

SOCIAL NETWORK ANALYSIS

II PROGETTO PER DATA SCIENCE

Anno accademico 2024-2025

AUTORI:

Gabriel Piercecchi

Tosca Pierro

Università Politecnica delle Marche

Facoltà di Ingegneria

Dipartimento di Ingegneria dell'Informazione

Corso di Laurea Magistrale in Ingegneria Informatica e dell'Automazione



Tesina di:

Data Science

Esame per il corso tenuto dal Prof. Domenico Ursino,
durante l'anno accademico 2024-2025

9 CFU

Professori:

Domenico Ursino

Michele Marchetti

Redazione del documento a cura di:

- ◆ **Piercecchi Gabriel** (Matr. 1120541) - s1120541@studenti.univpm.it
 - ◆ **Pierro Tosca** (Matr. 1120542) - s1120542@studenti.univpm.it
-

Anno Accademico 2024-2025

Anno II

SOCIAL NETWORK ANALYSIS

Indice

1 Introduzione	4
1.1 Social Network Analysis	4
1.1.1 NetworkX	4
2 Analisi del Dataset	5
2.1 Dataset	5
2.2 ETL	5
2.3 Visualizzazione Dataset	5
3 Social Network Analysis	7
3.1 NetworkX	7
4 Analisi della Centrality	13
4.1 Cos'è la Centrality	13
4.2 Closeness centrality	14
4.3 Betweenness Centrality	17
4.4 Eigenvector Centrality	19
4.5 Degree Centrality	23
5 Analisi delle strutture	26
5.1 Triade	26
5.2 Clique	27
5.3 K-core	28
5.4 Ego Graph	29

INTRODUZIONE

1.1 SOCIAL NETWORK ANALYSIS

La Social Network Analysis, o (SNA), è un metodo di ricerca utilizzato per visualizzare e analizzare le relazioni e le connessioni tra entità o individui all'interno di una rete.

Immagina di mappare le relazioni tra i diversi dipartimenti di un'azienda. Il risultato sarebbe un'immagine vivida di come ogni dipartimento interagisce con gli altri, permettendoci di vedere i modelli di comunicazione, le entità influenti e i colli di bottiglia.

La SNA consente di esplorare la struttura sottostante di un'organizzazione o di una rete, identificando le relazioni formali e informali che guidano i processi e i risultati ufficiali. Questa visione può favorire una comunicazione migliore, facilitare la gestione del cambiamento e ispirare una collaborazione più efficiente.

Questa metodologia aiuta anche a dimostrare l'impatto delle iniziative di costruzione di relazioni e di cambiamento dei sistemi, documentando i cambiamenti nella qualità e quantità delle relazioni prima e dopo l'iniziativa. Le mappe e le visualizzazioni prodotte dalla SNA sono un modo coinvolgente per condividere i progressi e l'impatto con gli stakeholder, i donatori e la comunità in generale.

1.1.1 NETWORKX

In questo progetto, per analizzare le dinamiche relazionali e le interazioni tra gli individui, è stata utilizzata NetworkX, una libreria Python specializzata nell'analisi e manipolazione di grafi e reti complesse. NetworkX ha permesso di modellare la struttura organizzativa sotto forma di grafo, dove i nodi rappresentano gli individui e gli archi indicano le connessioni tra di essi. Attraverso questa rappresentazione, sono state applicate diverse metriche di centralità, come degree centrality, betweenness centrality, closeness centrality ed eigenvector centrality per individuare gli individui più influenti nella rete.

ANALISI DEL DATASET

2.1 DATASET

In questa relazione si analizza il dataset relativo ai libri del Trono di Spade (disponibile al seguente indirizzo: <https://www.kaggle.com/datasets/mmmarchetti/game-of-thrones-dataset>) che raccoglie informazioni dettagliate sui personaggi dei libri.

2.2 ETL

L'analisi prende avvio con l'estrazione dei dati dai file del dataset, forniti in formato .csv, mediante l'utilizzo della libreria pandas per la lettura dei file.

Una volta importati i dati, viene condotta un'analisi preliminare tramite i metodi `isnull()`, `sum()`, `describe()` e `info()`, al fine di comprendere la struttura del dataset e verificare l'eventuale presenza di valori nulli. Ogni file CSV presenta le seguenti colonne:

- ◆ **Source:** Il personaggio di origine della relazione.
- ◆ **Target:** Il personaggio di destinazione della relazione.
- ◆ **Type:** Il tipo di relazione, che in questo caso è sempre *Undirected*, indicando che la relazione non ha una direzione prefissata.
- ◆ **Weight:** Il peso associato alla relazione, che può rappresentare l'intensità o la forza della connessione tra i personaggi. Il peso è suddiviso in vari intervalli numerici.
- ◆ **Book:** Il libro di riferimento in cui si trova la relazione. In questo caso, tutti i dati si riferiscono al primo libro.

2.3 VISUALIZZAZIONE DATASET

Nel contesto dell'analisi della rete sociale, è fondamentale comprendere quali personaggi abbiano un ruolo più centrale all'interno della narrazione. Un metodo efficace per ottenere una visione preliminare della centralità dei personaggi consiste nell'analizzare il numero di volte in cui ciascun personaggio compare nelle interazioni presenti nei diversi libri.

Attraverso l'esame delle connessioni nel dataset, viene effettuato il conteggio delle occorrenze di ogni personaggio in qualità di fonte (Source) o destinatario (Target) di

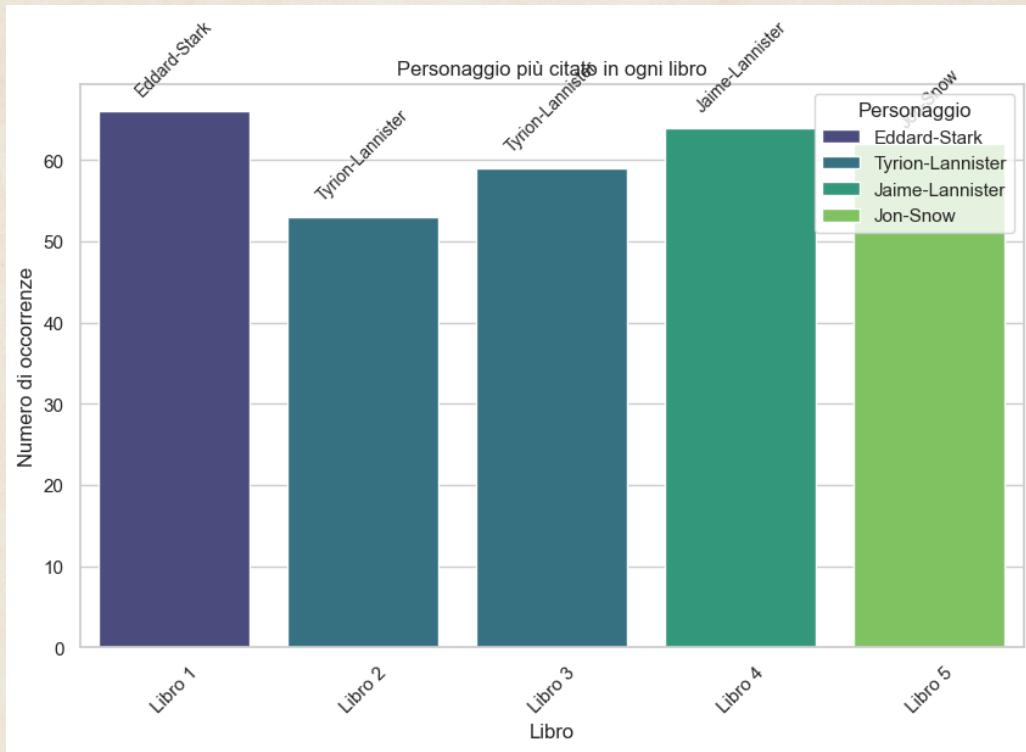


Figura 2.1: Personaggi più presenti nei 5 Libri

un'interazione.

I risultati sono presentati nel grafico 2.1, il quale mostra il personaggio più presente in ciascun libro. Si osserva come *Eddard Stark* risulti il personaggio maggiormente nominato nel primo volume, in quanto protagonista principale per gran parte della trama, fino al momento della sua morte.

Successivamente, l'attenzione narrativa si concentra su *Tyrion Lannister*, che assume un ruolo centrale nei due libri successivi. Nel secondo volume, Tyrion ricopre la carica di «Mano del Re» ad Approdo del Re, dove si trova a gestire complotti e intrighi politici; nel terzo libro, viene ingiustamente accusato della morte di Joffrey Lannister e arrestato, con il proprio destino sospeso in una situazione di forte tensione.

Nel quarto libro della saga emerge *Jaime Lannister* come figura centrale, la cui narrazione è incentrata su un percorso di redenzione e trasformazione personale. Dopo aver perso la mano destra, il personaggio affronta un difficile recupero sia fisico sia morale.

Infine, a partire dal quinto volume, *Jon Snow* assume un ruolo da protagonista, in virtù della sua posizione come Lord Comandante della Guardia della Notte. Le sue decisioni, cruciali per la difesa del Muro e per la lotta contro la minaccia rappresentata dagli Estranei, lo pongono al centro degli eventi principali della narrazione.

SOCIAL NETWORK ANALYSIS

3.1 NETWORKX

Come discusso nel capitolo precedente, nella sezione 1.1.1, per condurre l'analisi delle reti sociali è stata impiegata la libreria *NetworkX*, uno strumento potente per la creazione, la manipolazione e l'analisi di grafi complessi. NetworkX consente di rappresentare e studiare le relazioni tra i personaggi attraverso un grafo, in cui i nodi corrispondono ai personaggi e gli archi rappresentano le interazioni tra di essi.

Dopo aver analizzato il numero di apparizioni dei personaggi nei diversi libri, si procede con la visualizzazione delle relazioni tra di loro. L'analisi delle connessioni risulta fondamentale per comprendere la struttura delle interazioni nella rete sociale della saga.

A tal fine, viene utilizzato il grafo completo (Figura 3.1), che rappresenta tutte le connessioni tra i personaggi, basandosi sulle interazioni presenti nei dati. Ogni nodo del grafo rappresenta un personaggio, mentre ogni arco corrisponde a un'interazione tra due personaggi, con un peso che indica la frequenza di tali interazioni.

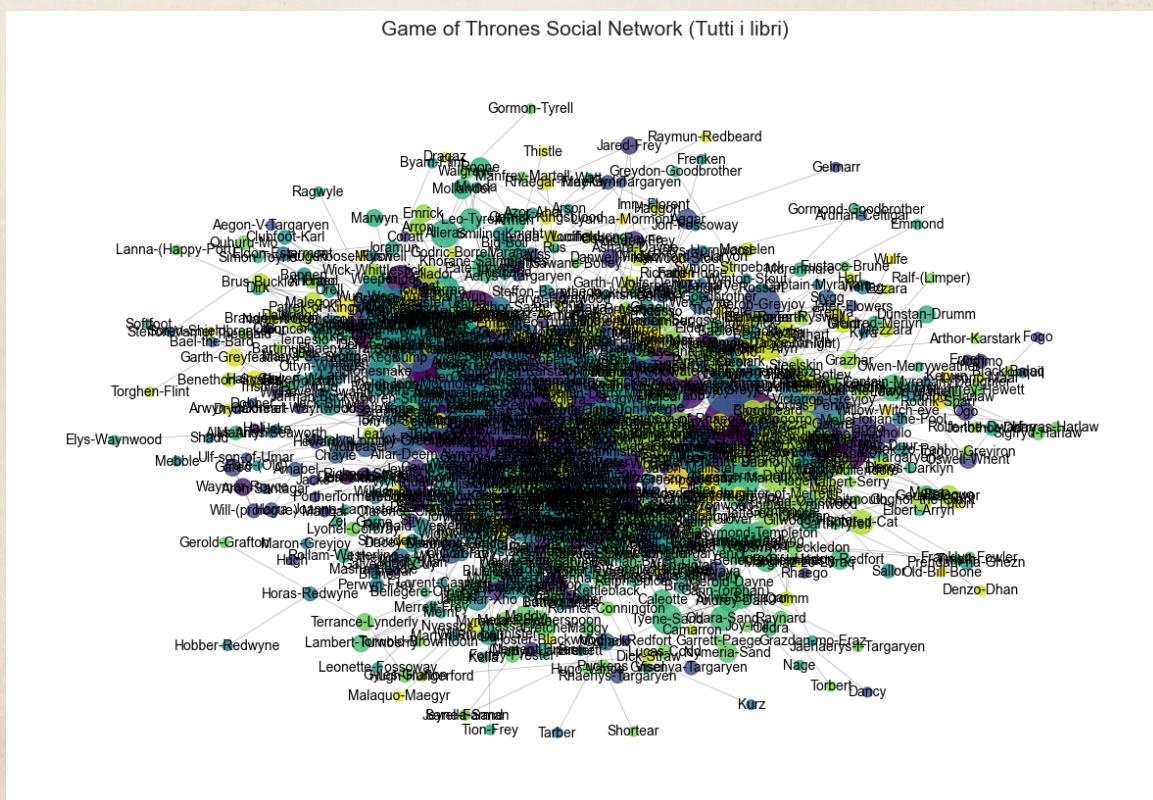


Figura 3.1: Grafo relazionale

Il grafo 3.1 rappresenta la rete sociale completa, evidenziando la distribuzione delle connessioni tra tutti i personaggi. Le dimensioni dei nodi sono proporzionali al numero di interazioni associate a ciascun personaggio: all'aumentare del numero di connessioni, corrisponde un nodo di dimensioni maggiori. Tale rappresentazione consente di individuare con facilità i personaggi centrali all'interno della trama. Tuttavia, a causa della complessità e densità della rete, la leggibilità del grafo risulta compromessa. Per ovviare a questo problema, l'analisi è stata focalizzata sui 20 personaggi con il maggior numero di connessioni, i quali ricoprono un ruolo cruciale nello sviluppo della storia, interagendo con numerosi altri personaggi e fungendo da snodi centrali nella rete sociale.

Il grafo 3.2 mostra esclusivamente questa sottorete, offrendo una visualizzazione più chiara delle relazioni tra i personaggi principali e permettendo di evidenziare le dinamiche tra i nodi più rilevanti.

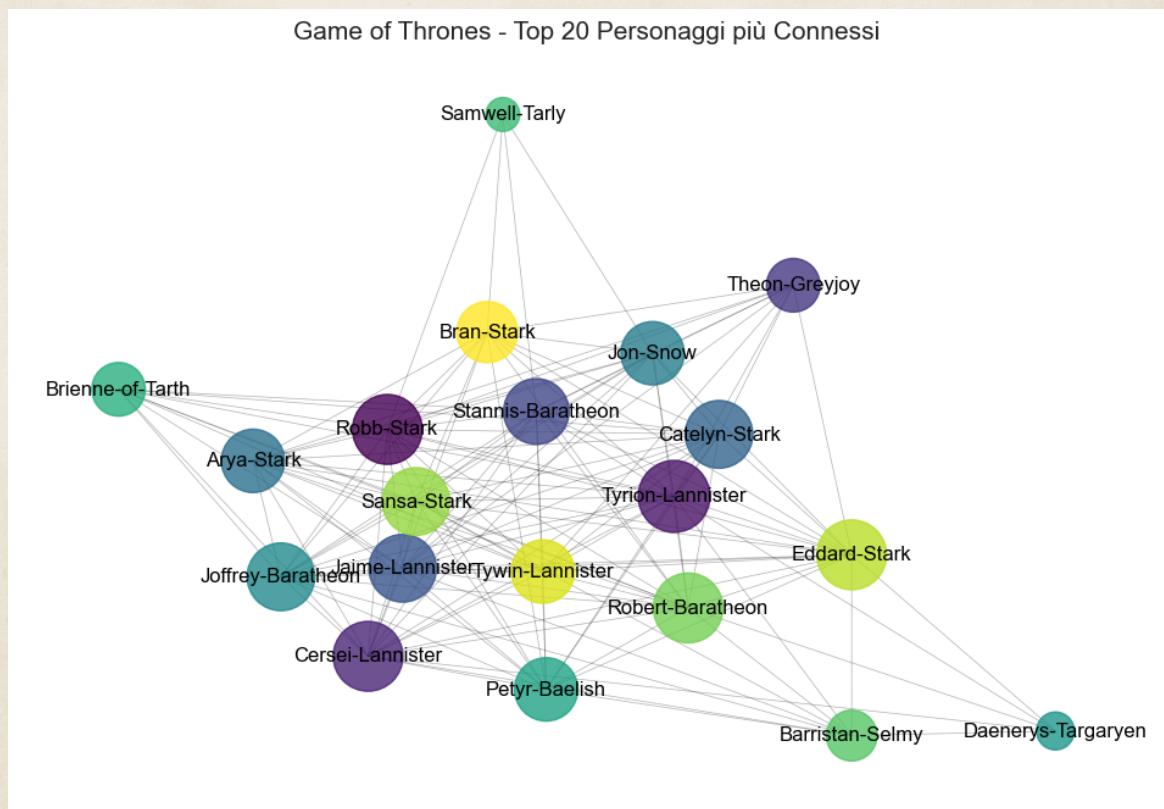


Figura 3.2: Grafo relazionale dei 20 personaggi con più relazioni

È possibile osservare come i personaggi più centrali nella trama de *Il Trono di Spade* appartengano frequentemente alle casate più influenti del continente di Westeros, quali i Lannister, gli Stark, i Targaryen e i Baratheon. Tale centralità riflette l'importanza politica e narrativa di queste famiglie all'interno della saga.

Nell'analisi rappresentata in Figura 3.3, sono stati estratti e ordinati gli archi della rete in base al peso, selezionando i dieci con valore più elevato. Il peso di ciascun arco

rappresenta l'intensità della relazione tra i nodi, che in questo contesto corrispondono ai personaggi e alle loro connessioni. Una maggiore intensità suggerisce una frequenza più elevata di interazioni tra due personaggi, indicando una relazione narrativa particolarmente significativa.

Nodo 1	Nodo 2	Peso
Daenerys-Targaryen	Hizdahr-zo-Loraq	96
Theon-Greyjoy	Ramsay-Snow	87
Daenerys-Targaryen	Barristan-Selmy	75
Jon-Snow	Stannis-Baratheon	73
Tyrion-Lannister	Penny	59
Daenerys-Targaryen	Quentyn-Martell	58
Jon-Snow	Tyrion-Lannister	56
Tyrion-Lannister	Haldon	54
Bran-Stark	Luwin	51
Daenerys-Targaryen	Daario-Naharis	50

Figura 3.3: Top 10 archi con maggior peso

Daenerys Targaryen emerge come una figura centrale all'interno della rete, risultando presente in numerosi archi con i pesi più elevati. Il legame più marcato si riscontra con *Hizdahr zo Loraq*, interpretato come un'alleanza di natura politica. Un'altra connessione significativa è quella tra *Theon Greyjoy* e *Ramsay Snow*, il cui alto peso suggerisce una relazione caratterizzata da una forte interazione; in questo caso, si tratta di un rapporto di sottomissione e abuso.

L'analisi è stata approfondita mediante l'applicazione di diverse metriche, illustrate in Figura 3.4, con l'obiettivo di comprendere meglio l'organizzazione e la distribuzione delle connessioni tra i nodi della rete. Il grafo analizzato è composto da 796 nodi e 2823 archi, configurandosi come una rete relativamente complessa, con un'elevata densità di interazioni tra i personaggi.

```

Nodi: 796
Archi: 2823
Density: 0.008921968332227173
Avarage Length: 10.769005847953217
Clustering: 0.486
Periferia: ['Murch', 'Simon-Toyne', 'Gerold-Grafton', 'Gormon-Tyrell', 'Torwold-Brown tooth',
'Manfrey-Martell', 'Lucifer-Long', 'Hugh-Hungerford']
IS Connect: True

```

Figura 3.4: Caratteristiche generali del grafo

La densità della rete risulta pari a 0.0089, valore che indica una struttura piuttosto sparsa. Ciò significa che il numero effettivo di connessioni tra i nodi è molto inferiore rispetto al numero massimo teoricamente possibile. Una tale configurazione è tipica dei grafi reali che seguono una distribuzione di tipo power-law, come nel caso delle reti sociali o delle reti di interazione tra personaggi, in cui pochi nodi risultano altamente connessi, mentre la maggior parte presenta un numero limitato di collegamenti.

La lunghezza media degli archi pesati è pari a 10.77, suggerendo che, in media, l'intensità delle interazioni tra i personaggi si distribuisce attorno a questo valore.

Il coefficiente di clustering medio è pari a 0.486, il che evidenzia una moderata tendenza alla formazione di gruppi o comunità di nodi interconnessi. Questo valore riflette un certo grado di coesione locale nella rete, con alcune regioni caratterizzate da una maggiore densità di connessioni interne.

La periferia del grafo è composta da nodi quali *Murch*, *Simon Toyne*, *Gerold Grafton*, *Gormon Tyrell*, *Torwold Brown tooth*, *Manfrey Martell*, *Lucifer Long* e *Hugh Hungerford*. Questi nodi presentano la massima eccentricità, risultando i più distanti all'interno della rete. La loro posizione indica una debole connessione con il nucleo della rete, collocandoli ai margini della struttura complessiva.

Per comprendere meglio la struttura e la dinamica di una rete, sono state esplorate due metriche: la **distribuzione della distanza tra i nodi** e la **distribuzione dei pesi degli archi**.

La distribuzione della distanza tra i nodi analizza quanto lontani siano i nodi tra loro in termini di cammino minimo, ovvero il numero di passaggi necessari per raggiungere un nodo da un altro. Questa metrica è utile per capire la «vicinanza» complessiva di un nodo all'interno della rete. Se la distanza tra i nodi è breve, significa che la rete è ben connessa, mentre una distanza elevata può suggerire una rete più frammentata.

Questa distribuzione evidenziata nella Figura 3.5, mostra un comportamento tipico delle reti «small-world», in cui la maggior parte dei nodi può essere raggiunta da qualsiasi altro nodo in un numero relativamente ridotto di passaggi. Il fatto che la distanza modale sia pari a 3 e che poche coppie di nodi si trovino a più di 6 – 7 passaggi di distanza dimostra che la rete di personaggi è altamente connessa: nonostante la

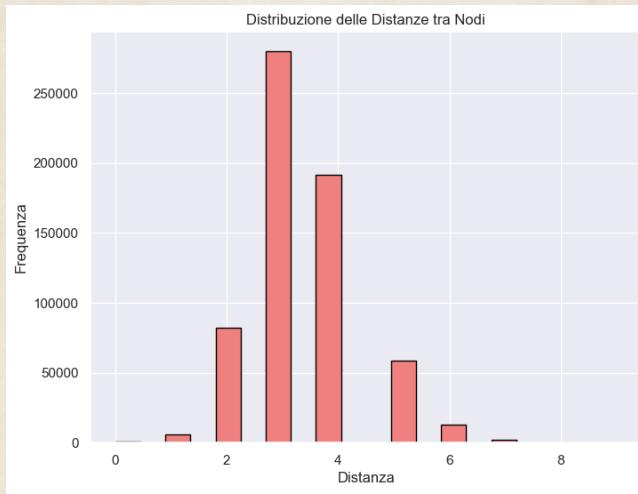


Figura 3.5: Distribuzione delle Distanze tra Nodi

vastità e complessità dell'universo narrativo, le relazioni tra i personaggi si articolano in modo compatto e centralizzato. Il picco pronunciato alla distanza 3 suggerisce l'esistenza di hub o personaggi centrali, che agiscono da veri e propri snodi narrativi, mettendo in collegamento sezioni della rete che altrimenti sarebbero lontane. Questo è coerente con l'impostazione narrativa dei romanzi, dove figure come Tyrion Lannister, Jon Snow o Daenerys Targaryen mantengono un ruolo centrale nella connessione tra sottotrame distanti.

Per quanto riguarda la distribuzione dei pesi degli archi si concentra invece sulle connessioni tra i nodi, che non sono sempre uguali in termini di forza o importanza. In una rete pesata, ogni arco (legame) può avere un peso che rappresenta l'intensità della connessione tra due nodi. L'analisi di come questi pesi sono distribuiti aiuta a capire se esistono legami particolarmente forti che influenzano il comportamento della rete.

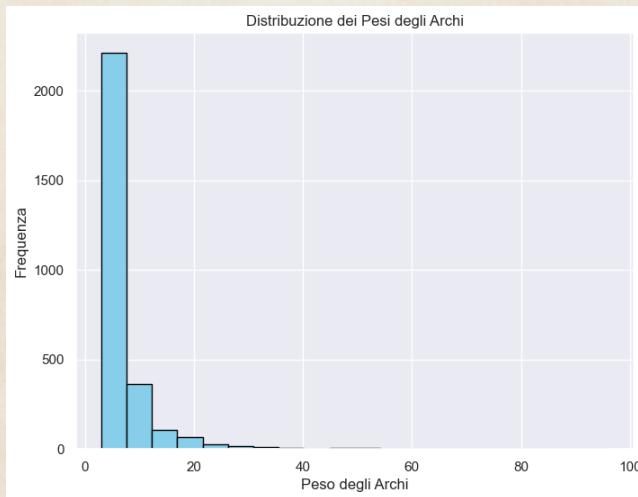


Figura 3.6: Distribuzione dei Pesi degli Archi

La Figura 3.6 mostra i risultati ed è possibile notare che i pesi degli archi seguono una

distribuzione altamente asimmetrica verso destra, suggerendo la presenza di una distribuzione di tipo «heavy-tailed» o con una legge potenzialmente esponenziale o lognormale. Questo implica che nella rete analizzata, la maggior parte degli archi ha un peso basso, mentre solo una piccola frazione ha un peso significativamente più alto. Infatti in Games of Thrones ci sono pochi personaggi principali che sono circondati da tanti personaggi secondari.

ANALISI DELLA CENTRALITY

In questo capitolo si procederà con un'analisi approfondita delle misure di centralità applicate alla rete di personaggi costruita a partire dal dataset narrativo di Game of Thrones. L'obiettivo principale è quello di identificare, tramite strumenti quantitativi, i personaggi più centrali nella rete delle relazioni, analizzando il ruolo che ciascuno di essi occupa nella struttura narrativa complessiva.

4.1 COS'È LA CENTRALITY

Nel contesto della Social Network Analysis (SNA), il concetto di **centralità** rappresenta una misura per valutare l'influenza e l'importanza di un nodo all'interno di una rete. Nel dataset in esame, relativo a *Game of Thrones*, i nodi corrispondono ai personaggi e gli archi rappresentano le interazioni tra di essi. La centralità fornisce indicatori quantitativi che permettono di misurare il «peso» di ciascun personaggio, non solo in termini di numero di connessioni, ma anche in relazione alla qualità e alla posizione strategica nella rete.

La centralità è fondamentale per comprendere quali personaggi ricoprono un ruolo centrale nelle dinamiche narrative e relazionali, nonché per analizzare come le informazioni o le influenze si propagano all'interno della trama. In altre parole, consente di identificare i personaggi chiave, quelli che, attraverso le loro azioni e interazioni, determinano il corso degli eventi.

A seconda della tipologia di centralità che si decide di misurare, è possibile ottenere diverse interpretazioni del ruolo di un personaggio. Ad esempio:

- ◆ **Closeness centrality** valuta quanto rapidamente un nodo può raggiungere gli altri, rilevando quei personaggi che sono al centro della rete e in grado di influenzare più velocemente l'ambiente circostante.
- ◆ **Betweenness centrality** misura la capacità di un nodo di fungere da ponte tra altri nodi, identificando quei personaggi che svolgono un ruolo cruciale nei passaggi tra differenti sottotrame.
- ◆ **Eigenvector centrality** considera non solo il numero di connessioni, ma anche l'importanza di quelle connessioni, premiando i nodi che sono collegati a nodi già centrali, come nel caso di personaggi che, pur avendo meno interazioni dirette, sono comunque influenti per la loro posizione strategica.

- ◆ **Degree centrality** misura semplicemente il numero di connessioni dirette di un nodo, aiutando a individuare i personaggi che interagiscono con il maggior numero di altri.

4.2 CLOSENESS CENTRALITY

La prima misura di centralità calcolata è la **Closeness Centrality**, che valuta la vicinanza di un nodo rispetto agli altri all'interno di un grafo. In altre parole, un nodo con alta closeness centrality è, in media, più vicino a tutti gli altri nodi, il che implica che può raggiungere qualsiasi altro nodo con un numero minimo di passaggi. Questo parametro è un indicatore dell'influenza di un nodo nella rete, poiché un nodo che può raggiungere rapidamente gli altri ha una maggiore capacità di diffondere informazioni o esercitare potere.

Per calcolare questa misura, è stata utilizzata la funzione `closeness_centrality()` fornita dalla libreria NetworkX.

I risultati ottenuti sono riportati nella tabella di Figura 4.1, che mostra i dieci nodi con i valori più alti di closeness centrality, seguita da una seconda tabella che elenca i dieci nodi con i valori più bassi di questa misura.

Dall'analisi della *closeness centrality* riportata in Figura 4.1, emerge che **Tyrion Lannister**, con un valore elevato, ha la capacità di influenzare e connettersi rapidamente con un ampio numero di altri personaggi, grazie alla sua posizione centrale nella rete. Al contrario, **Lucifer Long**, con un valore basso, risulta più isolato e distante dai nodi cruciali della trama, limitando così la sua capacità di interazione e influenza.

Nel dettaglio, *Tyrion Lannister* è un personaggio che, nel contesto della serie, ricopre un ruolo strategico, trovandosi al centro di molte dinamiche politiche e relazionali. La sua *closeness centrality* elevata suggerisce che Tyrion occupi una posizione privilegiata, capace di interagire facilmente con molte altre figure. Questo rispecchia il suo ruolo nella trama: come «mano del re» nel secondo libro e come consigliere strategico, Tyrion agisce da connettore tra diversi gruppi e poteri. Inoltre, la sua intelligenza, la capacità di mediazione e la propensione a cercare alleanze lo rendono un nodo cruciale nelle dinamiche di King's Landing. La sua *closeness centrality* elevata implica che potrebbe raggiungere e influenzare rapidamente i principali eventi e personaggi della trama. D'altra parte, *Lucifer Long*, un cavaliere del Continente Occidentale che milita nella banda mercenaria della Compagnia del Vento e nominato nel quinto libro della serie, presenta un valore basso di *closeness centrality*. Questo suggerisce che, pur essendo connesso ad altri nodi nella rete narrativa, la sua capacità di accedere rapidamente alle

	Personaggio	Closeness Centrality
0	Tyrion-Lannister	0.476333
1	Robert-Baratheon	0.459272
2	Eddard-Stark	0.455849
3	Cersei-Lannister	0.454545
4	Jaime-Lannister	0.451961
5	Jon-Snow	0.445378
6	Stannis-Baratheon	0.444631
7	Robb-Stark	0.444134
8	Joffrey-Baratheon	0.433952
9	Catelyn-Stark	0.433479

10 rows	v	10 rows x 2 cols
	Personaggio	Closeness Centrality
0	Lucifer-Long	0.187103
1	Manfrey-Martell	0.185835
2	Hugh-Hungerford	0.179742
3	Simon-Toyne	0.178571
4	Walgrave	0.177139
5	Quill	0.177060
6	Rosey	0.177060
7	Murch	0.175072
8	Torwold-Brownooth	0.171743
9	Gormon-Tyrell	0.150511

Figura 4.1: Tabella closeneess centrality

varie interazioni all'interno della trama è limitata. La sua posizione marginale nel contesto più ampio della narrazione ne riduce l'influenza e la velocità con cui può raggiungere altri personaggi chiave.

Per visualizzare meglio la distribuzione di questa misura, è stato utilizzato un *displot* (Figura 4.2), poiché il grafo contiene un numero elevato di nodi e includere tutti i valori di *closeness centrality* in un singolo grafico avrebbe reso difficile l'interpretazione. La distribuzione mostra che la maggior parte dei personaggi ha valori relativamente bassi, con picchi per alcuni dei protagonisti che occupano posizioni chiave all'interno della rete di relazioni.

La *closeness centrality* è concentrata principalmente attorno a valori di circa 0.30. Questo indica che la maggior parte dei nodi (personaggi) ha una prossimità media rispetto agli altri, riuscendo a raggiungere il resto della rete in un numero moderato di passaggi. Pochi personaggi presentano valori più elevati (oltre 0.40), segnalando un ruolo particolarmente centrale nel collegare diverse parti della rete. Al contrario, quelli con valori più bassi (vicini a 0.15) risultano più periferici, impiegando più passaggi per connettersi agli altri. Questo fenomeno è dovuto al fatto che in una trama così fitta e articolata, le relazioni non sono distribuite uniformemente, con i personaggi marginali o meno influenti che tendono a risultare più periferici all'interno della rete complessiva.

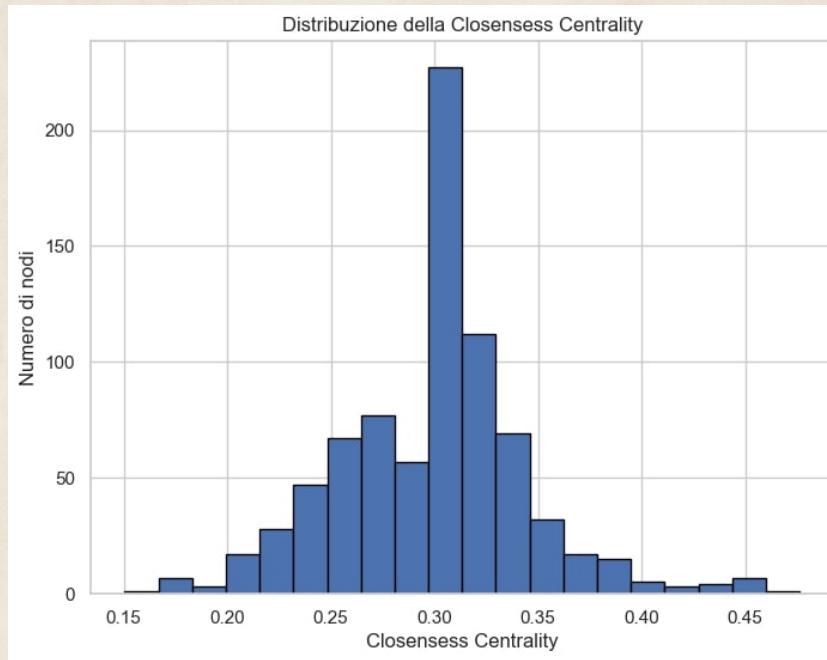


Figura 4.2: Distribuzione della Closeness Centrality

Infine, è stata generata una heatmap della *closeness centrality* (Figura 4.3), che evidenzia chiaramente i nodi più centrali e quelli più periferici. I nodi con valori di *closeness centrality* più elevati (tendenti al giallo) sono quelli che si trovano, in media, a pochi passaggi da tutti gli altri. Questi personaggi occupano una «posizione strategica» nel grafo, potendo diffondere informazioni (o influenza) più velocemente rispetto ai nodi periferici. Nel complesso, la mappa mostra un nucleo relativamente ampio di personaggi con valori intermedi (arancione/rosa), circondato da alcuni nodi con valori più bassi in periferia e da pochi nodi altamente centrali con valori vicini al massimo.

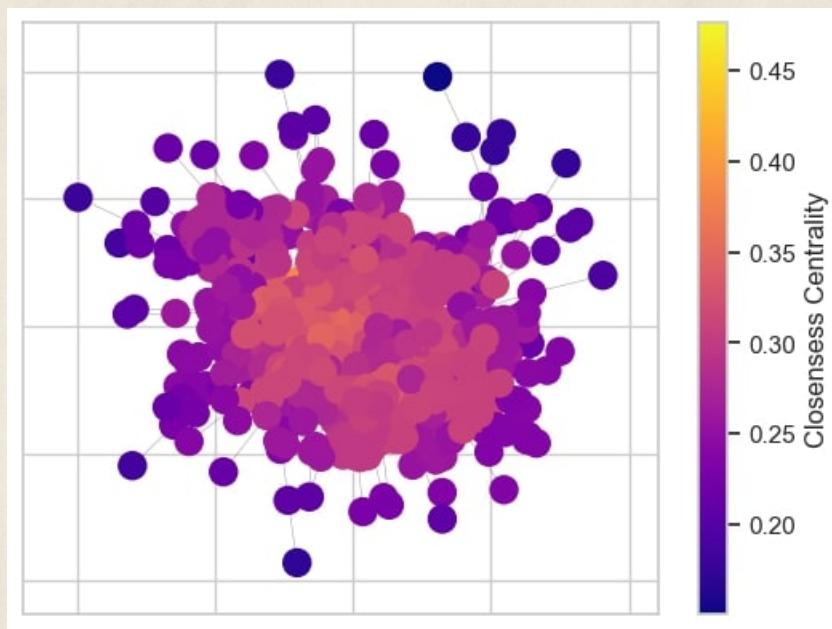


Figura 4.3: Heatmap Closeness Centrality

4.3 BETWEENNESS CENTRALITY

L'analisi prosegue con la **Betweenness Centrality**, una misura che calcola il grado in cui un nodo agisce da «ponte» tra altri nodi della rete. In altre parole, un nodo con una elevata betweenness centrality ha un ruolo cruciale nel connettere gruppi separati di nodi, fungendo da punto di passaggio per le comunicazioni tra di essi. Questa misura è particolarmente utile per identificare nodi che, pur non avendo un grande numero di connessioni dirette, svolgono un ruolo fondamentale nel mantenere la coesione e la connessione della rete.

La betweenness centrality viene calcolata come il numero di volte che un nodo appare sul cammino più breve tra due altri nodi, normalizzato rispetto al numero totale di cammini più brevi possibili nella rete. Se un nodo è molto centrale nella rete, ci sarà un alto numero di cammini più brevi che passano per quel nodo. Al contrario, un nodo con una betweenness centrality bassa ha meno influenza nel connettere parti distanti della rete.

Per calcolare la betweenness centrality, è stata utilizzata la funzione `betweenness_centrality()` della libreria NetworkX, che fornisce il valore di questa misura per ciascun nodo all'interno della rete.

Nel contesto dei personaggi di *Game of Thrones*, la betweenness centrality ci aiuta a identificare quei personaggi che, pur non essendo sempre al centro delle interazioni, sono cruciali per il passaggio delle informazioni tra gruppi separati di personaggi. I risultati ottenuti sono presentati in una tabella (Figura 4.4) contenente i dieci nodi con il valore di betweenness centrality più alto, seguiti da una tabella con i dieci nodi con il

valore di betweenness centrality più basso.

Personaggio	Betweenness Centrality
0 Jon-Snow	0.192120
1 Tyrion-Lannister	0.162191
2 Daenerys-Targaryen	0.118418
3 Theon-Greyjoy	0.111283
4 Stannis-Baratheon	0.110140
5 Jaime-Lannister	0.100838
6 Cersei-Lannister	0.088704
7 Arya-Stark	0.087242
8 Eddard-Stark	0.078732
9 Robert-Baratheon	0.078228

Personaggio	Betweenness Centrality
0 Yellow-Dick	0.0
1 Walda-Frey-(daughter-of-Merrett)	0.0
2 Roose-Ryswell	0.0
3 Scar	0.0
4 Shrouded-Lord	0.0
5 Theomore	0.0
6 William-Foxglove	0.0
7 Willow-Witch-eye	0.0
8 Thistle	0.0
9 Wulfe	0.0

Figura 4.4: Tabella betweenness centrality

Dall'analisi emerge che **Jon Snow** presenta il valore più alto di Betweenness Centrality, il che indica la sua posizione centrale all'interno degli eventi principali della trama. La sua influenza deriva dalla capacità di mediare tra alleanze e conflitti, fungendo da punto di connessione tra la Guardia della Notte, le casate nobili e le diverse fazioni. Il suo ruolo di intermediario tra gruppi altrimenti separati è cruciale per lo sviluppo delle dinamiche narrative e per il mantenimento dell'ordine in un contesto di alleanze fragili e tensioni interne.

D'altra parte, **Yellow Dick**, un personaggio marginale e un soldato fedele alla Casa Bolton, presenta un valore di Betweenness Centrality molto basso. Questo lo rende lontano dai principali snodi vitali della rete narrativa. La sua posizione periferica riflette la sua relativa insignificanza rispetto agli eventi principali e alle dinamiche di potere che definiscono la trama.

Per visualizzare i risultati di questa analisi, è stato deciso di utilizzare un displot (Figura 4.5) che rappresenta la distribuzione della betweenness centrality in tutta la rete, mostrando un forte sbilanciamento verso sinistra, con un picco molto marcato in prossimità dello zero. Questo indica che la quasi totalità dei nodi nella rete ha un valore nullo o prossimo allo zero, e quindi non è coinvolta nei cammini più brevi tra le altre coppie di nodi. Solo un numero molto ristretto di nodi presenta valori significativamente più alti, suggerendo che questi elementi svolgono un ruolo fondamentale come ponti critici o snodi centrali, facilitando la comunicazione e il collegamento tra diverse sezioni della rete. Tale configurazione evidenzia una struttura

non uniforme, in cui poche entità assumono una posizione strategica per il flusso informativo.

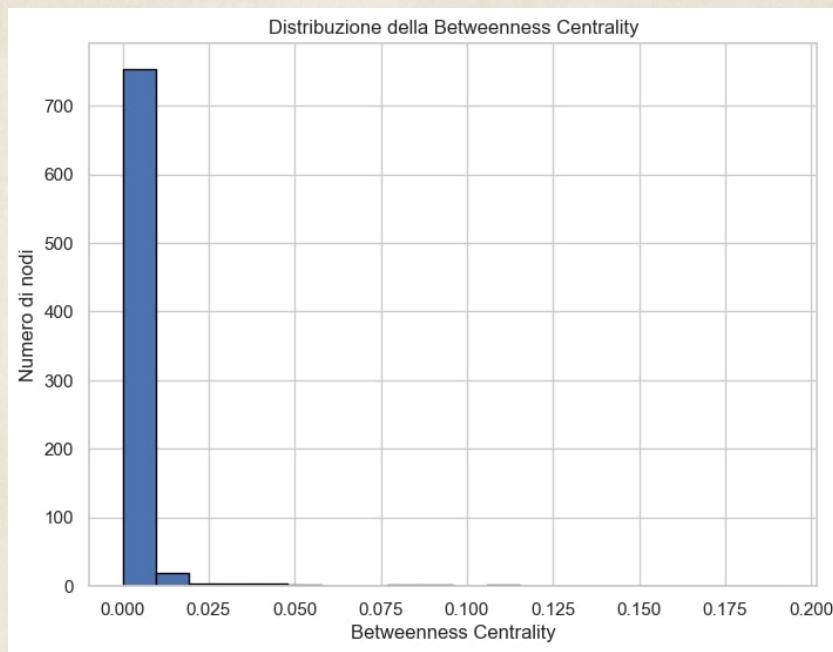


Figura 4.5: Distribuzione della Between Cetnrality

Inoltre, per una visione visiva della centralità di ciascun nodo, è stata utilizzata una heatmap della betweenness centrality (Figura 4.6), che evidenzia i personaggi con il ruolo più centrale attraverso colori più intensi, mentre i nodi periferici appaiono con colori più chiari. La maggior parte dei personaggi presenta valori molto bassi e non svolge ruoli chiave come “ponte” nel grafo. Un gruppo ristretto di nodi, con valori leggermente più alti, funge da passaggio intermedio moderato. Infine, pochissimi nodi (in giallo) agiscono da veri “broker”, collegando aree altrimenti poco connesse e garantendo la coesione dell’intera rete.

4.4 EIGENVECTOR CENTRALITY

Proseguendo con l’analisi delle diverse misure di centralità, è stata esaminata anche l’**Eigenvector Centrality**, la quale è una misura di centralità che non si limita a contare il numero di connessioni dirette di un nodo, ma considera anche la qualità e l’importanza di tali connessioni. In altre parole, un nodo è considerato centrale non solo in base al numero di legami che ha, ma anche in base a quanto siano centrali i nodi con cui è connesso. La logica alla base di questa misura è che se un nodo è collegato a nodi che sono anch’essi molto centrali, allora quel nodo acquisisce una centralità maggiore. In questo modo, l’Eigenvector Centrality premia i nodi che sono connessi a nodi già influenti all’interno della rete.

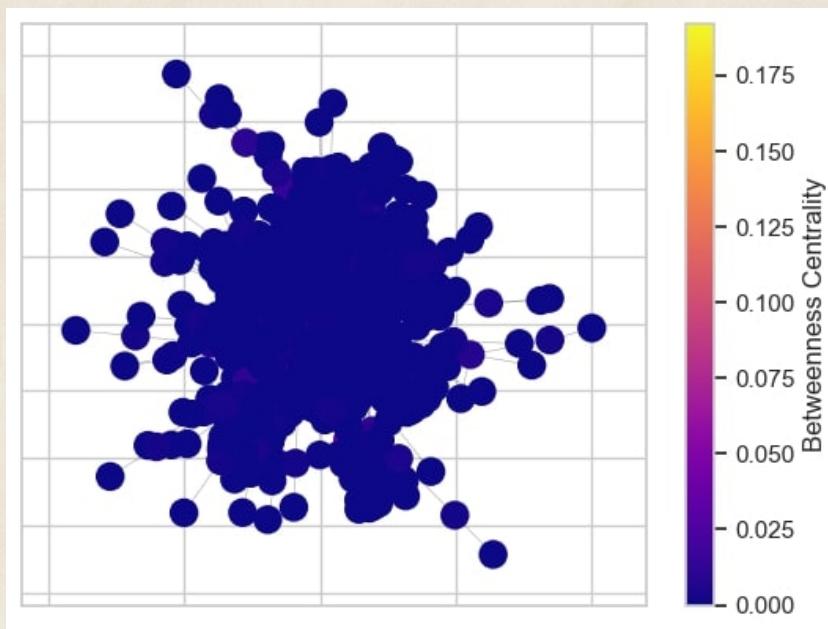


Figura 4.6: Heatmap Between Centrality

La formula matematica per calcolare l'Eigenvector Centrality si basa sull'autovettore della matrice di adiacenza del grafo. In pratica, la centralità di un nodo è determinata dalla somma delle centralità dei nodi con cui è connesso, ponderata dalla forza delle connessioni. Un nodo che è collegato a molti nodi centrali avrà una centralità elevata, mentre un nodo connesso a nodi periferici avrà una centralità bassa, anche se ha molte connessioni dirette.

Per calcolare la Eigenvector Centrality dei nodi, è stata utilizzata la funzione `eigenvector_centrality()` della libreria NetworkX, che fornisce una valutazione di questa misura per ogni nodo nella rete.

Nel contesto dei personaggi di Game of Thrones, la Eigenvector Centrality aiuta a identificare quei personaggi che non solo sono ben connessi, ma sono connessi con altri personaggi centrali e influenti.

I risultati ottenuti dalla tabella in Figura 4.7 mostrano dieci nodi con il valore di eigenvector centrality più alto e di seguito in basso un'altra tabella con i dieci nodi con il valore di eigenvector centrality più basso. Questo indica che nel caso di *Tyrion Lannister*, il suo valore elevato di Eigenvector Centrality può essere spiegato dal fatto che, sebbene non sia sempre il personaggio con il maggior numero di connessioni dirette, è frequentemente connesso a nodi di grande importanza, come le figure centrali della corte di King's Landing, la sua famiglia e le altre casate nobili. La sua abilità politica e il suo ruolo di stratega e consigliere lo mettono in una posizione privilegiata per influenzare eventi e alleanze, conferendogli una centralità maggiore rispetto ad altri personaggi. D'altra parte, *Manfrey Martell*, un cavaliere della Casa Martell, ha un valore di Eigenvector Centrality basso. Questo punteggio riflette la sua posizione



The figure consists of two vertically stacked tables from a data visualization tool like Jupyter Notebook. Both tables have a header row with columns for 'Personaggio' and 'Eigenvector Centrality'. The first table has 10 rows, and the second has 9 rows.

	Personaggio	Eigenvector Centrality
0	Tyrion-Lannister	0.251558
1	Cersei-Lannister	0.235771
2	Jaime-Lannister	0.226339
3	Joffrey-Baratheon	0.214376
4	Sansa-Stark	0.205842
5	Robert-Baratheon	0.194375
6	Eddard-Stark	0.191660
7	Stannis-Baratheon	0.182085
8	Catelyn-Stark	0.175886
9	Robb-Stark	0.173196

	Personaggio	Eigenvector Centrality
0	Manfrey-Martell	1.506822e-05
1	Steffon-Varner	1.479516e-05
2	Simon-Toyne	6.597588e-06
3	Walgrave	6.125952e-06
4	Quill	6.118404e-06
5	Rosey	6.118404e-06
6	Hugh-Hungerford	3.895174e-06
7	Murch	2.041418e-06
8	Torwold-Browntooth	1.278353e-06
9	Gormon-Tyrell	2.148365e-07

Figura 4.7: Tabella eigenvector centrality

relativamente periferica nella rete narrativa, nonostante il suo ruolo come castellano e membro della Casa Martell.

Anche in questo caso vengono mostrati i risultati con il grafico displot (Figura 4.8). L'istogramma evidenzia una distribuzione fortemente asimmetrica, con la maggior parte dei nodi concentrati vicino allo zero e una coda ridotta di nodi con valori di centralità più elevati. In termini di rete, ciò implica che la maggior parte dei personaggi è connessa in modo relativamente “debole” (o a nodi a loro volta poco influenti), mentre un ristretto gruppo di individui riveste un ruolo particolarmente rilevante: non solo possiedono molte connessioni, ma queste ultime sono verso altri nodi anch’essi centrali.

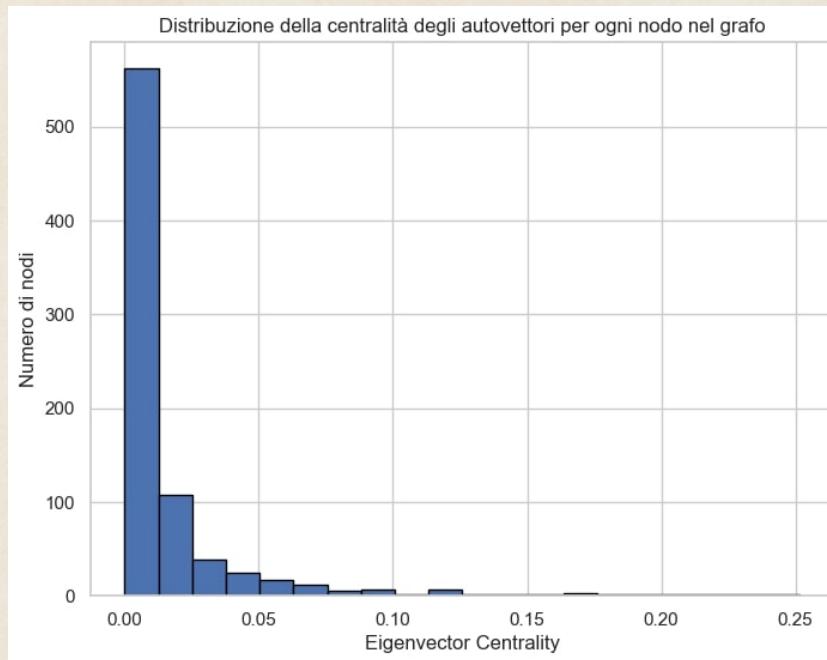


Figura 4.8: Distribuzione della Eigenvector Centrality

In aggiunta, è stata creata una heatmap della Eigenvector Centrality (Figura 4.9), che visualizza l'intensità della centralità di ogni nodo, dove i colori più scuri indicano una maggiore centralità e i colori più chiari denotano nodi meno centrali. Si può notare che ogni nodo è colorato in base al proprio valore di Eigenvector Centrality, che varia approssimativamente tra 0.05 e 0.25 (dal viola al giallo).

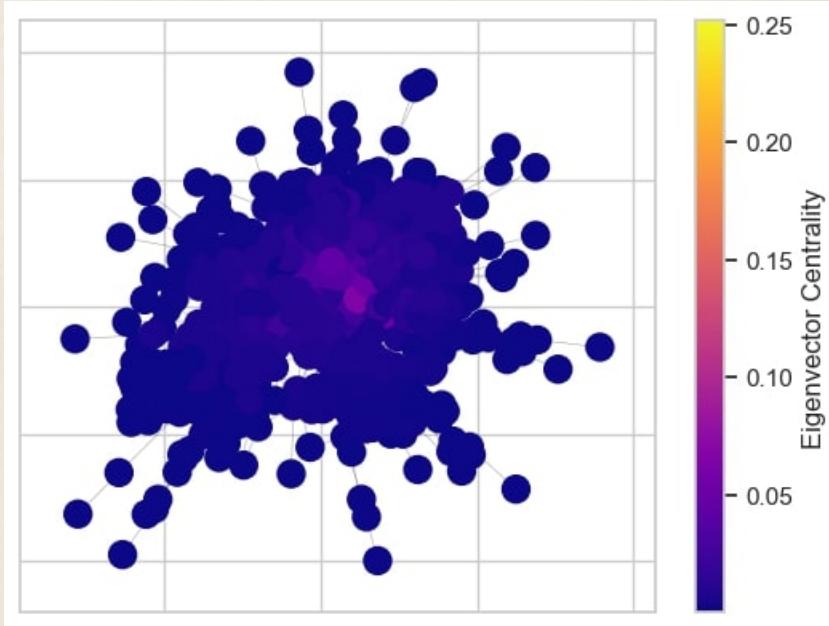


Figura 4.9: Heatmap Eigenvector Centrality

4.5 DEGREE CENTRALITY

Infine l'ultima tipologia di centralità analizzata è stata la **Degree Centrality**, la quale è una misura che calcola il numero di connessioni dirette che un nodo ha all'interno di una rete. Questo valore riflette l'importanza di un nodo in base alla sua capacità di connettersi direttamente con altri nodi. Tuttavia, va sottolineato che un nodo con una degree centrality alta non è necessariamente il più influente della rete, poiché potrebbe essere connesso a nodi che non sono centrali o che non svolgono un ruolo cruciale nella struttura complessiva della rete.

Per calcolare la Degree Centrality, è stata utilizzata la funzione `degree_centrality()` della libreria NetworkX, che permette di determinare il numero di connessioni dirette di ciascun nodo nel grafo.

I risultati ottenuti con la Figura 4.10 mostrano i dieci nodi con il valore di degree centrality più alto e, di seguito, in basso la tabella con i dieci nodi con il valore di degree centrality più basso.

Come già spiegato precedentemente, si è capito che *Tyrion Lannister*, uno dei personaggi principali della saga, è al centro di molte dinamiche politiche, alleanze e conflitti.

	Personaggio	Degree Centrality
0	Tyrion-Lannister	0.15345
1	Jon-Snow	0.14339
2	Jaime-Lannister	0.12704
3	Cersei-Lannister	0.12201
4	Stannis-Baratheon	0.11195
5	Arya-Stark	0.10566
6	Catelyn-Stark	0.09434
7	Sansa-Stark	0.09434
8	Eddard-Stark	0.09308
9	Robb-Stark	0.09308

10 rows	v	10 rows x 2 cols
	Personaggio	Degree Centrality
0	Lyanna-Mormont	0.001258
1	Malaquo-Maegyr	0.001258
2	Morgo	0.001258
3	Old-Bill-Bone	0.001258
4	Roose-Ryswell	0.001258
5	Shrouded-Lord	0.001258
6	William-Foxglove	0.001258
7	Willow-Witch-eye	0.001258
8	Thistle	0.001258
9	Wulfe	0.001258

Figura 4.10: Tabella degree centrality

D’altro canto, *Lyanna Mormont* è un personaggio marginale nella trama. Pur essendo una figura rispettata nella Casa Mormont, la sua presenza nel contesto narrativo è più limitata rispetto a quella di Tyrion. Lyanna non ha un numero significativo di connessioni con altri personaggi principali, poiché la sua storia è confinata principalmente all’interno della sua casata e alle dinamiche legate al Nord.

Per visualizzare la distribuzione dei valori di Degree Centrality, è stato utilizzato un displot (Figura 4.11) che consente di osservare come i valori siano distribuiti all’interno della rete. La coda della distribuzione, che si estende fino a circa 0.15 – 0.16, indica che non esistono nodi connessi a tutti o quasi tutti gli altri, ma esiste comunque un divario significativo tra i personaggi più centrali e quelli periferici. Ciò suggerisce una struttura a “stella multipla” o “scale-free”, in cui pochi nodi mantengono la coesione principale del grafo, mentre il resto si colloca su livelli di centralità più modesti.

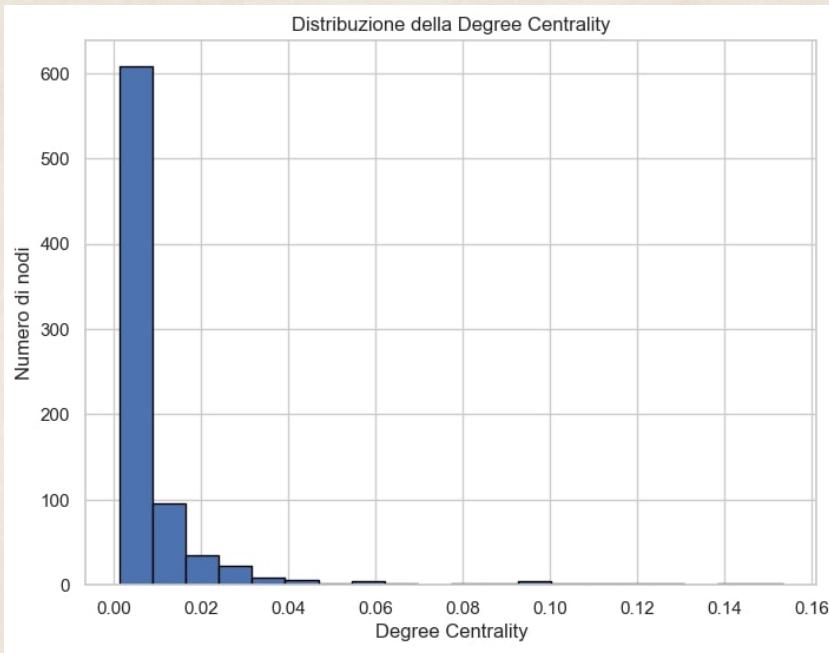


Figura 4.11: Distribuzione Degree Centrality

Infine, è stata creata una heatmap della Degree Centrality (Figura 4.12) che evidenzia i nodi con il numero più alto di connessioni. I nodi con tonalità più chiare (verso il giallo) sono gli hub del network, ovvero personaggi che interagiscono con molti altri. La maggior parte dei nodi si presenta in tonalità di viola più scure, indicando un numero di connessioni piuttosto limitato. L'assenza di valori estremamente elevati (oltre 0.14) suggerisce che, pur esistendo hub rilevanti, nessun nodo domina la rete in maniera assoluta, mantenendo comunque un certo equilibrio nelle interazioni.

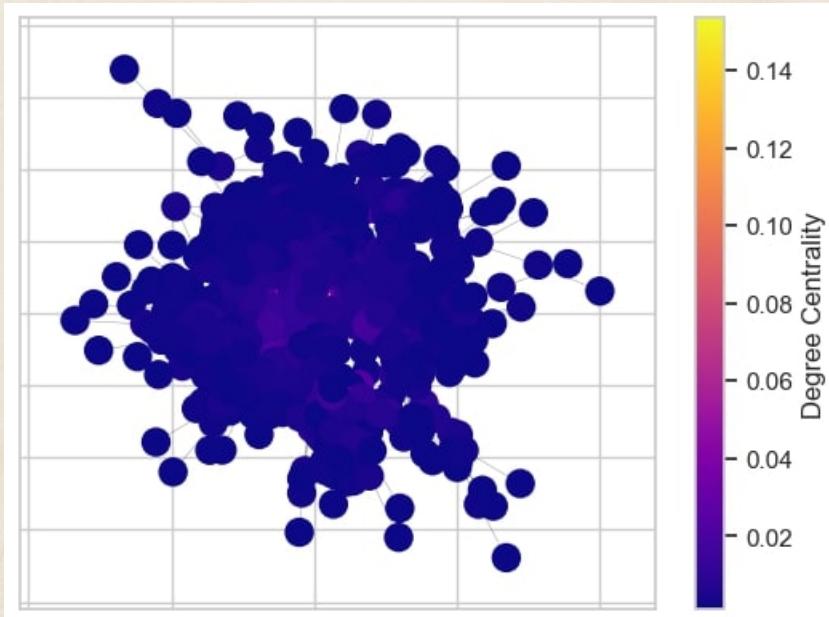


Figura 4.12: Heatmap Degree Centrality

ANALISI DELLE STRUTTURE

Mentre la centralità ci fornisce un'indicazione del grado di influenza di un nodo, l'analisi delle strutture consente di esplorare la configurazione delle connessioni tra i nodi stessi. Questo capitolo si concentrerà su alcune delle strutture fondamentali che emergono all'interno di una rete sociale: le **triadi**, le **clique**, i **k-core** e gli **ego graph**.

5.1 TRIADE

Partendo dalla **Triade**, essa rappresenta un sottoinsieme di tre nodi e delle loro connessioni reciproche, e rappresenta le interazioni dirette tra tre personaggi della rete. Nel grafo dei personaggi di Game of Thrones, le triadi spesso riflettono alleanze, conflitti o dinamiche familiari. Le triadi hanno permesso di identificare le interazioni più forti e più deboli all'interno della rete, e il loro numero totale è un indicatore della densità delle relazioni tra i personaggi.

Per analizzare le triadi all'interno di un contesto di Social Network Analysis (SNA), ha più senso considerare l'Eigenvector Centrality. Per cui si è scelto di concentrarsi sulle triadi formate da *Tyrion Lannister* e *Cersei Lannister*, in quanto questi due personaggi risultano avere il maggior numero di interazioni all'interno della rete. La centralità delle loro posizioni nel grafo, dovuta alle numerose connessioni con altri personaggi, li rende candidati ideali per analizzare le dinamiche di interazione a livello di triade. Le triadi che sono state trovate sono ben 58 (Figura 5.1).

Tabella delle Triadi:			
	Personaggio 1	Personaggio 2	Nodo Comune
0	Cersei-Lannister	Tyrion-Lannister	Brienne-of-Tarth
1	Cersei-Lannister	Tyrion-Lannister	Bronn
2	Cersei-Lannister	Tyrion-Lannister	Boros-Blount
3	Cersei-Lannister	Tyrion-Lannister	Robb-Stark
4	Cersei-Lannister	Tyrion-Lannister	Willas-Tyrell
5	Cersei-Lannister	Tyrion-Lannister	Bran-Stark
6	Cersei-Lannister	Tyrion-Lannister	Benjen-Stark
7	Cersei-Lannister	Tyrion-Lannister	Gregor-Clegane
8	Cersei-Lannister	Tyrion-Lannister	Gyles-Rosby

Figura 5.1: Triadi Tyrion Lannister e Cersei Lannister

L'analisi di queste triadi e dei loro legami evidenzia la non linearità delle interazioni in una rete sociale complessa, dove i conflitti non sono solo il risultato di legami diretti, ma anche delle strutture e relazioni indirette che collegano i nodi.

5.2 CLIQUE

In seguito all'analisi delle triadi, si è proceduto con l'analisi delle **Clique** utilizzando gli stessi personaggi principali, *Tyrion Lannister* e *Cersei Lannister*, come focus dell'indagine.

Le clique sono state selezionate in base a criteri specifici per analizzare le dinamiche sociali e politiche all'interno del grafo.

Sono state mostrate tutte le Clique presenti nel dataset in relazione con i due personaggi, ne sono state trovate 111.

Successivamente, sono state identificate le clique con il numero massimo di nodi (Figura 5.2), le quali rappresentano il gruppo più grande di personaggi che sono strettamente connessi tra loro. Nella ricerca è stata trovata solo una clique massima contenente 14 nodi.

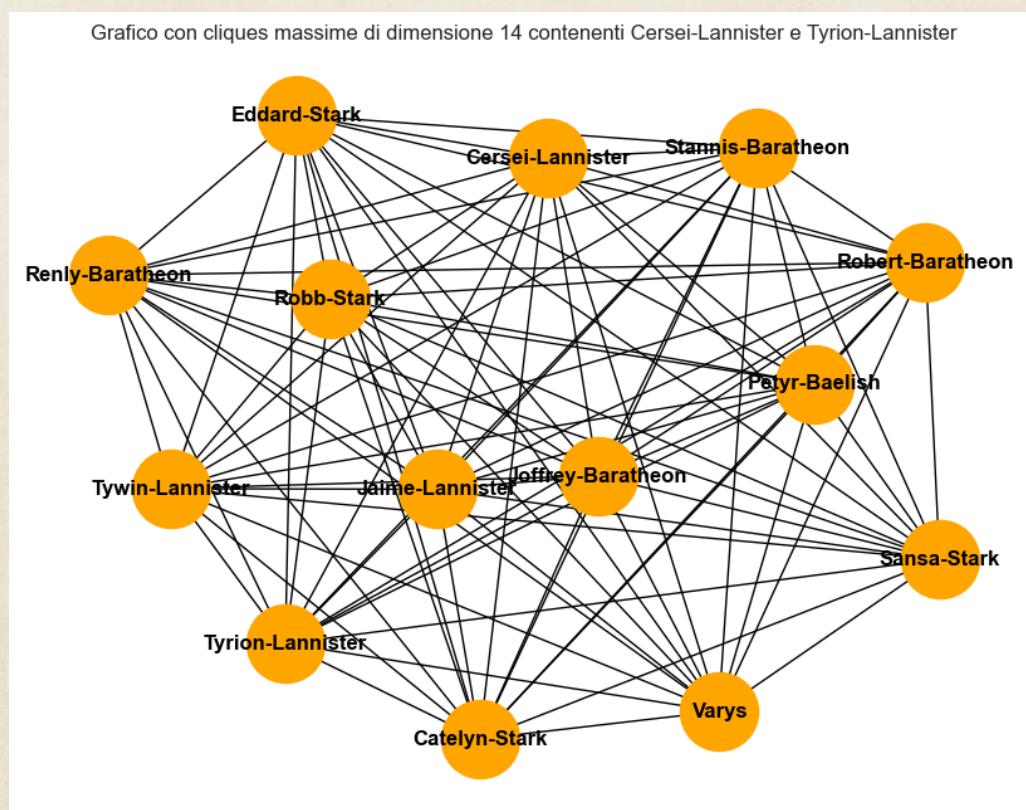


Figura 5.2: Clique massimale

La Figura 5.2 mostra un gruppo di personaggi, che parte da *Tyrion Lannister* e *Cersei Lannister*, legati tra loro da rapporti estremamente densi. La ragione di queste relazioni

così strette sta nella convergenza di interessi politici, familiari e di potere. Ad esempio, l'alleanza fra diversi Lannister, i Baratheon di Approdo del Re, alcuni Stark e altre casate nobili avviene perché, nella storia, i loro destini sono intrecciati da conflitti, matrimoni e alleanze strategiche.

5.3 K-CORE

Il **K-core** è un concetto usato nell'analisi delle reti per identificare un gruppo di nodi che sono densamente connessi tra di loro. La definizione di K-core può essere vista come una generalizzazione del concetto di clique. Mentre una clique è un sottoinsieme di nodi dove ogni nodo è connesso con ogni altro nodo, un K-core è più flessibile. Un nodo può far parte di un K-core se è connesso a un numero sufficiente di altri nodi, senza che sia necessario essere connesso con ogni singolo nodo del gruppo.

Il valore di k scelto è 13, il che significa che il K-core risultante include solo i nodi (personaggi) che sono connessi ad almeno 13 altri nodi all'interno del grafo (Figura 5.3). Questo gruppo evidenzia i personaggi più centrali e interconnessi nella trama, poiché hanno un numero sufficiente di legami con altri personaggi principali.

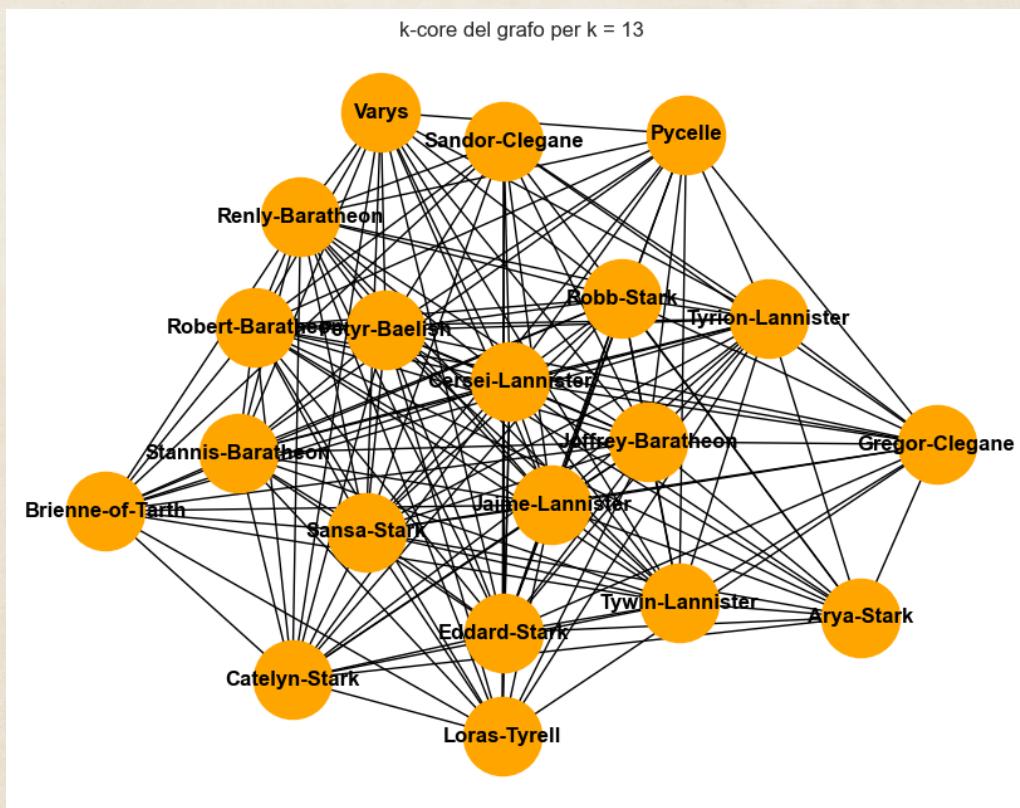


Figura 5.3: K-core

I rapporti con figure come Cersei, Jaime e Tywin Lannister derivano dal legame familiare e dalle trame di potere interne alla casata Lannister; le relazioni con altri

personaggi (ad esempio Varys, Petyr Baelish e i membri di diverse casate come Stark o Baratheon) riflettono l'intricato gioco politico e le alleanze variabili nella storia.

Di seguito vengono mostrati gli altri nodi. Nodi nel k-core con $k = 13$:

```
['Robert-Baratheon', 'Tyrion-Lannister', 'Sandor-Clegane',
'Catelyn-Stark', 'Sansa-Stark', 'Gregor-Clegane',
'Renly-Baratheon', 'Tywin-Lannister', 'Loras-Tyrell',
'Jaime-Lannister', 'Brienne-of-Tarth', 'Robb-Stark',
'Eddard-Stark', 'Cersei-Lannister', 'Pycelle', 'Joffrey-Baratheon',
'Arya-Stark', 'Petyr-Baelish', 'Varys', 'Stannis-Baratheon']
```

5.4 EGO GRAPH

Un **Ego-graph** è una rappresentazione di una rete che mostra un nodo centrale (chiamato ego) e tutti i nodi che sono direttamente connessi a esso (chiamati alter o vicini).

Nel caso specifico dell'analisi, è stato utilizzato un `ego_graph()` per esaminare le connessioni dirette di *Tyrion Lannister*. In questo ego-graph (Figura 5.4) *Tyrion Lannister* è stato considerato come il nodo centrale (l'ego), e sono state analizzate le interazioni dirette con tutti gli altri personaggi a lui collegati (gli alter). Questo ha permesso di visualizzare come *Tyrion Lannister* si relaziona con gli altri personaggi, senza considerare le connessioni tra i personaggi vicini a lui, ma concentrandosi esclusivamente sulle sue interazioni più immediate.

L'ego graph di *Tyrion Lannister* mostra tutti i personaggi a cui egli è direttamente connesso, evidenziando la rete locale attorno a lui e i legami fra questi nodi. La grande quantità di connessioni (e la loro ampia diffusione) conferma il ruolo di Tyrion come uno dei personaggi più “sociali” e politicamente attivi: egli è in contatto con membri di quasi ogni casata e con figure di rilievo come Varys (per scambi di informazioni e trame di corte) o Sansa Stark (tramite il matrimonio e la sua permanenza ad Approdo del Re). Le relazioni includono anche soldati, consiglieri e nobili, derivanti dal suo ruolo di Primo Cavaliere o di diplomatico in diverse fasi della storia. L'insieme dei legami che si irradiano da Tyrion rappresenta quindi la complessità della sua posizione: spesso si trova a mediare tra fazioni diverse e a costruire alleanze, spiegando la densità del suo ego graph.

Ego Graph di Tyrion-Lannister

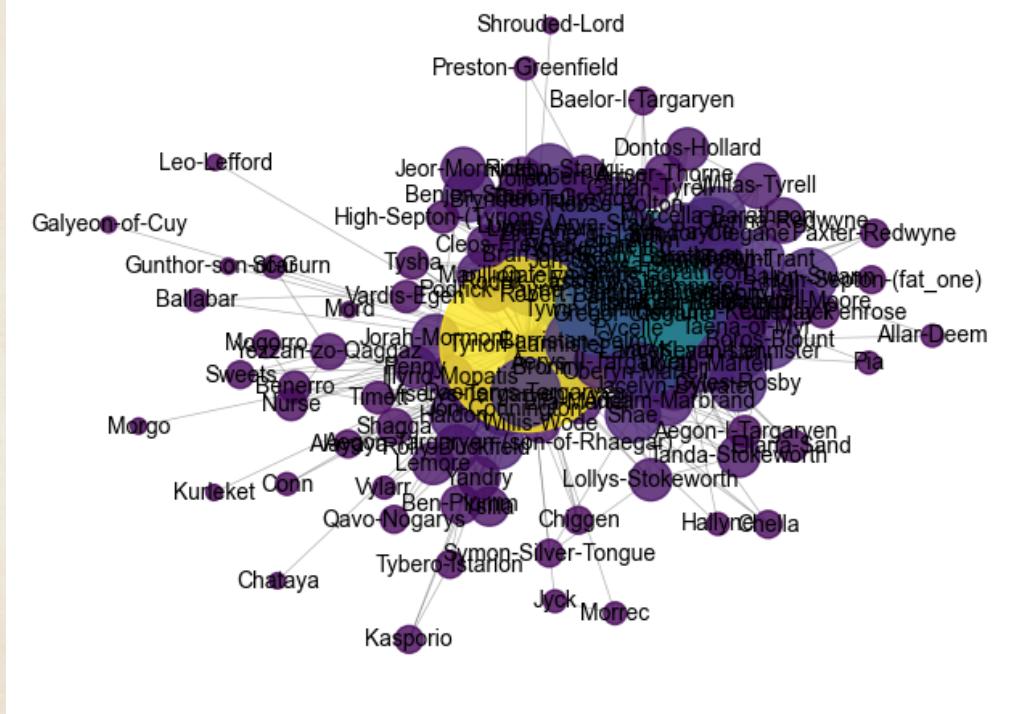


Figura 5.4: Ego graph