networks. Proc. Natl. Acad. Sci. 108, 20953–20958 (2011).

[10]. Marvel, S. A., Kleinberg, J., Kleinberg, R. D. & Strogatz, S. H. Continuous-time model of structural balance. Proc. Natl. Acad. Sci. 108, 1771–1776 (2011).

[11]. Antal, T., Krapivsky, P. L. & Redner, S. Dynamics of social balance on networks. Phys. Rev. E 72, 036121 (2005).

[12]. Beveridge, A. & Shan, J. Network of thrones. Math Horizons 23, 18–22 (2016).

[13]. Newman, M. Networks: an introduction (Oxford university press, 2010). 16. Newman, M. E. Assortative mixing in networks. Phys. review letters 89, 208701 (2002).