Social network analysis
of George R.R. Martin's A Song of Ice and Fire

Arianna Arruzzoli, Computer Science, ID 0001145336
Benedetta Bottari, Computer Science, ID 0001144057
Claudia Brunetti, Computer Science, ID 0001128544
Milena Mazza, Computer Science, ID 0001140174

1 Introduction

In this study, we focus on the epic fantasy series *A Song of Ice and Fire* by George R.R. Martin, a rich and intricate narrative where hundreds of characters interact within a sprawling, multifaceted storyline. *A Song of Ice and Fire* presents an ideal testbed for applying Social Network Analysis due to its large cast, complex political alliances, and interwoven plotlines that mirror real-world social dynamics. Our objective is to examine the underlying structure of the narrative's social network, identifying central characters, key relationships, and the broader organization of the network.

2 Problem and Motivation

The complex narrative of *A Song of Ice and Fire* provides a fertile ground for exploring the dynamics of social interactions within a fictional world. Our analysis addresses several interconnected problems, each aimed at uncovering the structural principles that underpin the story's network of characters. A key focus is understanding the roles of individual nodes, or characters, in the network. For instance, identifying the most influential or visible figures helps reveal power dynamics and leadership structures, including hidden leaders who exert influence behind the scenes. Similarly, recognizing strategic mediators, characters who act as bridges between different groups or factions, offers insight into how alliances are maintained and how information flows across the narrative. Additionally, pinpointing individuals capable of rapidly spreading information highlights the pivotal roles some characters play in driving the story forward, particularly during moments of tension or crisis. Finally, analyzing characters embedded in tightly-knit groups sheds light on the stability and cohesion of key alliances and factions.

At the community level, our study examines the broader social structure of the narrative by identifying major factions or alliances, such as noble houses or political coalitions. Mapping these communities helps us understand the sociopolitical landscape of the story, while analyzing their cohesion reveals the strength of internal relationships and their resilience to external pressures. Beyond this, examining the connections between different communities allows us to explore the dynamics of inter-group interactions, highlighting critical links that sustain the narrative's overall integrity. Within each community, identifying local leaders provides further insight into how influence is distributed and exercised within smaller, more localized networks.

On a global scale, we assess the overall structure of the network to understand how interconnected or fragmented it is. A cohesive network suggests tight integration of storylines and character interactions, whereas fragmentation points to isolated subplots or narrative segmentation. Furthermore, evaluating the network's efficiency in transmitting information, based on metrics such as path length and connectivity, offers a quantitative perspective on how the narrative facilitates interactions between distant groups or characters. Finally, identifying critical nodes and links—those whose removal would disrupt the network—highlights the key figures and relationships that underpin the entire story.

Together, these analyses provide a comprehensive framework for understanding *A Song of Ice and Fire* through the lens of social network analysis, offering insights into the interplay between individual characters, communities, and the global structure of the narrative.

3 Datasets

3.1 Dataset

For this project, we used a publicly available dataset[1], which is released under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.
The data are provided in both CSV and JSON formats. The dataset represents the interaction networks between characters from the literary series *A Song of Ice and Fire* by George R.R. Martin, based on the occurrence of characters' names (or nicknames) within 15 words of each other in the first five books of the series.
The dataset comprises a total of 796 nodes, representing the characters, and 2823 undirected edges, forming a single undirected monomodal social network. The data includes edge weights, which indicate the number of interactions between characters.
The dataset provides an interesting foundation for analyzing the social and narrative dynamics of the saga, offering insights into the interconnections and influences between the characters.

3.2 Tools

To analyze the data, Python was used along with libraries specialized in handling structured files. The CSV and JSON files provided in the dataset were loaded, preprocessed, and stored in optimized Python data structures via the pandas library. To calculate metrics on the data, we used advanced network science tools through the NetworkX and igraph libraries.
To calculate the centrality measures within the entire dataset and to visualize the main graphs, we used the Gephi program[2]. Gephi provided the main statistics regarding the Network Overview, the Community Detection, the Node Overview and the Edge Overview.
These tools allowed us to thoroughly explore the dynamics within the Seven Kingdoms, identifying key narrative plots and the strategic roles played by the characters within the social network of the saga.

[1] Dataset's repository on GitHub (chosen files: `asoiaf-all-edges.csv` and `asoiaf-all-nodes.csv`)
[2] Gephi - The Open Graph Viz Platform

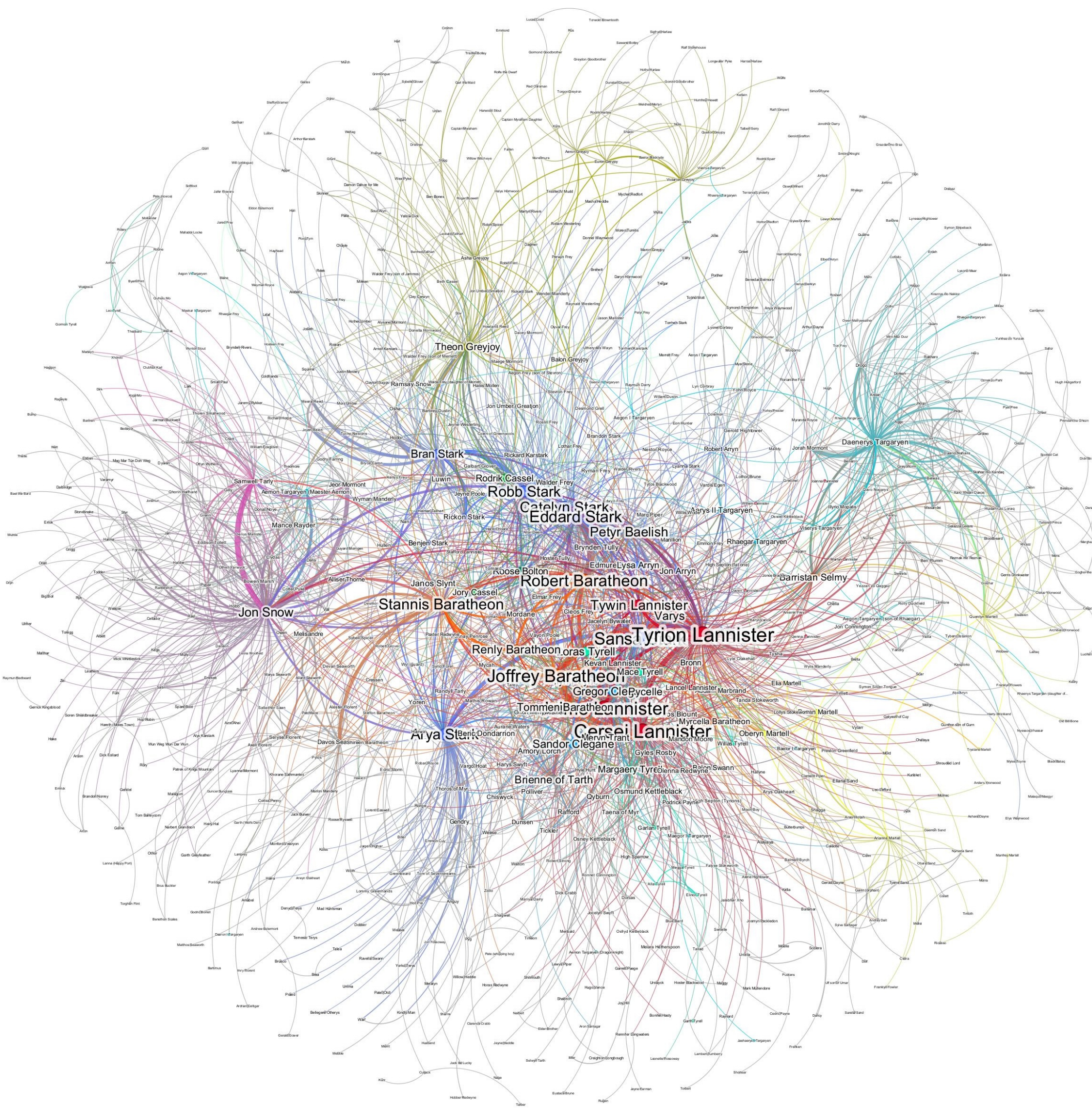


Figure 1: Complete graph created via Gephi.
The size of the node and its label is proportional to the authority value of the node

4 Validity and Reliability

4.1 Validity

The data collected inside the dataset are not an accurate representation of the reality of the relationships between the characters. The authors' decision to consider names appearing within a 15-word distance does not accurately reflect the nature of relationships between characters and arbitrarily excludes certain connections.
Furthermore, other key informations necessary for character identification are missing; for instance, houses (see section 4.2), which are fundamental within the context of the books, are not specified. Additionally, critical details such as gender, age, and whether characters are alive or deceased are absent. The dataset also includes references to characters from centuries prior to the events of the story, introducing irrelevant information that undermines the analysis.

4.2 Reliability

To perfectly reproduce the study, is necessary to first edit the dataset inside the `asoiaf-all-nodes.csv` file. Given that the context of *A Song of Ice and Fire* is set in a medieval-like era, the role of houses and families represents a crucial factor in understanding the underlying social dynamics. To this end, a column labeled "house" has been included in the dataset to identify the house affiliation of individual nodes based on their surnames. This labeling is essential not only for analyzing these dynamics but also for comparing them with the communities identified among the newtork.
The addition of this information does not alter the topology of the graph, but rather enhances the value of the analysis conducted.

5 Measures and Results

This section outlines the measures used to analyze the *A Song of Ice and Fire* social network and their significance in understanding the structural and functional properties of the network. Additionally, the results derived from these measures are discussed, emphasizing their application to the narrative and structural analysis of the network.

5.1 Measures

To explore the structural and functional properties of the social network, we employed several key measures from network science. These measures, along with their mathematical definitions and interpretations, are detailed below:

5.1.1 Degree Centrality

Degree centrality measures the number of direct connections a node i has in the network. It is calculated as:

$$C_D(i) = \deg(i) = \sum_{j \in N} A_{ij}$$

where A_{ij} is the adjacency matrix of the network, $A_{ij} = 1$ if there is an edge between node i and node j, and $A_{ij} = 0$ otherwise. Nodes with high degree centrality are directly connected to many other nodes and often serve as prominent communicators or hubs in the network.

5.1.2 PageRank

PageRank identifies the importance of a node by distributing the centrality score iteratively based on the importance of its neighbors. It is defined as:

$$PR(i) = (1-\alpha) + \alpha \sum_{j \in N(i)} \frac{PR(j)}{\deg(j)}$$

where $PR(i)$ is the PageRank of node i, α is a damping factor (typically set to 0.85), and $N(i)$ represents the set of neighbors of i. Nodes connected to highly influential nodes gain higher scores, making this measure ideal for identifying critical actors within the network.

5.1.3 Eigenvector Centrality

Eigenvector centrality evaluates a node's importance by considering the centrality of its neighbors. It is computed by solving the eigenvector equation:

$$\mathbf{C_E} = \lambda \mathbf{A} \mathbf{C_E}$$

where $\mathbf{C_E}$ is the vector of eigenvector centrality scores, $\mathbf{A}$ is the adjacency matrix, and λ is the largest eigenvalue. Nodes with high eigenvector centrality are embedded in influential clusters, highlighting their role in the broader network structure.

5.1.4 Betweenness Centrality

Betweenness centrality measures the frequency with which a node lies on the shortest paths between pairs of nodes. It is defined as:

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

where σ_{st} is the total number of shortest paths between nodes s and t, and $\sigma_{st}(i)$ is the number of those paths passing through node i. High betweenness centrality indicates a node's ability to act as a bridge or mediator between different network regions.

5.1.5 Closeness Centrality

Closeness centrality quantifies how quickly a node can access other nodes in the network. It is defined as the inverse of the sum of the shortest path distances from node i to all other nodes:

$$C_C(i) = \frac{1}{\sum_{j \in N} d(i, j)}$$

where $d(i, j)$ is the shortest path distance between nodes i and j. Nodes with high closeness centrality are well-positioned to disseminate information efficiently across the network.

5.1.6 Authority (HITS - Hyperlink-Induced Topic Search)

The authority score evaluates nodes that serve as authoritative sources of information. It is computed as part of the HITS algorithm, which iteratively assigns two scores, authority (a_i) and hub (h_i), based on the equations:

$$a_i = \sum_{j \in N(i)} h_j, \quad h_i = \sum_{j \in N(i)} a_j$$

Nodes with high authority scores are heavily referenced by influential hubs, emphasizing their significance in the network's information flow.

5.1.7 Density

The density of a network indicates the proportion of realized connections compared to all possible connections. It is calculated as:

$$D = \frac{2E}{N(N-1)}$$

where E is the number of edges, and N is the number of nodes. Density reflects the overall connectivity of the network, distinguishing tightly interconnected structures from sparse ones.

5.1.8 Clustering Coefficient

The clustering coefficient measures the degree to which a node's neighbors form a complete graph (i.e., are directly connected). For a node i, the local clustering coefficient is given by:

$$C(i) = \frac{2e_i}{k_i(k_i - 1)}$$

where e_i is the number of edges between the neighbors of i, and k_i is the degree of i. The global clustering coefficient is the average of local coefficients across all nodes and highlights the presence of cohesive subgroups or communities.

5.1.9 Small-Worldness

Small-worldness quantifies a network's ability to balance short average path lengths with high clustering, characteristics of small-world networks. It is evaluated using the small-world coefficient:

$$S = \frac{C/C_{rand}}{L/L_{rand}}$$

where C and L are the clustering coefficient and average path length of the network, and C_{rand} and L_{rand} are those of a random network with the same size and density. A small-world network typically has $S > 1$, reflecting high local cohesion and efficient global connectivity.

5.1.10 Scale-Free Property

A network is scale-free if its degree distribution follows a power-law distribution:

$$P(k) \sim k^{-\gamma}$$

where $P(k)$ is the probability of a node having k connections, and γ is a constant typically between 2 and 3. This property indicates the presence of a few highly connected hubs and many less connected nodes, a hallmark of hierarchical and resilient networks.

5.1.11 K-Core Decomposition

K-core decomposition identifies nested subgraphs in which each node has at least k connections. The coreness of a node reflects its maximal k-core membership. Nodes with high coreness are located at the center of the network, forming a dense, resilient core, while those with lower coreness reside in the periphery, highlighting their weaker or more isolated roles.

5.1.12 Louvain Modularity

Louvain modularity detects communities by optimizing the modularity score Q, which measures the density of connections within communities relative to connections between communities:

$$Q = \frac{1}{2E} \sum_{ij} \left[A_{ij} - \frac{\deg(i) \cdot \deg(j)}{2E} \right] \delta(c_i, c_j)$$

where c_i and c_j are the community assignments of nodes i and j, and $\delta(c_i, c_j)$ is 1 if i and j are in the same community and 0 otherwise. This method reveals densely connected subgroups that may correspond to factions, families, or alliances within the narrative.

5.2 Results

5.2.1 Global Network Analysis

The global analysis of the *A Song of Ice and Fire* social network aims to investigate its overall structure and identify the key mechanisms through which the narrative maintains connectivity. This analysis addresses essential questions, such as how interconnected the network is, which characters hold the story together, and how relationships between central and peripheral characters influence the development of the plot. Each measure was chosen specifically to provide insights into these aspects of the network, ensuring a robust framework for evaluating its structure and dynamics.

To understand the degree of cohesion in the network, we first assessed its density, which revealed a sparse structure at an initial value of $D = 0.0089$. This result reflects the sprawling and fragmented nature of the narrative, where numerous peripheral characters and isolated subplots contribute to the overall complexity. However, as nodes with lower connectivity (degree ≤ 70) are filtered out, the density increases dramatically, reaching $D = 0.8545$ (fig. 2). This transformation underscores the presence of a highly cohesive core group of characters who anchor the narrative. These results highlight the layered nature of the plot: a diffuse periphery supports the intricate world-building, while a concentrated central group drives the main storyline.

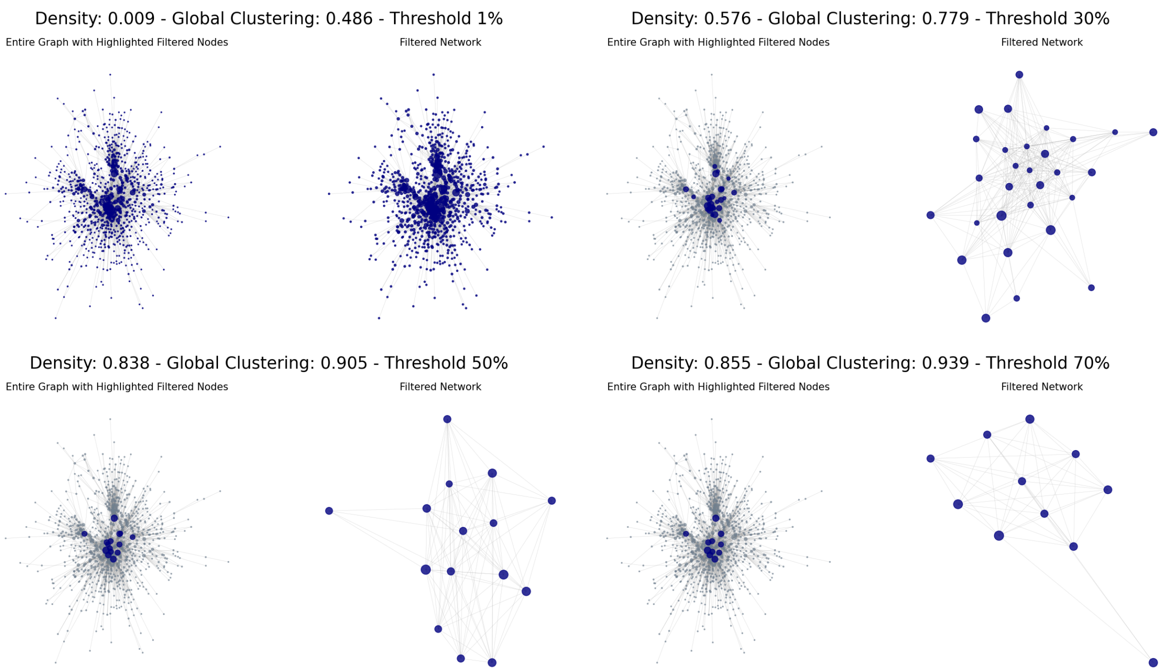


Figure 2: Network Density and Clustering Coefficient as Functions of Degree Centrality

To further examine the formation of cohesive subgroups, we analyzed the clustering coefficient, which measures the extent to which tightly-knit groups exist within the network. Starting at $C = 0.4858$, the coefficient increases to $C = 0.9389$ as peripheral nodes are excluded (fig. 2). This significant rise confirms that central characters not only play critical individual roles but also form dense, interconnected clusters.

This measure demonstrates how the narrative's complexity is anchored by strongly bonded groups that interact strategically to advance the story. The diameter of the network (9) and the average distance ($\bar{d} = 3.42$) reveal a small-world structure, characterized by short paths between nodes even in a sprawling network. This property reflects how the narrative ensures interconnectivity: despite its complexity, distant characters are linked by only a few intermediaries.

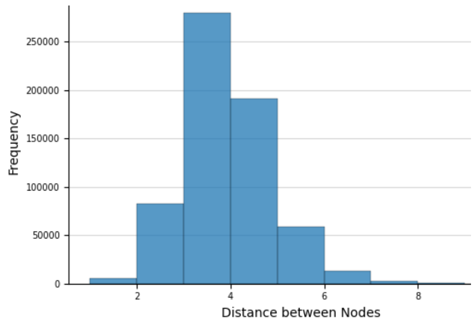


Figure 3: Distribution of Shortest Distances with Network Diameter 9

The small-worldness coefficient, $S = 5.837$, calculated from the clustering coefficient ($C_{\text{real}} = 0.4859$) and average path length ($L_{\text{real}} = 3.4162$) compared to a random network baseline ($C_{\text{random}} = 0.0810$, $L_{\text{random}} = 3.3245$), confirms this observation. Such a configuration facilitates rapid dissemination of information or influence across the network, ensuring coherence in the storytelling and allowing pivotal events to impact a wide range of characters.

To further investigate the stratified structure of the network, we performed a k-core decomposition. This method identifies nested subgraphs where each node is connected to at least k others. Characters with high coreness values ($k = 13$) are located at the network's center, forming its most densely connected and resilient core (fig. 4). In contrast, nodes with lower coreness occupy the periphery, reflecting characters with weaker connections or isolated roles. This decomposition emphasizes the centrality of key characters in maintaining the network's structural stability while illustrating how peripheral nodes contribute to broader world-building.

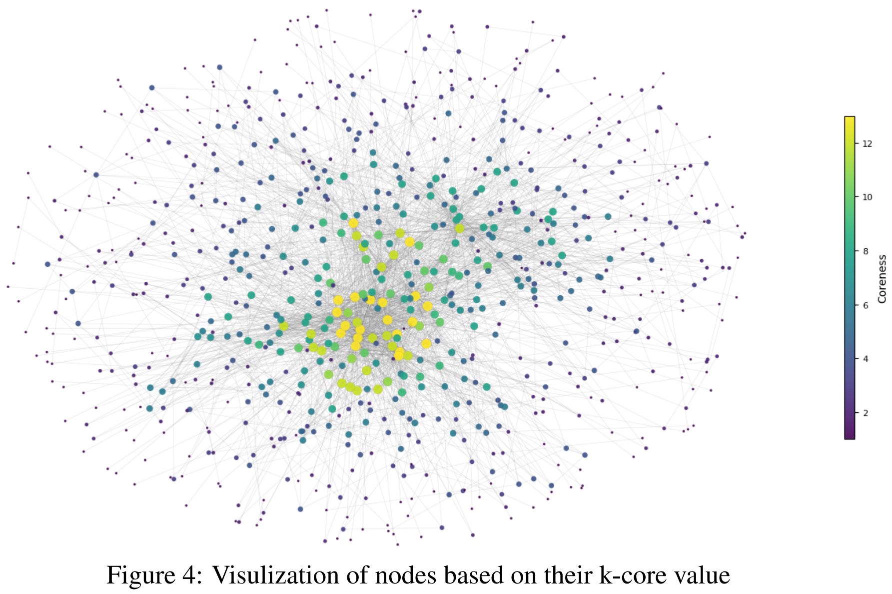


Figure 4: Visulization of nodes based on their k-core value

An analysis of articulation points (fig. 5) revealed 109 critical nodes whose removal would fragment the single connected component of the network into disconnected subgraphs. Prominent characters such as **Jon Snow**, **Daenerys Targaryen**, and **Tyrion Lannister** emerge as vital to the network's overall cohesion, capable of splitting it into five or more subnetworks if removed. This aligns with their narrative roles as central figures who bridge various factions and drive the story's major conflicts. However, the analysis also highlights less visible characters, such as **Aemon Targaryen**, whose structural importance might otherwise be overlooked. Their roles as articulation points reveal their latent influence in maintaining connections across disparate subplots.

Figure 5: Network with critical nodes and arcs colored red

Finally, the degree distribution (fig 6) of the network confirms a scale-free configuration, where many nodes exhibit low connectivity, while a few hubs dominate interactions. This pattern reflects the hierarchical organization of the narrative, where key hubs sustain the network's stability and serve as focal points for major plot developments. Meanwhile, peripheral nodes, though less connected, enrich the narrative by contributing diverse subplots and secondary characters that add depth to the story world.

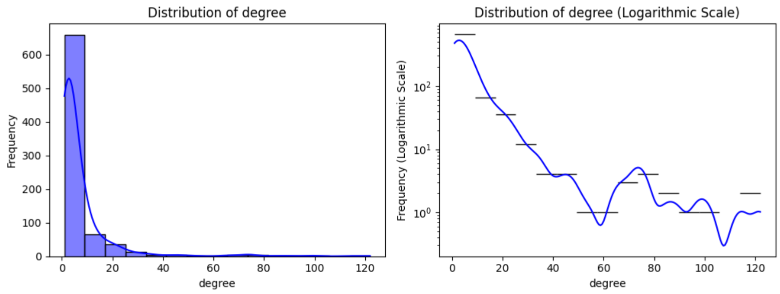


Figure 6: The histogram shows the frequency of node degrees, while the right-hand plot applies a logarithmic scale to better highlight the characteristics of the distribution

5.2.2 Nodes analysis

In *A Song of Ice and Fire*, the complex dynamics of power between characters shape the plot, focusing on shifting alliances, betrayals, and struggles for control. Identifying the key figures—leaders, mediators, and authoritative figures—is essential for understanding how power is distributed and how the interactions among characters influence the events of the series.

Identifying Prominent Leaders Prominent leaders in A Song of Ice and Fire are central figures in the narrative, shaping alliances, conflicts, and power dynamics.
To identify prominent leaders, we utilized **Degree Centrality** and **PageRank**. These metrics were selected because they capture complementary aspects of leadership. Degree Centrality measures the direct connectivity of a character, making it ideal for identifying individuals with extensive interactions who hold prominent positions in the network. PageRank, on the other hand, evaluates the quality of a character's connections by weighting the importance of their neighbors, revealing those whose influence extends beyond direct interactions to the broader structure of the network. Together, these metrics allow us to comprehensively analyze both the visibility and systemic influence of leadership within the narrative.

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

Personaggio	PageRank
Jon Snow	0.019000
Tyrion Lannister	0.018362
Jaime Lannister	0.015461
Stannis Baratheon	0.013650
Arya Stark	0.013448
Cersei Lannister	0.013362
Daenerys Targaryen	0.013048
Theon Greyjoy	0.012040
Eddard Stark	0.011485
Catelyn Stark	0.011187

Table 1: Top 5 characters identified from the analysis of prominent leader

The results (tab. 1) show that **Tyrion Lannister**, **Jon Snow**, and **Jaime Lannister** emerge as central figures. Tyrion has the highest Degree Centrality (122), reflecting his widespread connections. However, Jon Snow, despite a slightly lower Degree Centrality (114), has the highest PageRank (0.019),indicating the strategic significance of his ties across factions. Jaime Lannister ranks high in both Degree Centrality (101) and PageRank (0.015), underscoring his importance in military and political narratives.
The correlation between Degree Centrality and PageRank (fig. 20) highlights that prominent leaders are both highly connected and strategically positioned. Degree Centrality identifies their visible roles through direct relationships, while PageRank refines this by emphasizing their influence within the broader power dynamics. This dual perspective confirms the centrality of these leaders in driving the series' key events and power struggles.

Identifying Hidden Leaders Hidden leaders wield influence behind the scenes, leveraging strategic connections rather than occupying visible positions of power. To identify these individuals, we applied **Degree Centrality** and **Eigenvector Centrality**.
Degree Centrality was utilized to filter characters with limited direct connections, isolating those with lower visibility in the network. A threshold was applied, excluding nodes with Degree values above the 90th percentile—a choice justified by the scale-free nature of the network, where a small number of highly connected hubs dominate. This ensured focus on less connected characters who could otherwise be overlooked. Eigenvector Centrality then provided insight into the quality of these characters' connections. High Eigenvector values indicated that despite their relatively low number of direct ties, these individuals were strategically connected to central and influential figures in the network.

Personaggio	Degree	Eigenvector Centrality
Mandon Moore	9	0.206368
Olenna Redwyne	9	0.196215
Cleos Frey	7	0.188948
Elmar Frey	9	0.170334
Jacelyn Bywater	9	0.166741

Table 2: Top 5 characters identified from the analysis of hidden leader

The analysis (tab. 2)revealed that some characters, like Mandon Moore, Olenna Redwyne, and Cleos Frey, are important behind the scenes. They don't interact directly with many characters, but their connections to influential figures make them significant. Mandon Moore's role in the Kingsguard connects him to key political leaders, Olenna Redwyne wields influence through strategic family alliances, and Cleos Frey's family ties position him as a connector within the broader network.

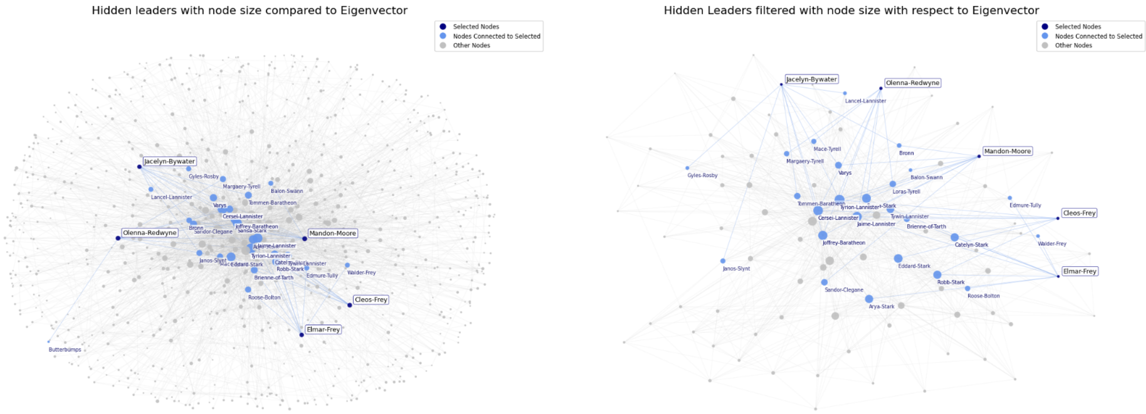


Figure 7: Complete and filtered graph on the first 100 nodes sorted on the eigenvector, with hidden leader

Identifying Key Mediators Key mediators are essential in bridging isolated groups within a network, ensuring the flow of information, resources, and influence. Using **Betweenness Centrality** (tab. 3), which identifies characters frequently positioned along shortest paths between other nodes, we identified **Jon Snow**, **Tyrion Lannister**, and **Daenerys Targaryen** as critical mediators. Jon Snow emerged as the primary mediator (0.192), connecting diverse factions such as the Wildlings and northern houses. Tyrion Lannister (0.162) facilitates communication between politically fragmented groups, using his diplomatic skills. Daenerys Targaryen (0.118) links disparate factions across continents, underscoring her strategic importance in aligning culturally distinct allies. These results highlight the significance of mediation in maintaining the network's structural and functional integrity.

Personaggio	Betweenness Centrality
Jon Snow	0.192120
Tyrion Lannister	0.162191
Daenerys Targaryen	0.118418
Theon Greyjoy	0.111283
Stannis Baratheon	0.110140

Table 3: Top 5 characters identified from the analysis of key mediators

Identifying Hidden Mediators Hidden mediators are characters who facilitate critical connections between groups while maintaining a low profile. To identify these figures, we combined **Betweenness Centrality**, which highlights nodes critical to the flow of information across the network, with a **Degree Centrality** filter set below the 9th percentile—a threshold chosen to exclude the network's highly connected hubs, typical of its scale-free structure, and focus on less prominent nodes. This approach ensured a focus on characters with relatively fewer direct connections who still play a strategic role in linking otherwise disconnected groups. The high Betweenness Centrality values identify their importance as bridges in the network, while the Degree filter excludes overly central figures to uncover those operating subtly. This approach (tab 4) identified **Myrcella Baratheon**, **Doran Martell**, and **Ramsay Snow** as key mediators, who brindge factions through diplomacy or manipulation.

Personaggio	Degree	Betweenness Centrality
Myrcella Baratheon	22	0.020193
Arianne Martell	20	0.017878
Ramsay Snow	23	0.016713
Doran Martell	21	0.014537
Alleras	8	0.014199

Table 4: Top 5 characters identified from the analysis of hidden mediators

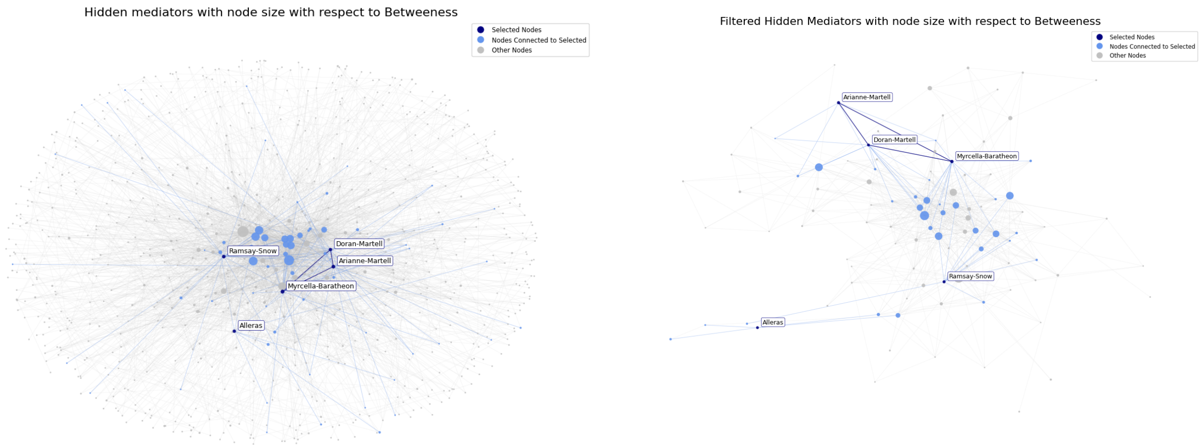


Figure 8: Complete and filtered graph on the first 100 nodes sorted on the betweeness, with hidden mediators

5.2.3 Community Analysis

The objective of this analysis is to identify how characters organize into sub-networks and to analyse the social dynamics between groups such as houses, alliances, and political factions inside the *A Song of Ice and Fire* saga.

Identifying Principal Houses The role of the houses within the context of the *A Song of Ice and Fire* saga is pivotal in shaping the political alignments and potential alliances throughout the narrative. To identify the principal houses and evaluate the extent to which they act as protagonists, it was necessary to exclude all nodes with the attribute “house” set to “Other”, indicating that they did not belong to any recognized house. The process resulted in the identification of sixteen houses A.1.
For each of the Houses identified, the analysis of **intra-house density** reveals that cohesion correlates with the number of nodes per house (tab. 6). Due to the dataset’s validity (section 4.1), there are single-node houses, which inevitably yield an internal density measure of zero. For the same reasons, two-node families have a cohesion value of one, representing a single edge connection between the two nodes. Larger houses show lower cohesion overall, but **House Stark** is notably cohesive, while **House Frey** appears highly strewn.

House	Leader	Degree	Mediator	Betweenness
House Arryn	Jon Arryn	3	Jon Arryn	0.33
House Baratheon	Stannis Baratheon	7	Stannis Baratheon	0.28
House Bolton	Ramsay Snow	1	Ramsay Snow	0.00
House Cassel	Rodrik Cassel	2	Rodrik Cassel	1.00
House Clegane	Gregor Clegane	1	Gregor Clegane	0.00
House Frey	Walder Frey	7	Walder Frey	0.35
House Greyjoy	Euron Greyjoy	6	Euron Greyjoy	0.10
House Jon Snow	Jon Snow	0	Jon Snow	0.00
House Lannister	Jaime Lannister	8	Jaime Lannister	0.41
House Martell	Arianne Martell	3	Oberyn Martell	0.19
House Selmy	Barristan Selmy	0	Barristan Selmy	0.00
House Stark	Eddard Stark	9	Eddard Stark	0.25
House Targaryen	Daenerys Targaryen	6	Daenerys Targaryen	0.35
House Tarly	Randyll Tarly	1	Randyll Tarly	0.00
House Tully	Brynden Tully	2	Brynden Tully	0.00
House Tyrell	Margaery Tyrell	6	Margaery Tyrell	0.27

Table 5: Nodes identified as authority, leader and mediator for each House

House	Intra-community density	Number of nodes	Graph
House Arryn	0.40	5	14a
House Baratheon	0.64	8	16a
House Bolton	1.00	2	14b
House Cassel	0.67	3	14c
House Clegane	1.00	2	14d
House Frey	0.08	24	16c
House Greyjoy	0.61	8	16b
House Jon-Snow	0.00	1	14g
House Lannister	0.33	12	15a
House Martell	0.25	8	14e
House Selmy	0.00	1	14f
House Stark	0.50	12	15b
House Targaryen	0.10	21	15c
House Tarly	1.00	2	14h
House Tully	1.00	3	14i
House Tyrell	0.25	11	15d

Table 6: Intra-density values for each House identified

Analysis of Inter-Houses Relationships The inter-house connections reveal the strength of relationships between different houses. **House Stark** and **House Lannister** have the strongest inter-house connection (31 connections), reflecting their central conflict, followed closely by **House Stark** and **House Baratheon** (30), which highlights their historic and political ties. Other prominent links include **House Lannister** and **House Baratheon** (27) and **House Stark** and **House Frey** (24), driven by major events such as alliances and betrayals. These metrics highlight Stark, Lannister, and Baratheon as central to the story's political and narrative framework 9.

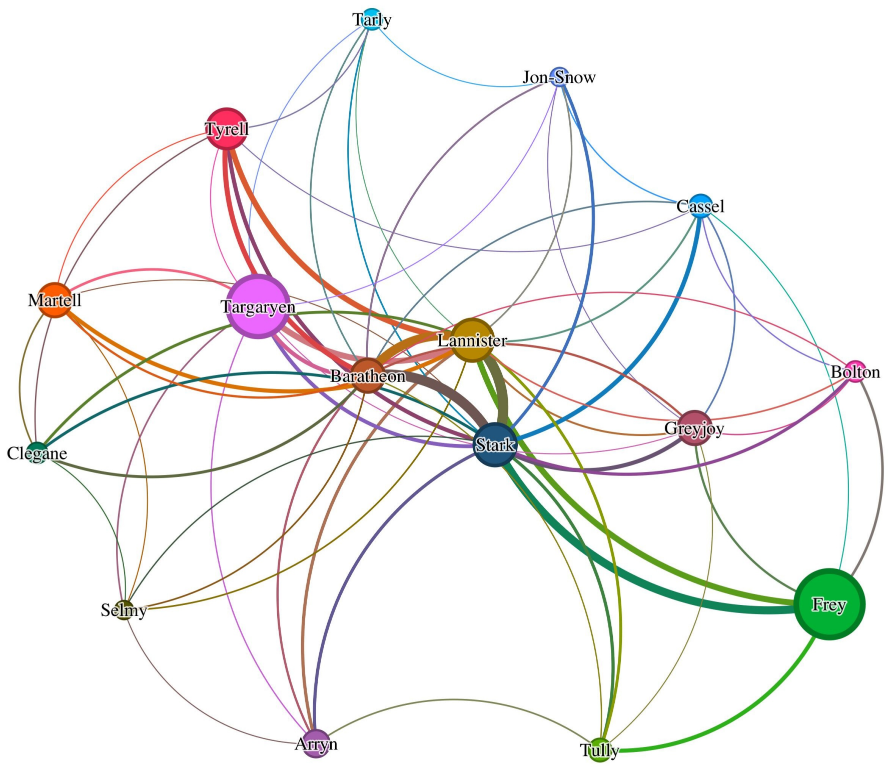


Figure 9: How the Houses interact with each other.
The size of the nodes is proportional to the number of family's members, while the edges have sizes proportional to the number of connections

House's Leaders The analysis aims to determine the roles of characters within their respective families, by examining centrality metrics to determine local leaders and key figures driving internal dynamics. To achieve this, we analyzed the **Degree Centrality**, to highlight the nodes with the highest number of connections within the family, identifying local leaders. **Eddard Stark** emerges as the central leader in House Stark, while **Jaime Lannister** and **Stannis Baratheon** hold significant influence in their respective families, aligning with their narrative prominence.

House's Mediator The analysis aims to identify characters who serve as bridges between different houses, emphasizing their role in facilitating inter-house relationships through the **Betweenness Centrality** metric, calculated with respect to the house's cluster. **Jaime Lannister** and **Walder Frey**'s positions are crucial in forming alliances or conflicts, while **Daenerys Targaryen**'s outreach to other factions in order to form alliances and gain power throughout the story, align with these metrics. Betweenness centrality measures a character's role as a bridge between houses. Zero centrality, as seen in houses like Bolton, Clegane, Snow and Tarly occurs in smaller subgraphs with few nodes. It is important to highlight how the leadership role (degree centrality) and mediation role (betweenness centrality) do not always overlap (tab. 5).

Identifying Principal Communities Following the analysis of the houses, it became crucial to identify the communities within the graph. The **Louvain Modularity** was used to identify clusters by maximizing modularity, a measure that compares the density of connections within communities to those between them, highlighting groups of characters with frequent interactions. It revealed nineteen clusters of closely connected characters, visible in section A.2.
Additionally, **inter-community connections** and **intra-community density** metric were used. Intra-community density highlights smaller communities with full cohesion, while larger groups exhibit lower cohesion (tab. 7). Inter-community values reveals how **Community 10** is connected to all others, likely due to Tyrion Lannister's pivotal role as a mediator between groups, as mentioned in section 5.2.2. This underscores the importance of specific nodes in bridging relationships and spreading information across clusters.

Community	Intra-community density	Number of nodes	Graph
Community 4	1.00	2	19b
Community 8	1.00	2	19d
Community 12	1.00	2	19e
Community 17	1.00	2	19j
Community 3	1.00	3	19a
Community 13	1.00	3	19f
Community 18	1.00	3	19k
Community 19	0.67	3	19l
Community 5	0.43	7	19c
Community 14	0.35	11	19g
Community 16	0.27	16	19i
Community 15	0.21	23	19h
Community 7	0.08	44	18a
Community 9	0.05	58	18b
Community 2	0.07	60	17c
Community 0	0.09	63	17a
Community 6	0.06	89	17d
Community 11	0.05	114	18d
Community 1	0.05	115	17b
Community 10	0.04	176	18c

Table 7: Intra-density values for each community identified

Communities' Leader To identify the leader nodes inside each community, it has been calculated the **Degree centrality** measure, within the community. Additionally, the **Authority value** from the full graph was used to identify a representative, as it captures the node's broader influence and reliability across the network.
The analysis reveals a discrepancy between communities' leaders and main authorities, indicating that nodes identified as authorities do not always serve as the primary sources of information dissemination within communities (tab. 9).

Communities' Mediator To identify which nodes play the role of mediators within each community, connecting different communities, we referred to the **Betweenness Centrality** measure, calculated relative to the community (tab. 8). Results show that mediators often align with community leaders, except in Communities 5 and 16. In smaller communities, such as those with two or three nodes, Betweenness Centrality is either zero or predictable due to the lack of intermediary paths, though no subgraphs are isolated from the broader network.

Community	Leader	Mediator	Authority	Authority value
Community 0	Arya Stark	Arya Stark	Arya Stark	0.172624
Community 1	Theon Greyjoy	Theon Greyjoy	Catelyn Stark	0.175885
Community 2	Stannis Baratheon	Stannis Baratheon	Stannis Baratheon	0.182076
Community 3	Moelle	Moelle	Unella	0.009642
Community 4	Alys Arryn	Alys Arryn	Alys Arryn	0.00279
Community 5	Lewyn Martell	Arthur Dayne	Gerold Hightower	0.027488
Community 6	Daenerys Targaryen	Daenerys Targaryen	Barristan Selmy	0.098095
Community 7	Victarion Greyjoy	Victarion Greyjoy	Balon Greyjoy	0.043302
Community 8	Hobber Redwyne	Hobber Redwyne	Horas Redwyne	0.007511
Community 9	Jaime Lannister	Jaime Lannister	Jaime Lannister	0.226345
Community 10	Tyrion Lannister	Tyrion Lannister	Tyrion Lannister	0.251562
Community 11	Jon Snow	Jon Snow	Jon Snow	0.144182
Community 12	Florian the Fool	Florian the Fool	Florian the Fool	0.007212
Community 13	Alla Tyrell	Alla Tyrell	Alla Tyrell	0.012282
Community 14	Pate (novice)	Pate (novice)	Alleras	0.002925
Community 15	Arianne Martell	Arianne Martell	Myrcella Baratheon	0.07536
Community 16	Aegon Targaryen	Jon Connington	Jon Connington	0.033735
Community 17	Raynard	Raynard	Raynard	0.008261
Community 18	Gared	Gared	Waymar Royce	0.002029
Community 19	Willem Lannister	Willem Lannister	Willem Lannister	0.008856

Table 8: Leader and Mediator for each community identified

Community	Leader	Degree	Mediator	Betweenness
Community 0	Arya Stark	50	Arya Stark	0.71
Community 1	Theon Greyjoy	47	Theon Greyjoy	0.45
Community 2	Stannis Baratheon	31	Stannis Baratheon	0.49
Community 3	Moelle	2	Moelle	0.00
Community 4	Alys Arryn	1	Alys Arryn	0.00
Community 5	Lewyn Martell	4	Arthur Dayne	0.53
Community 6	Daenerys Targaryen	58	Daenerys Targaryen	0.68
Community 7	Victarion Greyjoy	22	Victarion Greyjoy	0.47
Community 8	Hobber Redwyne	1	Hobber Redwyne	0.00
Community 9	Jaime Lannister	34	Jaime Lannister	0.77
Community 10	Tyrion Lannister	75	Tyrion Lannister	0.38
Community 11	Jon Snow	82	Jon Snow	0.70
Community 12	Florian the Fool	1	Florian the Fool	0.00
Community 13	Alla Tyrell	2	Alla Tyrell	0.00

Community 14	Pate (novice)	7	Pate (novice)	0.65
Community 15	Arianne Martell	16	Arianne Martell	0.57
Community 16	Aegon Targaryen	9	Jon Connington	0.44
Community 17	Raynard	1	Raynard	0.00
Community 18	Gared	2	Gared	0.00
Community 19	Willem Lannister	2	Willem Lannister	1.00

Table 9: Nodes identified as authority, leader and mediator for each community

Analysis of Inter-Community Relationships The analysis aims to explore inter-community relationships, focusing on the strength and nature of connections between communities, identifying key nodes that serve as bridges or connectors between groups. To do so two primary metrics are employed: **Edge Weight Between Communities**, which measures the strength of inter-group connections, and **Overlap of Shared Nodes**, which identifies individuals interacting with multiple communities. **Tyrion Lannister** plays a central role, connecting several factions, notably with Community 1 (edge weight 106) and Community 0 (edge weight 96). Robert Baratheon, Tyrion, Cersei, Sansa Stark, and Eddard Stark serve as critical bridges, interacting with more than ten communities, highlighting the narrative importance of House Lannister and House Stark. All nodes identified are part of Community 10.

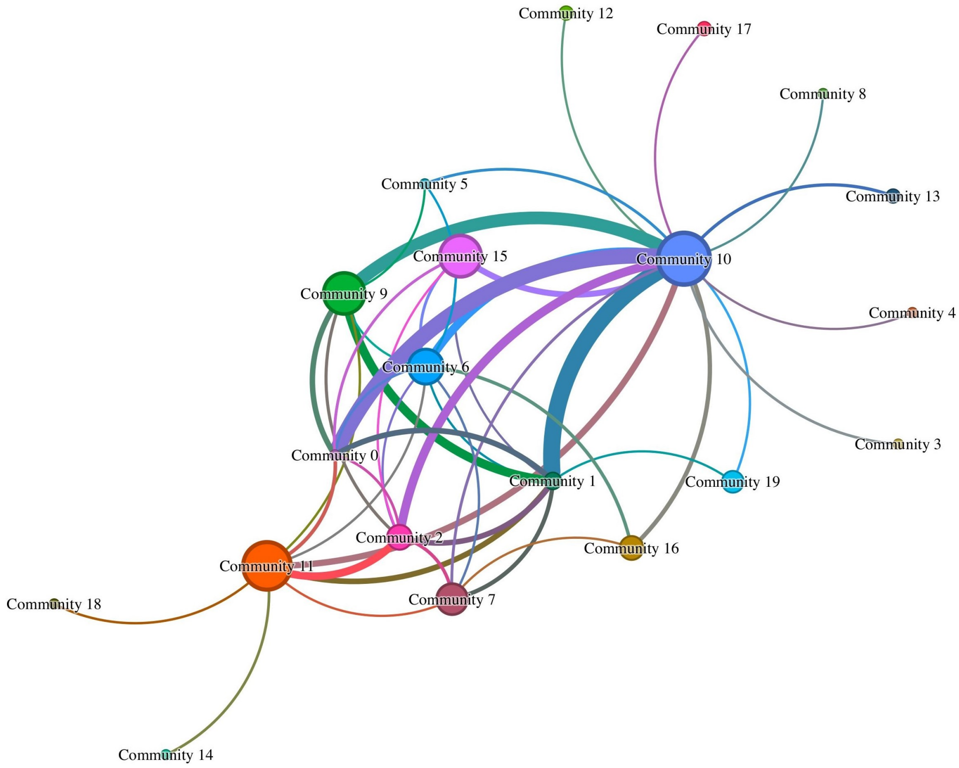


Figure 10: How the communities interact with each other.
The size of the nodes is proportional to the number of family's members, while the edges have sizes proportional to the number of connections

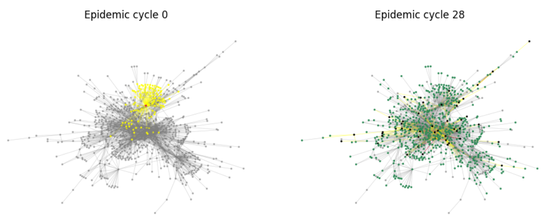
5.2.4 Epidemic Simulation

Inspired by George R.R. Martin's "Great Spring Epidemic", an analysis was conducted to model epidemic spread using realistic parameters for contagion, recovery, and mortality. The objective was to explore how the network structure and character roles influence disease dynamics. Two scenarios were simulated:

Epidemic	Infection Rate	Mortality Rate	Recovery Rate	Cycle Duration
COVID-19	0.2%	0.02%	0.3%	2 viral cycles
Black Plague	0.3%	0.5%	0.05%	7 viral cycles

Table 10: Epidemic Parameters

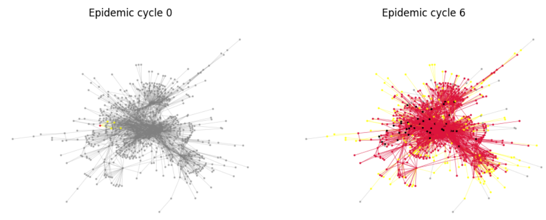
The epidemic begins with a "patient zero" infecting neighboring nodes based on arc weights, representing interaction frequency. Recovery and death probabilities are adjusted dynamically depending on the accessibility to medical resources and the number of deceased neighbors. Simulations end when 75% of nodes are either infected or dead, no further spread occurs, or after 100 cycles. Each character served as patient zero across 250 iterations, and results were averaged to reduce stochastic effects.



(a) COVID-19 Simulation (Patient Zero: Jon Snow)

(b) Plague Simulation (Patient Zero: Jon Snow)

(a) COVID-19 Simulation (Patient Zero: Hother-Umber)



(b) Plague Simulation (Patient Zero: Hother-Umber)

Results of the Analysis The results of the simulations are summarized below:

Character	Total Infected
Robert Baratheon	411.586
Asha Greyjoy	409.981
Varys	407.931
Eddard Stark	407.736
Bran Stark	407.000

Table 11: Top 5 by Infections

Character	Total Dead
Tywin Lannister	55.594
Varys	55.199
Gendry	55.077
Petyr Baelish	55.027
Jeor Mormont	54.992

Table 12: Top 5 by Deaths

Character	Total Cured
Jon Snow	605.801
Mance Rayder	604.540
Asha Greyjoy	604.280
Joffrey Baratheon	604.188
Sansa Stark	604.180

Table 13: Top 5 by Recoveries

The analysis revealed that the choice of patient zero significantly influences both the pace and extent of the epidemic, as illustrated in Figure 26. We further explore this through boxplots of centrality metrics divided by quartiles (e.g., Degree, Betweenness), in Appendix A.6 and figure

13 showing that higher centrality values are associated with more severe epidemic outcomes, including more infected, deaths, and longer epidemic cycles. This shows how highly central individuals amplify the spread of the epidemic due to their strategic positions in the network.

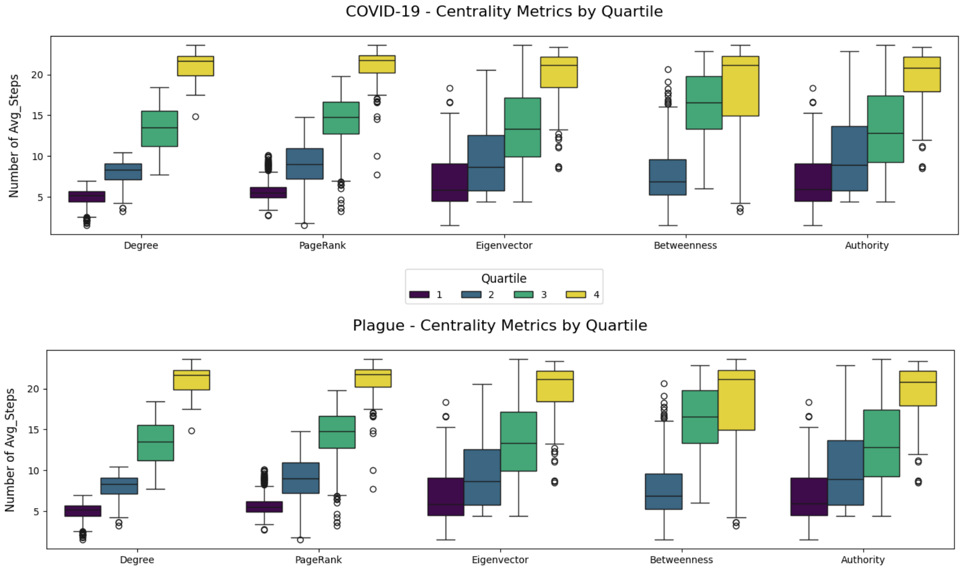


Figure 13: Distribution of epidemic cycles depending on the choice of patient zero, according to the quartiles of the different metrics

When patient zero has weak connections or limited interactions, the epidemic fails to spread extensively, often extinguishing after a few cycles with minimal casualties. This is especially evident in the plague simulation, where characters often die before spreading the disease (Figures 23 and 24).

Overall, as shown in Figures A.7, the two epidemics behaved as expected: COVID-19 displayed greater overall spread, infecting over 50% of the network in some iterations due to its low mortality and high recovery rates. The Black Plague, while achieving significant spread (under 50%), resulted in substantially higher mortality rates and limited recovery (below 5%) across all simulations.

6 Conclusion

The quantitative analysis of the social network in *A Song of Ice and Fire* highlighted a complex and layered structure that reflects the narrative depth and the intricate web of social, political, and familial dynamics within the saga.

Centrality measures such as Degree Centrality, PageRank, and Eigenvector Centrality enabled the identification of both visible and hidden central figures. The results reveal that characters like Tyrion Lannister, Jon Snow, and Cersei Lannister exert critical influence not only through their direct connections but also due to their strategic positions within the global network.

The analysis of communities emphasized the importance of alliances and familial bonds, uncovering highly cohesive subgroups representing houses and political factions. Structural measures such as density and clustering further revealed a balance between a diffuse periphery, which enriches the narrative world-building, and a cohesive central core, which drives the development of the main plot. The small-world property of the network ensures that key events can spread rapidly, connecting seemingly distant characters and maintaining narrative coherence.

The inclusion of epidemic simulations has added another layer of understanding, demonstrating how the network's structure can influence the spread of diseases, information, or influences. While the simulation scenarios are hypothetical, the results enhanced the understanding of certain characters' strategic roles and others' vulnerabilities.

The combination of simulations and metrics offered new insights into the network's vulnerabilities and the power dynamics shaping the story, revealing characters ignored in the main event of the story that results in playing powerful roles within the social network.

7 Critique

This study effectively addresses the research questions concerning the structural and narrative dynamics of the *A Song of Ice and Fire* social network, offering valuable insights into the roles of central characters, mediators, and community structures. These findings align with the narrative's organization, particularly in terms of identifying key figures and their roles in connecting different factions. However, the analysis does not fully resolve all aspects of the research problem due to several methodological and data-related limitations.

The primary issue lies in the representation of relationships. Relationships were inferred from co-occurrences of character names within a 15-word window, failing to distinguish between different types of interactions, such as alliances, enmities, or neutral ties. This method limits the depth of understanding regarding the nature of relationships. Furthermore, the use of undirected edges prevents the capture of relationship directionality, which is crucial for analysing power dynamics, such as hierarchical or reciprocal ties. Additionally, inferring house affiliations based solely on character names introduces inaccuracies, limiting the understanding of familial and political alliances constituted via marriages.

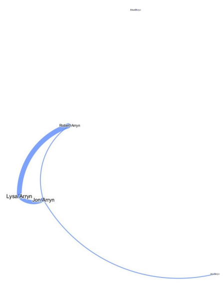
Another limitation arises from the lack of interaction metadata. Without attributes like interaction type or strength, it was not possible to apply equivalence measures to investigate structural redundancy or functional roles among characters.

Moreover, the epidemic models used assumed static interactions, overlooking temporal variations in the narrative that could impact influence dynamics.

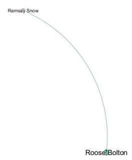
In conclusion, while the study addresses many aspects of the research question, it does not fully resolve it due to these methodological and data-related constraints. These limitations prevent a more nuanced and comprehensive analysis of the network's dynamics. Future research could improve upon this by incorporating more detailed data on interactions and relationships, such as directional edges, character affiliations, and interaction types, to enable a deeper exploration of the network's structure and narrative significance.

A Appendix

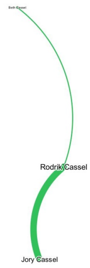
A.1 House Analysis



(a) House Arryn

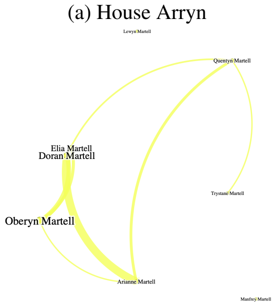


(b) House Bolton



(c) House Cassel

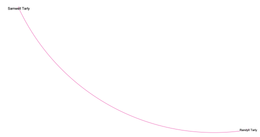
(d) House Clegane



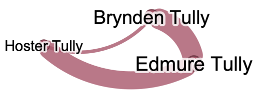
(e) House Martell

(f) House Selmy

(g) House Snow

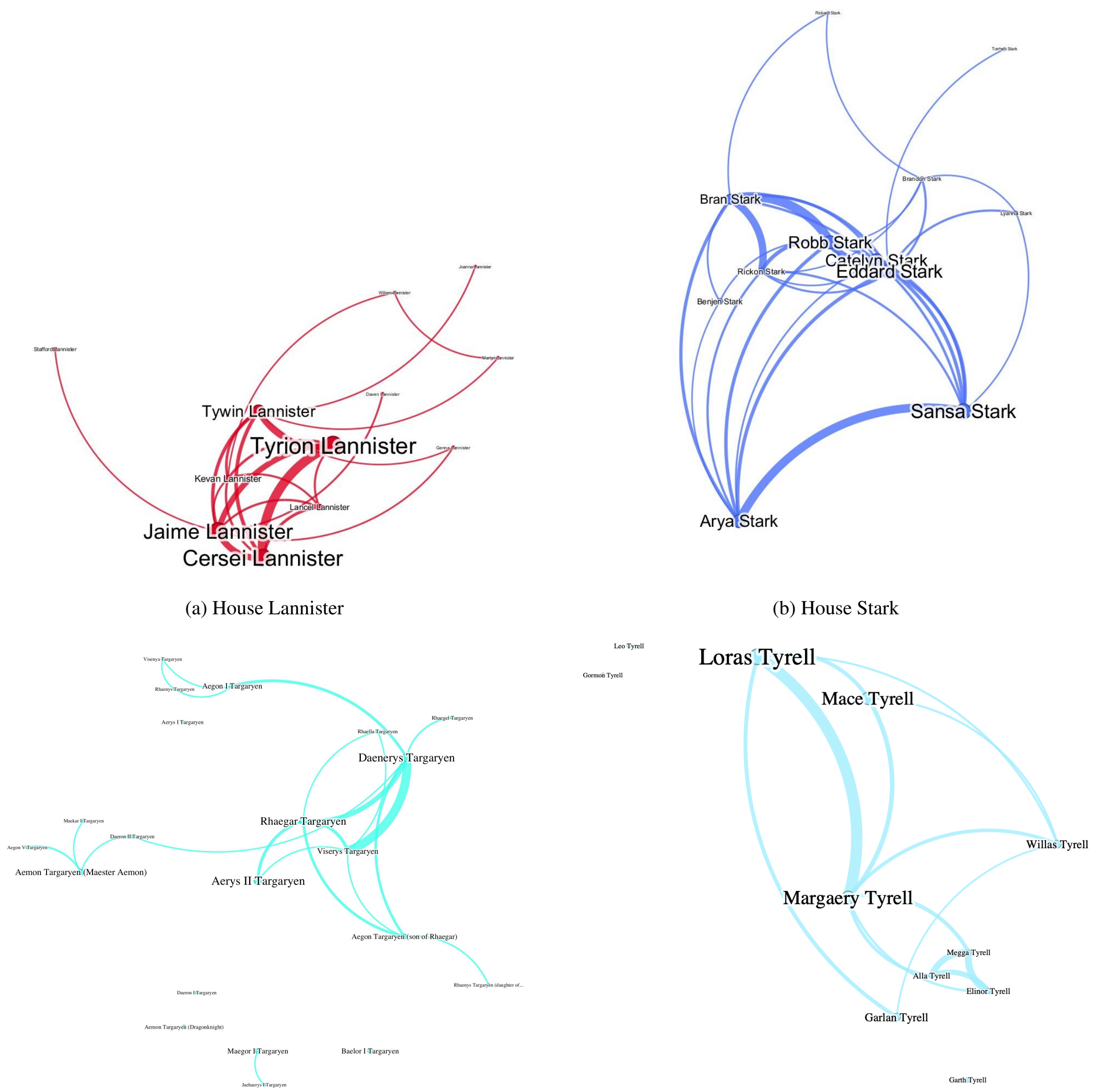


(h) House Tarly



(i) House Tully

Figure 14: House's graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight



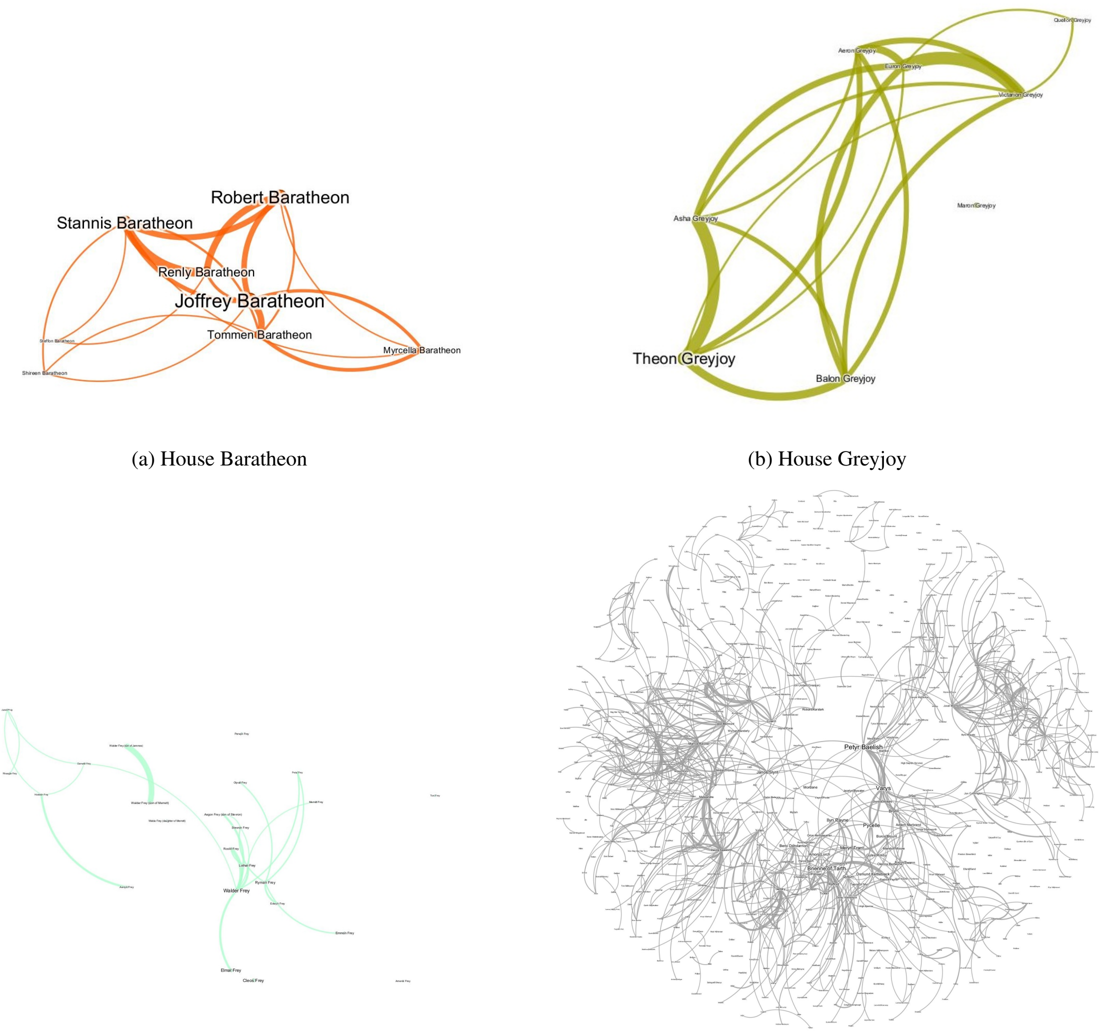
(a) House Lannister

(b) House Stark

(c) House Targaryen

(d) House Tyrell

Figure 15: House's graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight



(a) House Baratheon

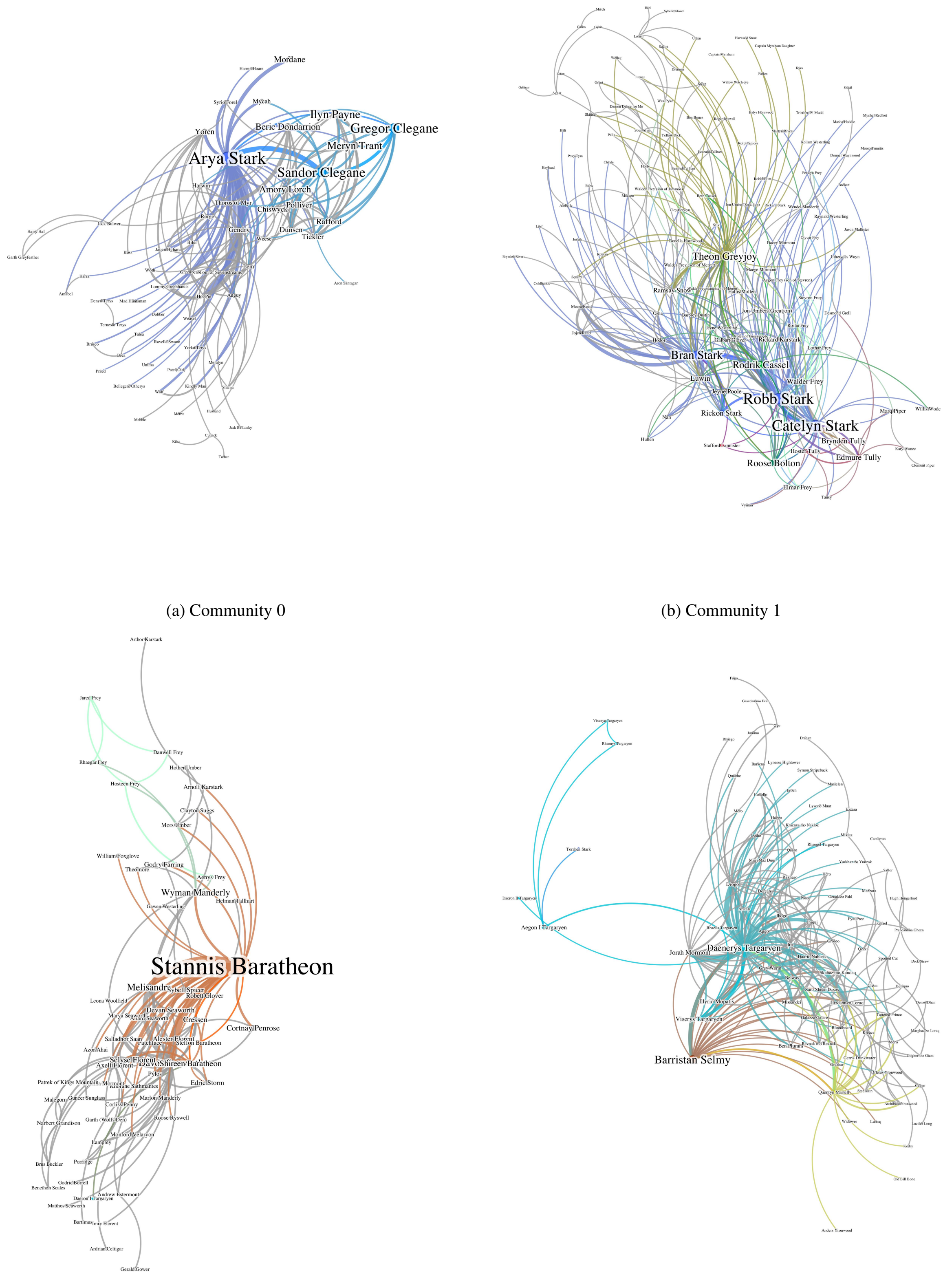
(b) House Greyjoy

(c) House Frey

(d) No house associated

Figure 16: House's graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight

A.2 Community Analysis



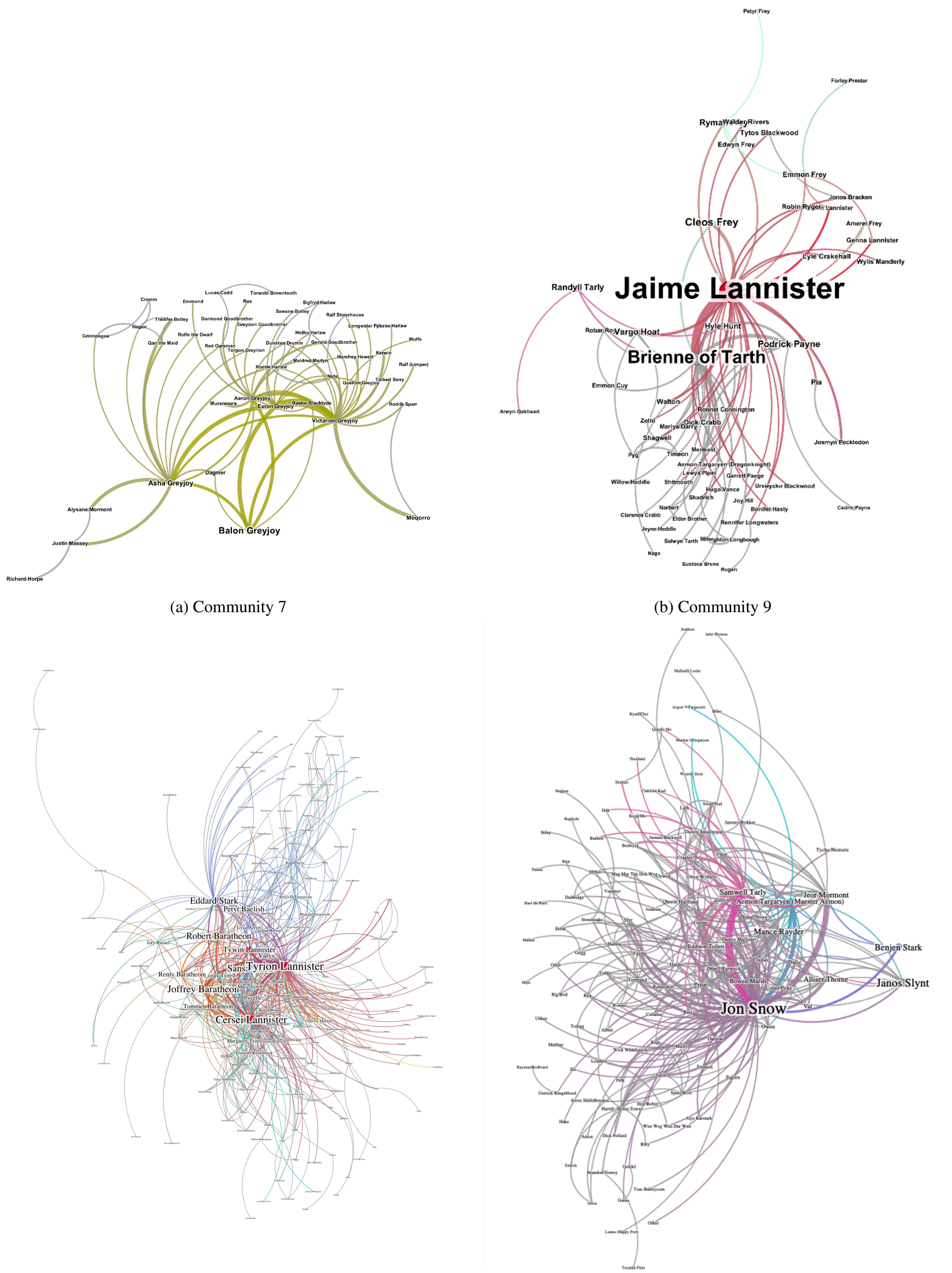
(a) Community 0

(b) Community 1

(c) Community 2

(d) Community 6

Figure 17: Communities' graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight



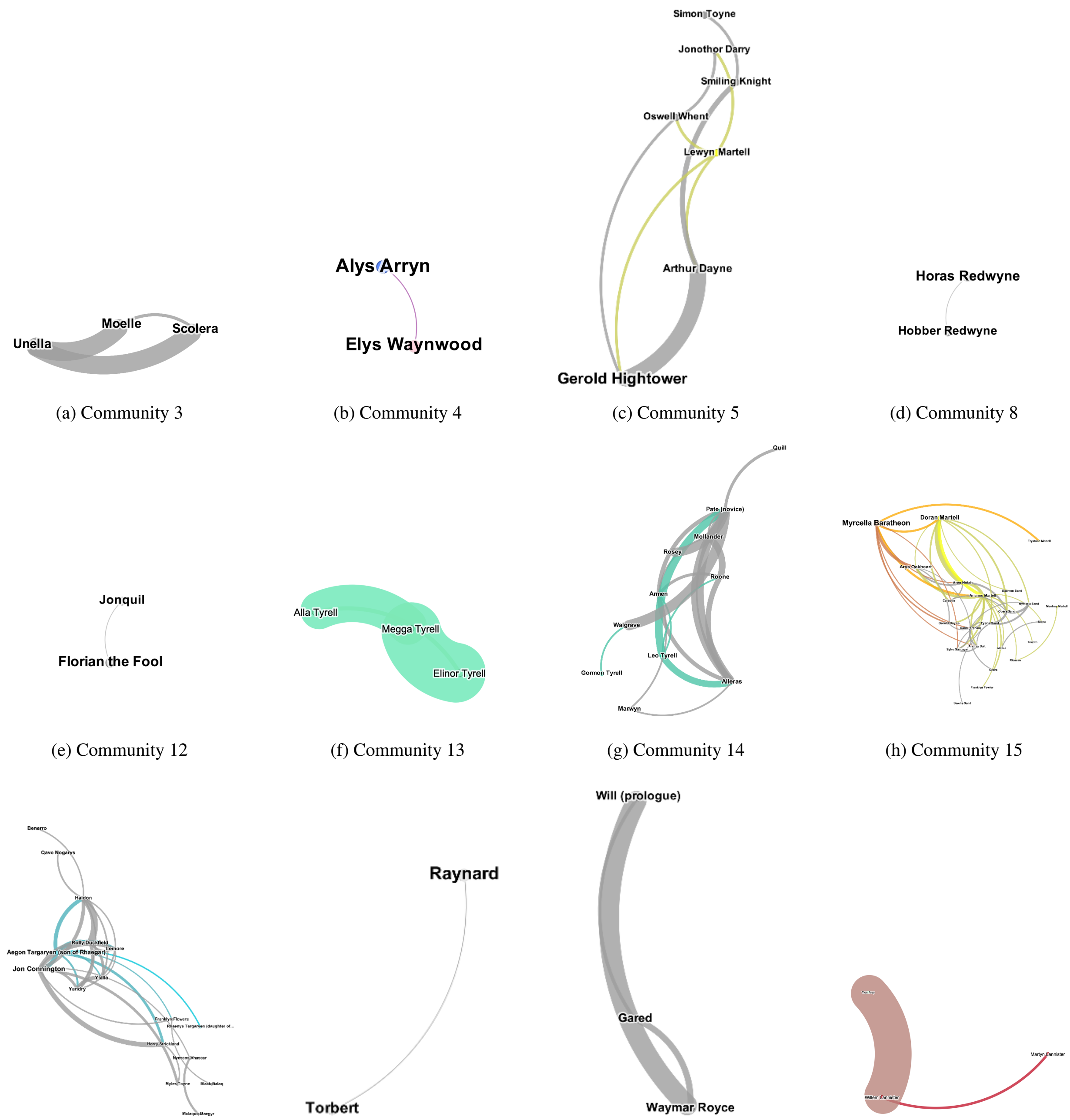
(a) Community 7

(b) Community 9

(c) Community 10

(d) Community 11

Figure 18: Communities' graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight



(a) Community 3 (b) Community 4 (c) Community 5 (d) Community 8

(e) Community 12 (f) Community 13 (g) Community 14 (h) Community 15

(i) Community 16 (j) Community 17 (k) Community 18 (l) Community 19

Figure 19: Communities' graphs, the size of the nodes are proportional to their authority values, the size of the edges is proportional to their weight

A.3 Measures of centrality Analysis

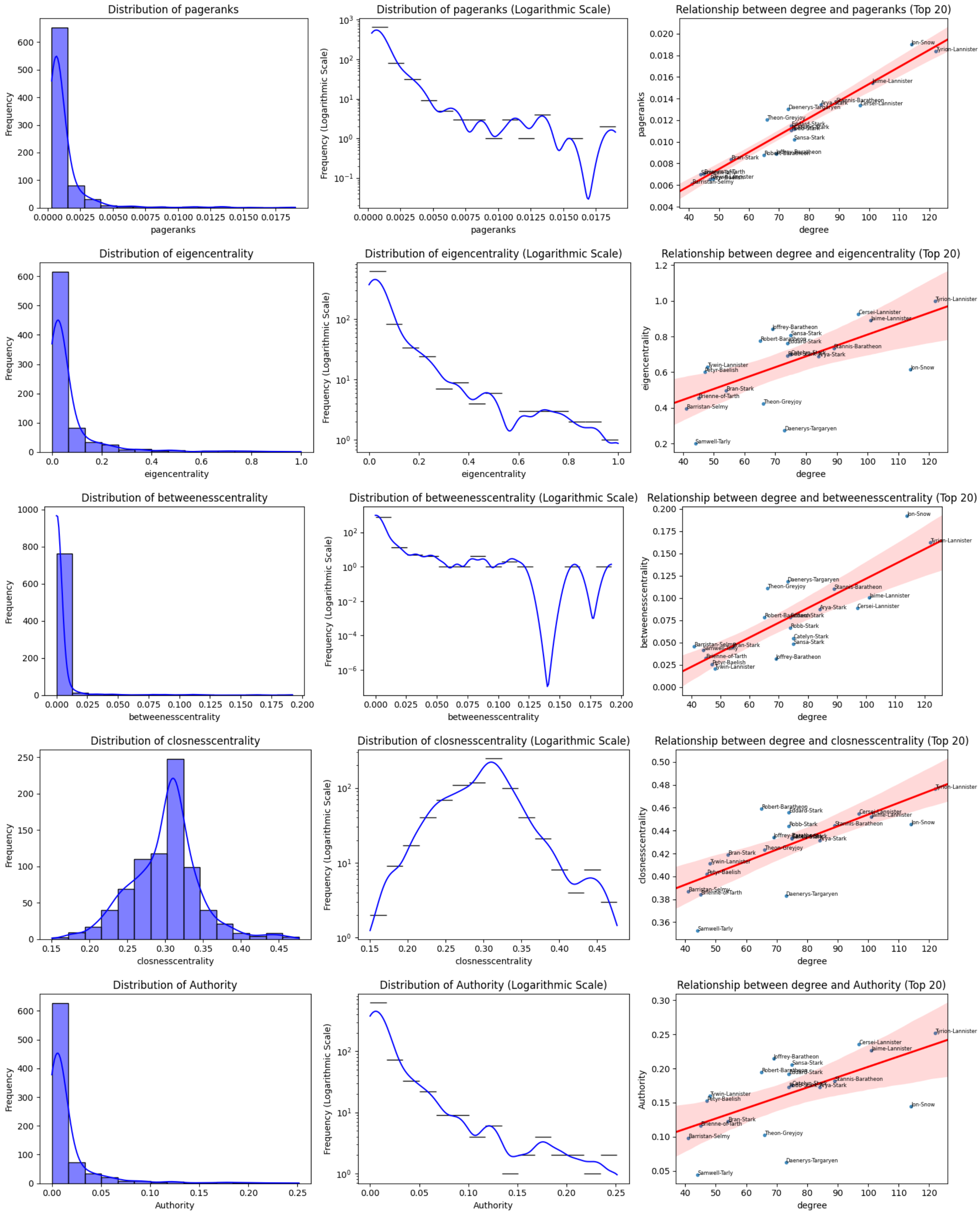


Figure 20: Distribution and relationship of measures of centrality analysis

A.4 Epidemic simulation

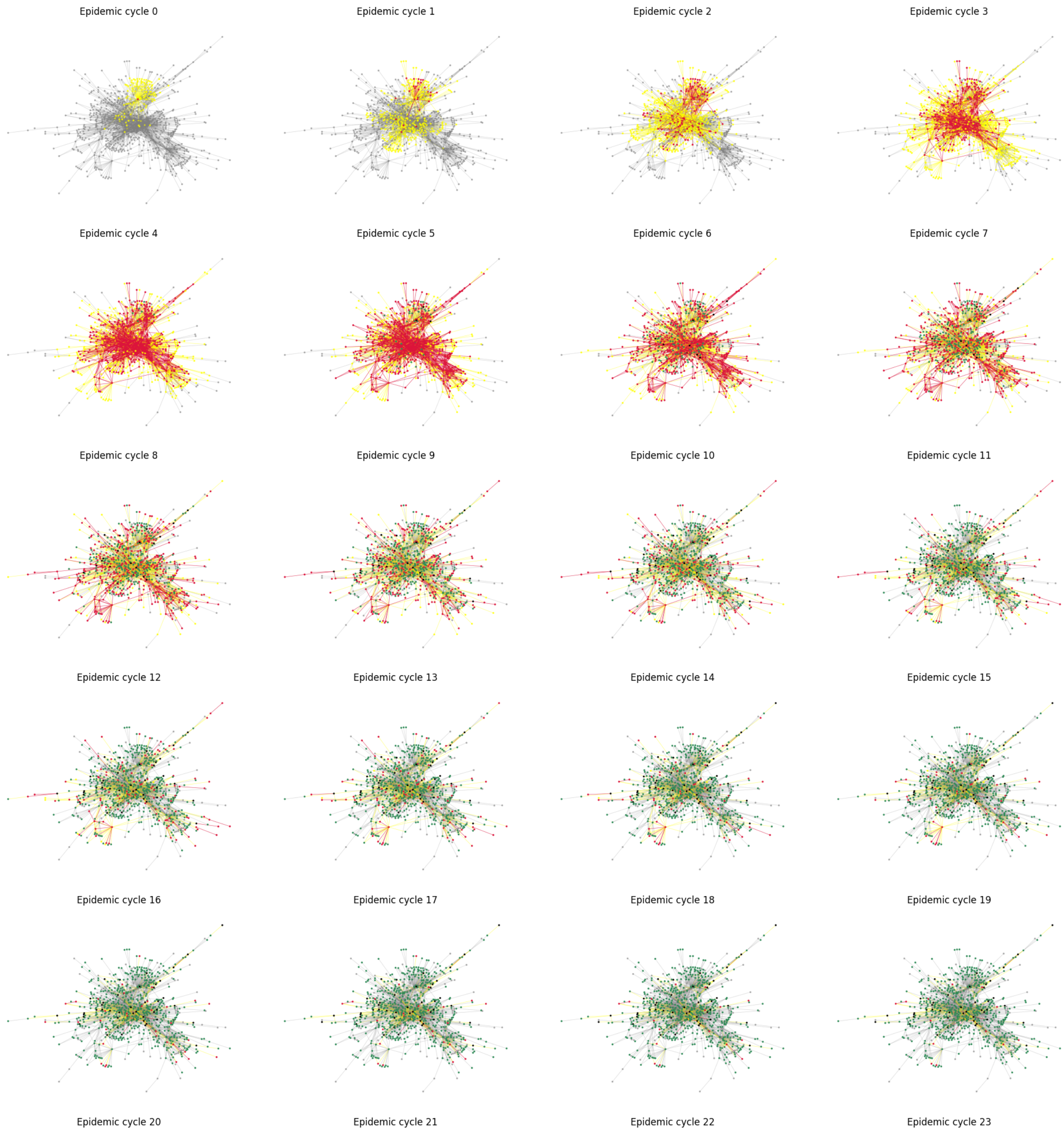


Figure 21: Epidemic (COVID-19) stages related to simulation in figure 12a on Jon Snow

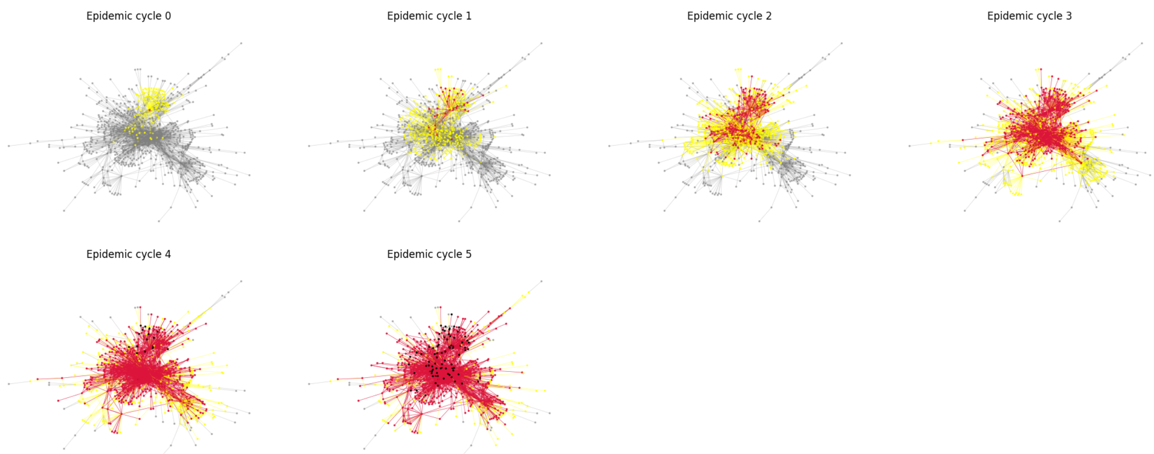


Figure 22: Epidemic (plague) stages related to simulation in figure 12b on Jon Snow

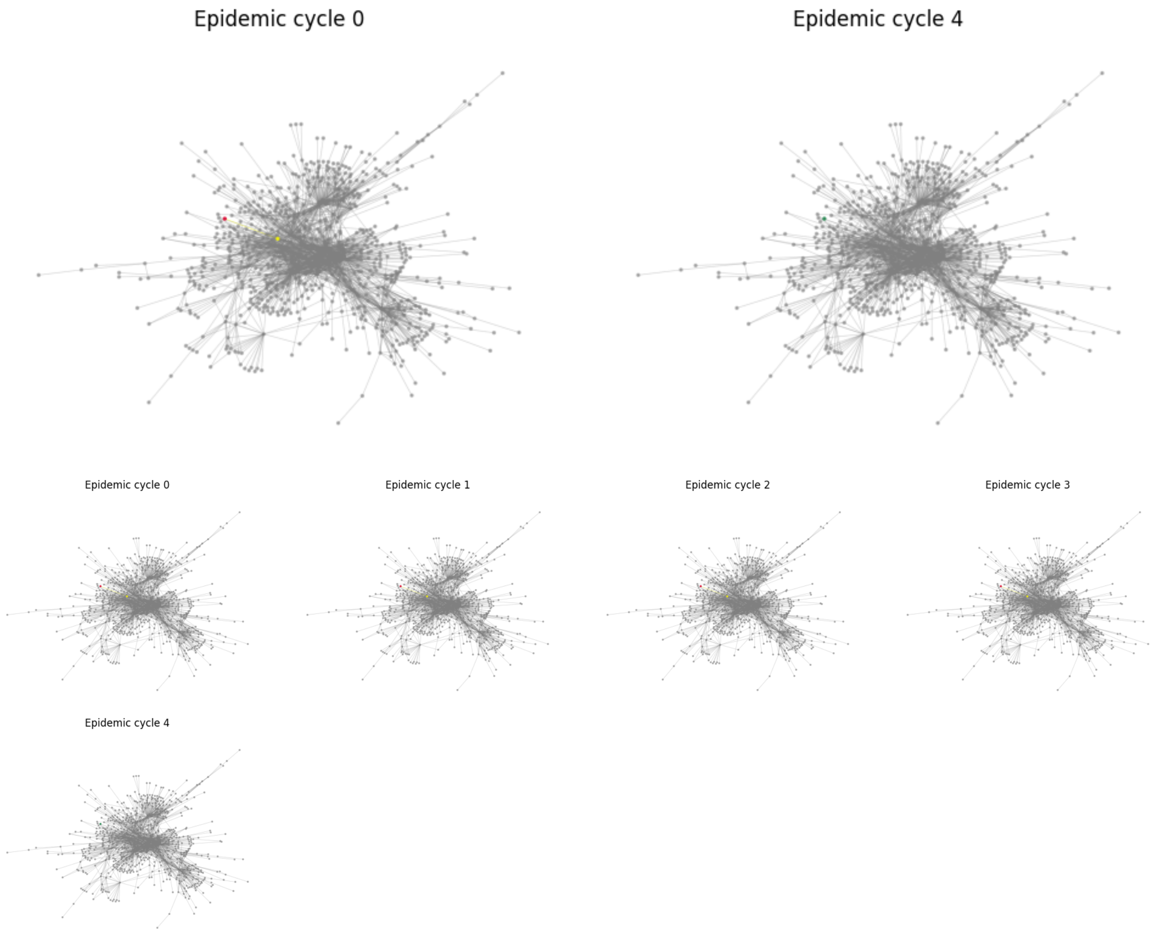


Figure 23: Epidemic (covid) interruption with low infection with patient 0 Poxy-Tym

Figure 24: Epidemic (plague) interruption with low infection with patient 0 Steffon-Varner

A.5 Epidemic trend Epidemic trend versus epidemic cycles

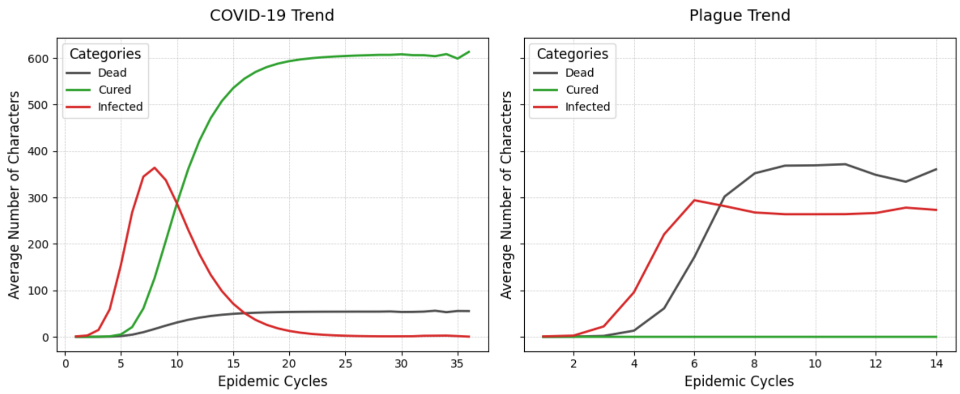
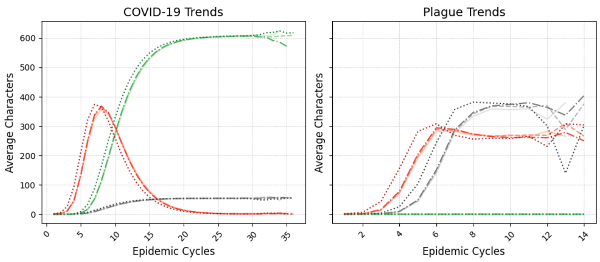
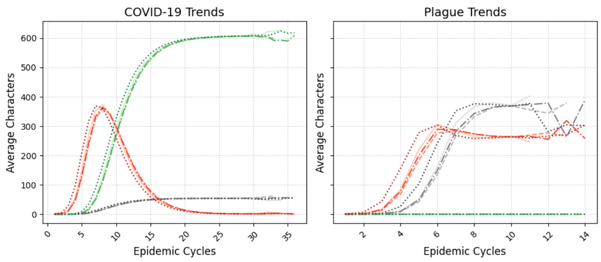


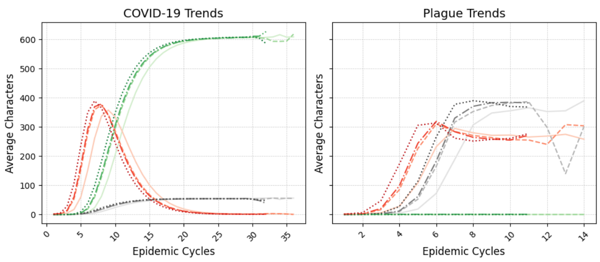
Figure 25: Epidemic trend versus epidemic cycles



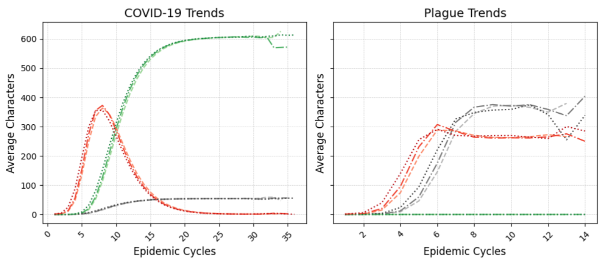
(a) Epidemic trend by quartiles of degree measure



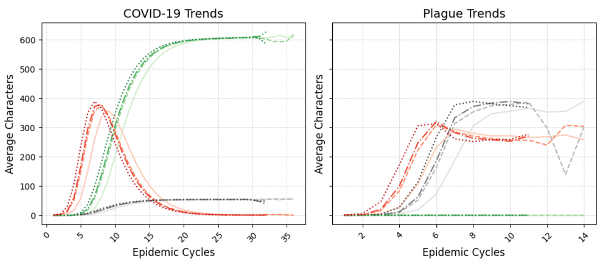
(b) Epidemic trend by quartiles of pagerank measure



(c) Epidemic trend by quartiles of betweeness measure



(d) Epidemic trend by quartiles of the betweeness measure

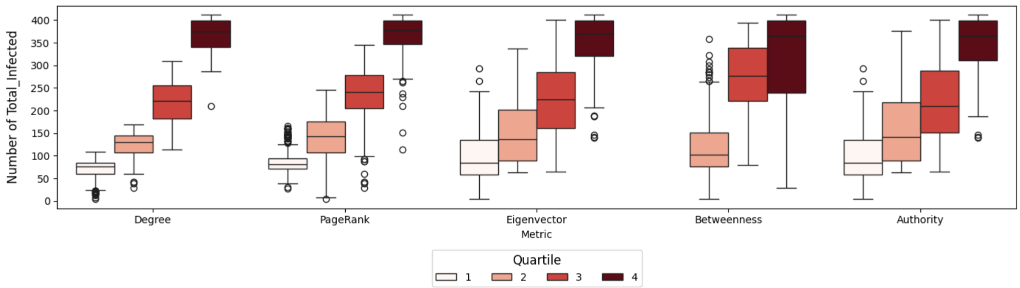


(e) Epidemic trend by quartiles of the authority measure

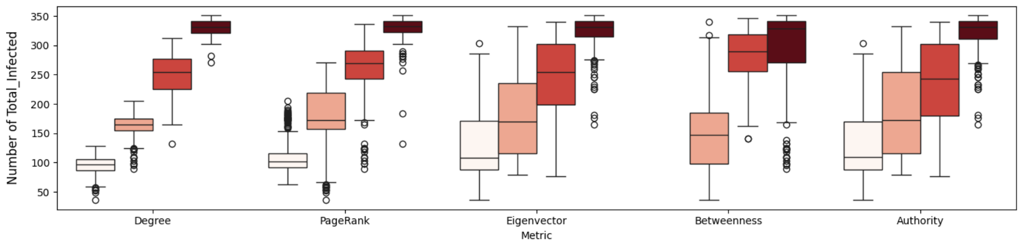
Figure 26: Epidemic trend versus epidemic cycles

A.6 Simulation results versus quartile measures

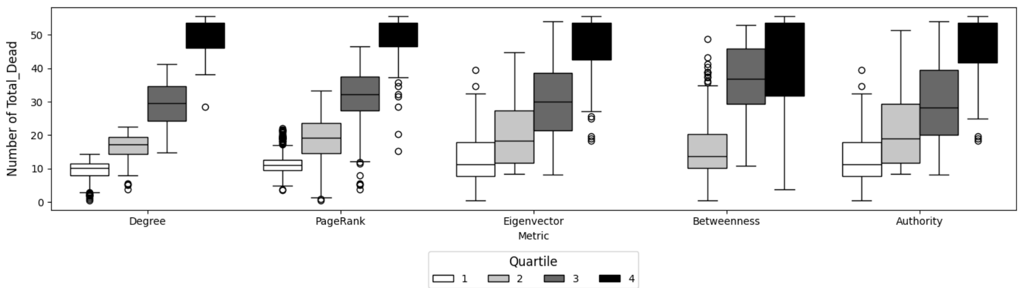
COVID-19 - Centrality Metrics by Quartile



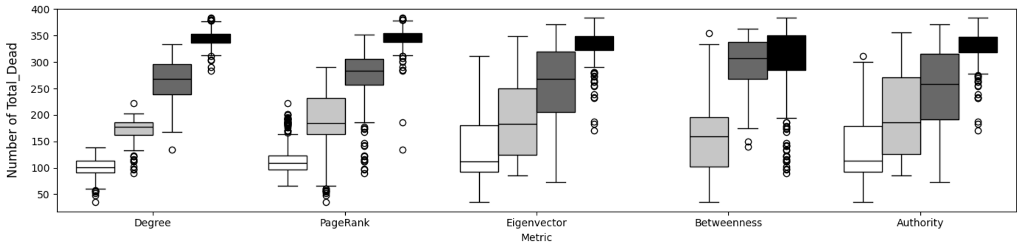
Plague - Centrality Metrics by Quartile



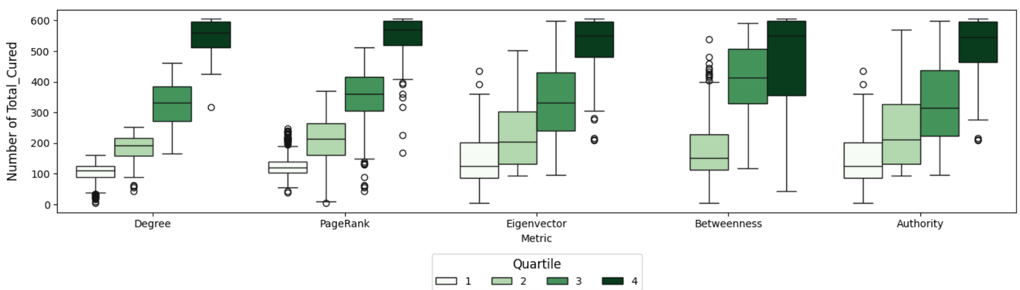
COVID-19 - Centrality Metrics by Quartile



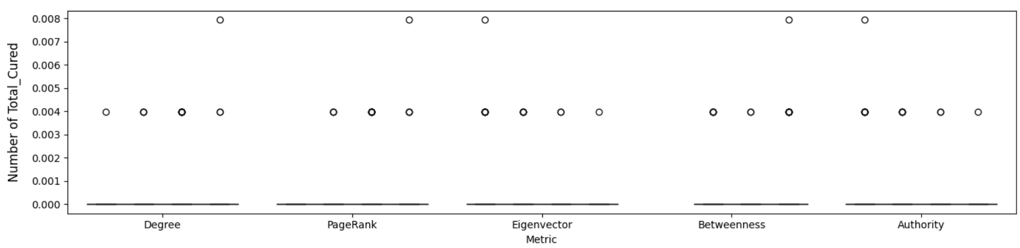
Plague - Centrality Metrics by Quartile



COVID-19 - Centrality Metrics by Quartile



Plague - Centrality Metrics by Quartile



A.7 Distribution of percentage results of epidemics