

UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Dipartimento di Ingegneria dell'Informazione
Corso di Laurea in Ingegneria Informatica e dell'Automazione



CORSO DI DATA SCIENCE

**Analisi delle strutture relazionali nell'Universo Marvel tramite
Social Network Analysis**



Autori

Mattia Sbattella
Simone Di Battista
Niccolò Ciotti
Luca Renzi

ANNO ACCADEMICO 2024-2025

Indice

1	Introduzione	1
1.1	Social Network Analysis	1
1.2	Obiettivo e ambiti applicativi della SNA	1
1.3	NetworkX	2
2	Dataset	3
2.1	Struttura del dataset	3
2.2	Preprocessing	4
2.3	Analisi esplorativa del grafo	5
3	Analisi delle centralità nella rete	10
3.1	La centralità nelle reti	10
3.2	Degree Centrality	10
3.3	Betweenness Centrality	14
3.4	Closeness Centrality	17
3.5	Eigenvector Centrality	20
4	Analisi delle strutture	24
4.1	Triadi	24
4.2	Clique	25
4.3	K-core	27
4.4	Ego network	29
4.5	Communities	32
4.5.1	Risultati con algoritmo Louvain	33
4.5.2	Risultati con metodo Greedy Modularity	36

Elenco delle figure

2.1	Rappresentazione del grafo dei personaggi Marvel	5
2.2	Analisi del grafo dei personaggi Marvel	6
2.3	Rappresentazione del grafo dei personaggi Marvel con Kamada-Kawai Layout	7
2.4	Rappresentazione del grafo dei personaggi Marvel con Spring Layout .	8
2.5	Distribuzione degli archi per categoria di interazione	9
3.1	I 10 nodi con degree centrality più alta e più bassa	12
3.2	Distribuzione della degree centrality nella rete	13
3.3	Rappresentazione della degree centrality nella rete	14
3.4	I 10 nodi con betweenness centrality più alta e più bassa	15
3.5	Distribuzione della betweenness centrality nella rete	16
3.6	Rappresentazione della betweenness centrality nella rete	17
3.7	I 10 nodi con closeness centrality più alta e più bassa	18
3.8	Distribuzione della closeness centrality nella rete	19
3.9	Rappresentazione della closeness centrality nella rete	20
3.10	I 10 nodi con eigenvector centrality più alta e più bassa	21
3.11	Distribuzione dell'eigenvector centrality nella rete	22
3.12	Rappresentazione dell'eigenvector centrality nella rete	23
4.1	Visualizzazione delle triadi più pesanti nella rete	25
4.2	Visualizzazione delle clique nella rete	26
4.3	Visualizzazione della clique massima	26
4.4	Rappresentazione grafica della clique massima	27
4.5	Visualizzazione del k-core di grado massimo	28
4.6	Rappresentazione grafica del k-core, con k=56	29
4.7	Rappresentazione grafica della ego-network di Captain America	31
4.8	Grafo della ego-network dei 10 vicini con centralità maggiore	32

4.9 Comunità rilevate tramite algoritmo di Louvain	33
4.10 Rappresentazione grafica della comunità X-Men	34
4.11 Rappresentazione grafica della comunità Avengers	34
4.12 Tabella con i primi 5 personggi per centralità di ogni comunità	35
4.13 Rappresentazione grafica della comunità X-Men, con i top 25 nodi per centralità	36
4.14 Rappresentazione grafica della comunità Avengers, con i top 25 nodi per centralità	36
4.15 Comunità rilevate tramite metodo Greedy Modularity	37
4.16 Rappresentazione grafica della comunità Avengers	37
4.17 Rappresentazione grafica della comunità X-Men	38
4.18 Rappresentazione grafica della comunità Avengers, con i top 25 nodi per centralità	38
4.19 Rappresentazione grafica della comunità X-Men, con i top 25 nodi per centralità	39
4.20 Tabella con i primi 5 personggi per centralità di ogni comunità	39
4.21	40

CAPITOLO 1

Introduzione

1.1 Social Network Analysis

La *Social Network Analysis (SNA)* è l'insieme di metodi e teorie che studiano le relazioni e le interazioni tra entità sociali. A differenza degli approcci che si concentrano sugli attributi individuali degli attori, la SNA pone l'attenzione sulla struttura dei legami come elemento determinante per comprendere il comportamento e gli esiti del sistema.

Una rete sociale è formalmente rappresentata utilizzando la *Teoria dei Grafi*. Un grafo è una coppia $G = (V, E)$, dove:

- i Nodi (o Vertici, V) rappresentano gli attori sociali all'interno della rete. Possono essere individui, organizzazioni, paesi, siti web, o qualsiasi entità in grado di interagire.
- gli Archi (o Spigoli, E) rappresentano le relazioni o le connessioni tra i nodi. La natura degli archi può variare notevolmente:
 - Direzionali: La relazione va in un'unica direzione (es. A segue B su una piattaforma social).
 - Non Direzionali: La relazione è reciproca (es. A è amico di B).
 - Pesati: Viene assegnato un valore (peso) per indicare l'intensità, la frequenza o la forza del legame.

1.2 Obiettivo e ambiti applicativi della SNA

La Social Network Analysis ha l'obiettivo di esplorare e comprendere la struttura delle relazioni all'interno di una rete. Attraverso questo approccio è possibile indi-

viduare quali nodi rivestono un ruolo centrale, quali gruppi o comunità si formano naturalmente e in che modo le connessioni influenzano il comportamento complessivo del sistema. L'analisi permette inoltre di osservare come l'informazione o l'influenza si propagano, evidenziando punti di forza e di vulnerabilità della rete. In questo modo la SNA offre una visione più profonda delle dinamiche che caratterizzano un insieme complesso di interazioni.

Questo tipo di analisi è trasversale e viene applicata in svariati domini:

- *Sociologia*: Studio della diffusione di informazioni, innovazioni o malattie.
- *Economia e Management*: Analisi delle strutture di collaborazione aziendale e dei flussi di conoscenza.
- *Scienze Politiche*: Mappatura delle alleanze internazionali e delle reti di influenza.
- *Informatica*: Analisi dei collegamenti ipertestuali (Web Graph) o delle interazioni su piattaforme social.

1.3 NetworkX

NetworkX è una libreria Python ampiamente utilizzata per l'analisi e la visualizzazione di grafi e reti complesse. Consente di rappresentare sistemi relazionali sotto forma di nodi e archi e mette a disposizione numerosi strumenti per studiare proprietà strutturali, centralità, percorsi, comunità e molte altre caratteristiche tipiche della Social Network Analysis. La libreria è particolarmente utilizzata per la sua semplicità, la flessibilità nella gestione dei dati e la ricchezza di funzioni già pronte, che permettono di costruire e analizzare reti anche molto grandi con poche righe di codice. Grazie a NetworkX diventa quindi possibile esplorare in modo intuitivo le dinamiche interne a una rete e generare visualizzazioni utili per interpretarne la struttura globale e i rapporti tra i singoli elementi.

CAPITOLO 2

Dataset

2.1 Struttura del dataset

Per studiare le relazioni tra i personaggi dell'Universo Marvel è stato scelto un dataset che modella una rete di co-appearance (o co-occurrence network) dei suoi personaggi principali.

Il dataset è organizzato in due file CSV distinti:

- *marvel-unimodal-nodes.csv*: rappresenta l'elenco completo dei personaggi Marvel considerati contenente una tabella costituita da due attributi, `Id` e `Label`.

Id	Label
Black Panther / T'challa	Black Panther / T'challa
Loki [asgardian]	Loki [asgardian]

Tabella 2.1: Esempio della tabella nel file *marvel-unimodal-nodes.csv*

- *marvel-unimodal-edges.csv*: contiene le relazioni tra i personaggi, includendo un peso associato, tramite l'attributo `Weight`, a ciascun legame e indicando, tramite le colonne `Source` e `Target`, i due personaggi coinvolti in ogni arco.

Source	Target	Weight
Black Panther / T'challa	Loki [asgardian]	10
Captain America	Black Panther / T'challa	131

Tabella 2.2: Esempio della tabella nel file *marvel-unimodal-edges.csv*

L’elemento centrale di questo dataset è il concetto di co-apparizione, rappresentato dal peso dell’arco. In particolare, viene creato un arco tra due personaggi quando essi compaiono insieme almeno una volta nello stesso fumetto o film. Poiché la relazione è simmetrica (se A appare con B, automaticamente B appare con A), la rete è modelata come un grafo non orientato. Questo permette di analizzare in modo immediato l’intensità e la struttura delle interazioni tra i vari protagonisti dell’universo Marvel.

2.2 Preprocessing

Per analizzare le relazioni tra i personaggi, abbiamo voluto classificare gli archi della rete in base alla loro “forza”, rappresentata dal peso associato a ciascun legame. In questo contesto, il peso di un arco non è un semplice numero: esso riflette quanto frequentemente due personaggi compaiono insieme. In altre parole, un peso elevato indica una co-occorrenza frequente, che può essere interpretata come un’interazione più forte tra i due personaggi. Viceversa, un peso basso suggerisce una presenza congiunta meno frequente, quindi un’interazione più debole.

Per rendere l’analisi più interpretabile, i valori di peso sono stati suddivisi in quattro gruppi basati sui quartili della distribuzione. In questo modo, ogni categoria contiene un numero simile di archi, garantendo una classificazione equilibrata. A ciascun quartile è stata assegnata un’etichetta descrittiva, da *Interazione Debole* a *Interazione Forte*, che permette di comprendere rapidamente il livello di interazione tra i personaggi.

Inoltre, sono stati calcolati i range di peso effettivi corrispondenti a ciascuna categoria, rendendo evidente quali valori numerici rappresentano interazioni deboli, medie o forti.

Le tabelle riportano i risultati concreti di questa discretizzazione. La prima mostra i punti di taglio utilizzati, ossia i valori dei quartili della distribuzione dei pesi degli archi:

Quantile	Peso (Weight)
Q0 (Min)	5.0
Q1 (25%)	7.0
Q2 (50%)	11.0
Q3 (75%)	21.0
Q4 (Max)	744.0

Tabella 2.3: Punti di taglio per la discretizzazione dei pesi degli archi

Questi valori rappresentano i confini tra le categorie di interazione, garantendo che ciascun gruppo contenga un numero simile di archi.

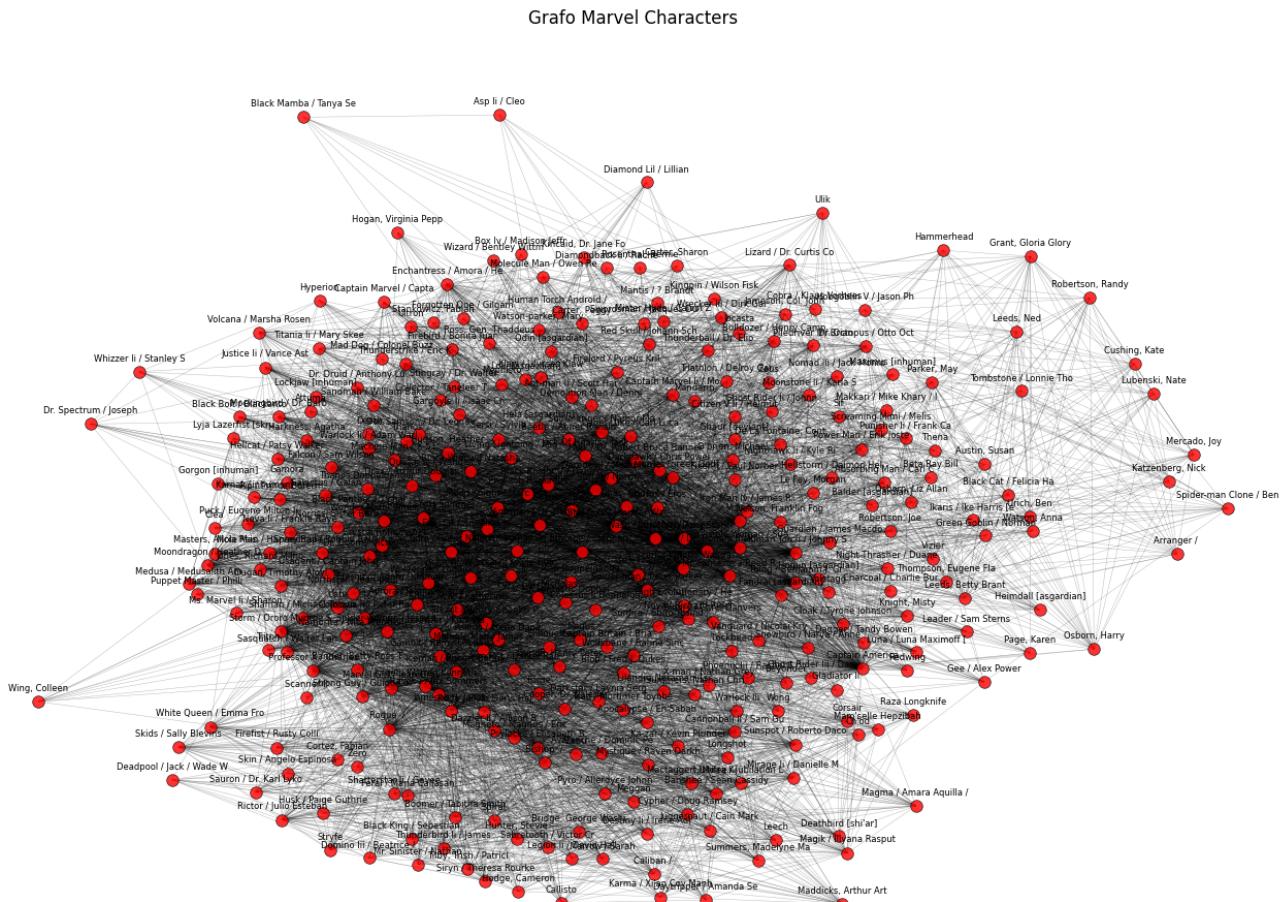
La seconda tabella mostra il numero di archi assegnati a ciascuna categoria:

Categoria di Interazione	Numero di Archi
Interazione Debole	3.212
Interazione Media-Bassa	2.154
Interazione Media-Alta	2.164
Interazione Forte	2.361

Tabella 2.4: Conteggio degli archi per categoria di interazione

2.3 Analisi esplorativa del grafo

Per avere una prima impressione visiva della rete dei personaggi Marvel, nella Figura 2.1 viene mostrata la rappresentazione grafica della rete.

**Figura 2.1:** Rappresentazione del grafo dei personaggi Marvel

Successivamente è stata effettuata un'analisi esplorativa del grafo, rappresentata nella Figura 2.2.

Numero Totale di Nodi ($ V $)	327
Numero Totale di Archi ($ E $)	9891
Componenti Connesse Totali	1
Grafo Connesso?	True
Densità della Rete	0.1856
Somma Totale dei Pesi degli Archi	216765
Densità Pesata (Custom, rispetto al Max Peso)	0.0055
Diametro (Componente più Grande)	3
Raggio (Componente più Grande)	2
Coefficiente di Clustering Medio (Componente più Grande)	0.7047

Figura 2.2: Analisi del grafo dei personaggi Marvel

Dall'analisi emerge che la rete è composta da 327 *nodi* e 9.891 *archi*. La rete è completamente connessa, come indicato da *Componenti Connesse Totali* = 1 e *Grafo Connesso?* = *True*, il che significa che ogni personaggio è raggiungibile da qualsiasi altro attraverso le relazioni presenti. La *densità* della rete è pari a 0.1856, un valore relativamente alto che suggerisce che molti personaggi sono direttamente collegati tra loro. La somma totale dei pesi degli archi ammonta a 216.765, indicando il numero complessivo di co-occorrenze, mentre la *densità pesata* (0.0055) fornisce una misura della connettività considerando l'intensità delle interazioni.

Per quanto riguarda le proprietà della componente principale, il *diametro* è 3, indicando che la distanza massima tra due nodi qualsiasi è molto breve, mentre il *raggio* pari a 2 mostra che alcuni nodi centrali sono vicini a gran parte degli altri. Infine, il *coefficiente di clustering medio* della componente principale è 0.7047, un valore elevato che evidenzia la tendenza dei personaggi a formare sottogruppi densamente connessi, ossia comunità locali di interazioni frequenti.

Dopo le operazioni di preprocessing, è stato ottenuto un grafo aggiornato in cui gli archi sono colorati in base alla forza dell'interazione tra i personaggi, permettendo di distinguere visivamente le relazioni deboli da quelle forti.

Per visualizzare la rete sono state prodotte due rappresentazioni principali: una basata sul *Kamada-Kawai layout*, e l'altra sul *Spring layout*. Il *Kamada-Kawai layout*, Figura 2.3, posiziona i nodi in modo da minimizzare le differenze tra le distanze geometriche e le distanze ideali lungo il grafo, risultando in una visualizzazione chiara delle strutture interne e dei cluster locali. Questo layout è particolarmente utile per evidenziare sottogruppi e comunità in grafi di dimensioni medio-piccole, ma può diventare computazionalmente pesante e creare sovrapposizioni quando il numero di nodi è elevato.

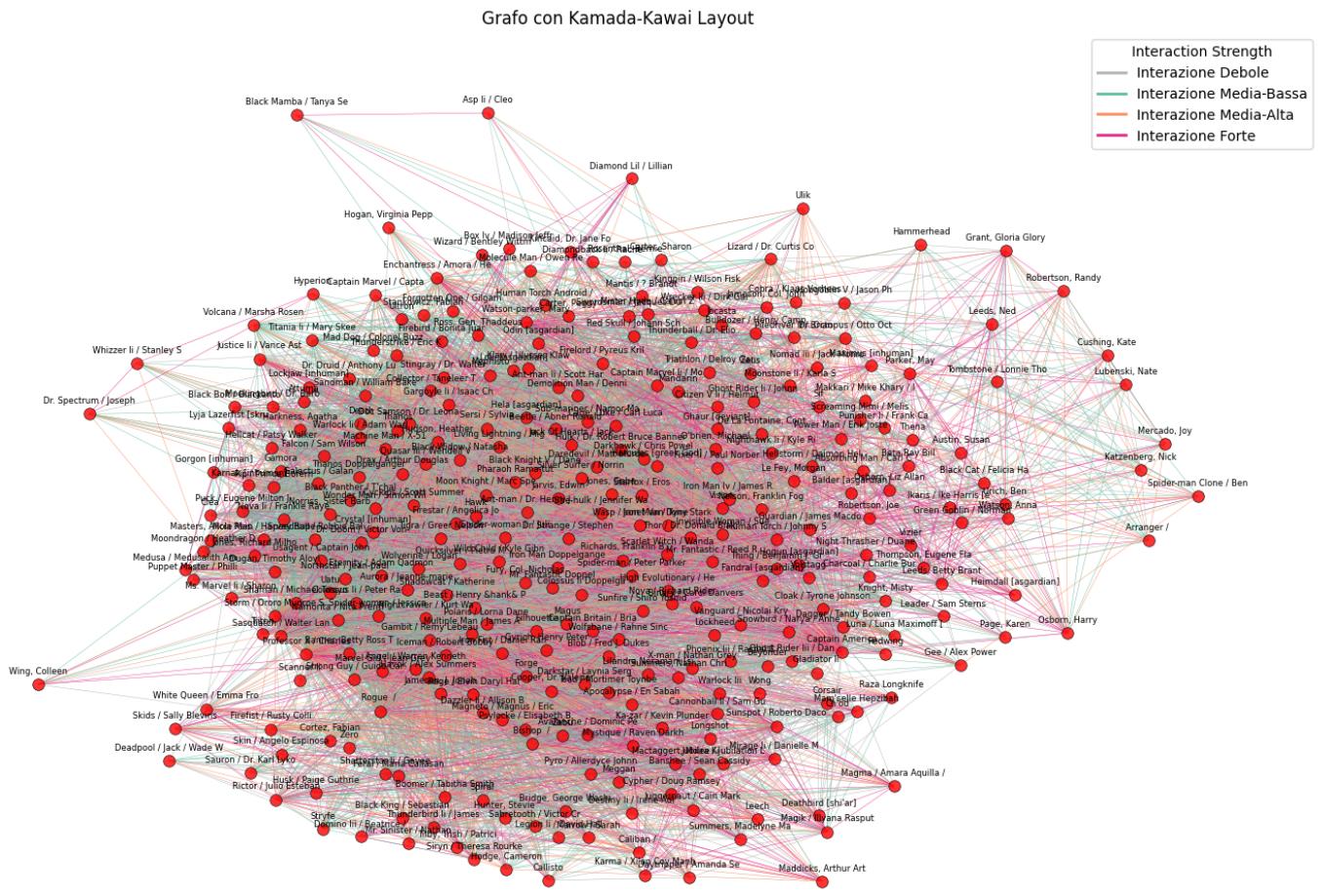


Figura 2.3: Rappresentazione del grafo dei personaggi Marvel con Kamada-Kawai Layout

Lo *Spring layout*, Figura 2.4, invece, simula una rete elastica in cui i nodi si respingono tra loro mentre gli archi agiscono come molle attrattive; questo layout tende a mostrare più chiaramente la distribuzione generale dei nodi nello spazio e la struttura complessiva della rete. Grazie a questa caratteristica, questo layout è più adatto per grafi grandi, in quanto distribuisce i nodi in modo uniforme e mantiene leggibile la rete anche con molte connessioni, pur sacrificando un po' di precisione nella visualizzazione dei cluster più densi.

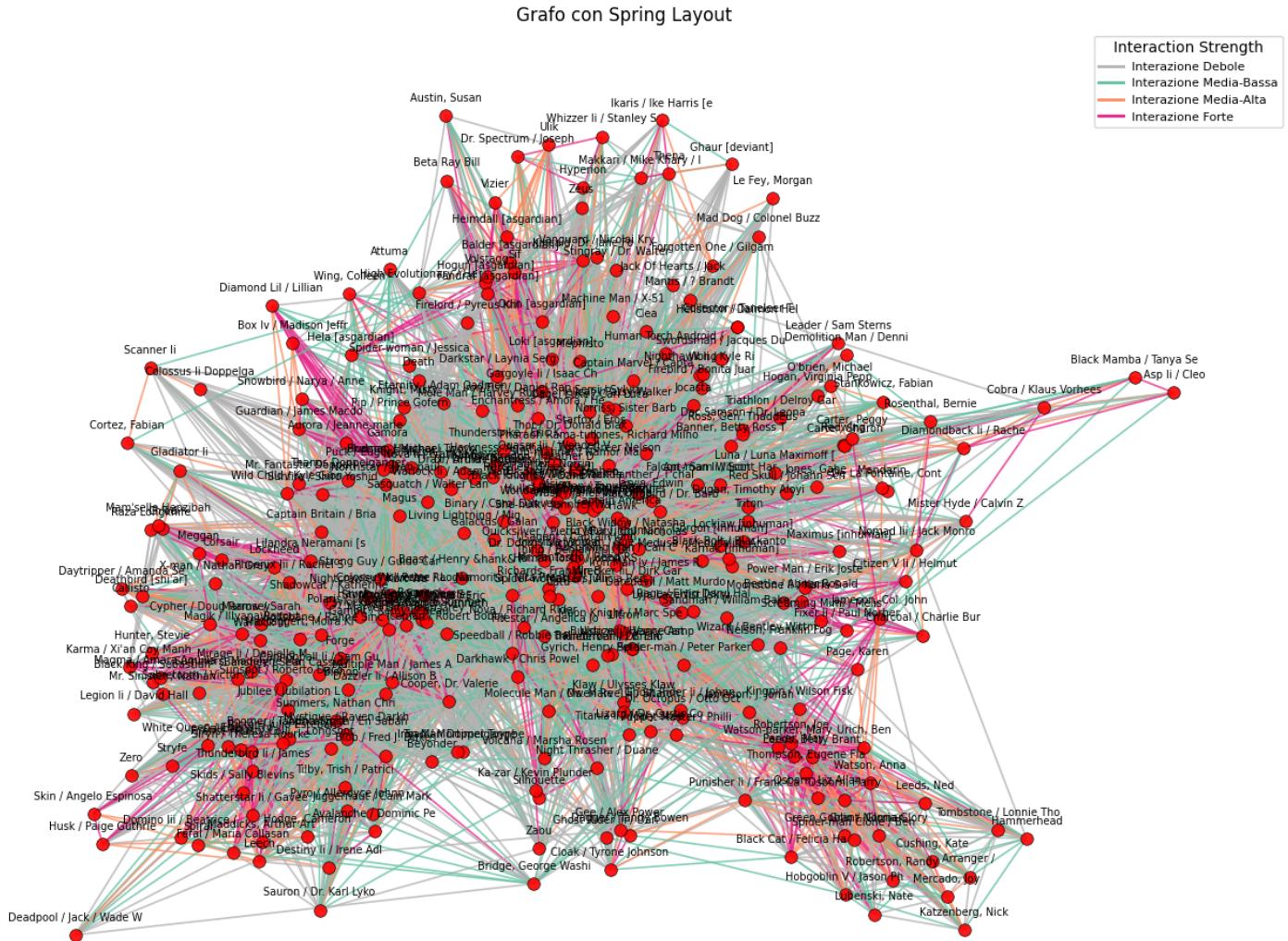


Figura 2.4: Rappresentazione del grafo dei personaggi Marvel con Spring Layout

Infine, per comprendere meglio la distribuzione delle interazioni all'interno della rete, è stato creato un istogramma che mostra la frequenza degli archi per categoria di interazione, come riportato nella Figura 2.5.

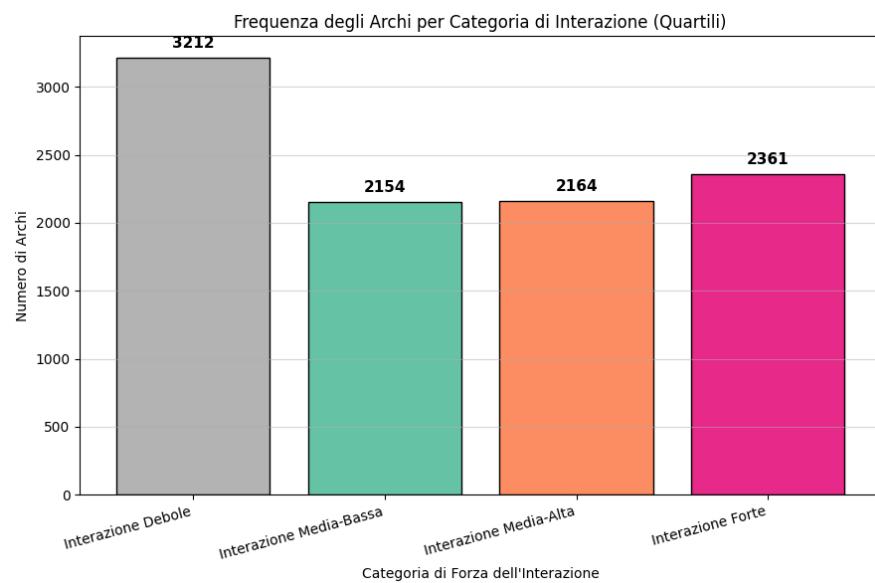


Figura 2.5: Distribuzione degli archi per categoria di interazione

CAPITOLO 3

Analisi delle centralità nella rete

3.1 La centralità nelle reti

La centralità rappresenta uno dei concetti fondamentali della Social Network Analysis e consente di valutare il ruolo e l'importanza dei nodi all'interno di una rete. Attraverso diverse misure, la centralità permette di identificare quali attori occupano posizioni strategiche, influenzano maggiormente il flusso di informazioni, fungono da intermediari o risultano particolarmente integrati nel sistema relazionale.

L'idea di fondo è che non tutti i nodi di una rete hanno la stessa rilevanza strutturale: alcuni sono più connessi, altri più vicini agli altri nodi, altri ancora svolgono funzioni di ponte tra comunità differenti. Le diverse misure di centralità sono state sviluppate proprio per cogliere queste sfumature e per descrivere in modo quantitativo il ruolo che ciascun nodo ricopre.

Nel contesto di una rete sociale — come quella dei personaggi Marvel analizzata — la centralità consente di individuare quali personaggi risultano più “centrali” nelle dinamiche narrative: quelli più ricorrenti, più interconnessi o che fungono da collegamento tra gruppi differenti.

3.2 Degree Centrality

La *degree centrality* è la misura di centralità più semplice e intuitiva all'interno della Social Network Analysis. Essa quantifica l'importanza di un nodo contando il numero di connessioni che intrattiene con gli altri nodi della rete. In altre parole, indica quanto un attore è direttamente coinvolto nelle relazioni che definiscono il sistema.

Nel caso della nostra rete Marvel non orientata, il *degree* di un nodo corrisponde al numero totale di archi che lo collegano agli altri personaggi. Una *degree centrality* elevata

suggerisce quindi che il nodo occupa una posizione altamente interattiva, poiché entra in contatto con molti altri attori della rete.

Dal punto di vista interpretativo, questa misura permette di individuare i personaggi più “popolari” o maggiormente presenti nelle interazioni. Nelle reti sociali, infatti, avere molte connessioni significa spesso disporre di maggiore visibilità e influenza locale.

Operativamente, la *degree centrality* è stata ottenuta attraverso la funzione, della libreria NetworkX, `nx.degree_centrality(G)`, che restituisce per ogni personaggio un valore normalizzato compreso tra 0 e 1. Questo valore rappresenta la proporzione dei personaggi della rete con cui ciascun nodo è direttamente collegato.

Una volta calcolate le misure di centralità, queste sono state integrate in un unico DataFrame contenente l’elenco di tutti i personaggi e la relativa quantità di connessioni. Successivamente, i personaggi sono stati ordinati in maniera decrescente per individuare i nodi più centrali — ovvero quelli che intrattengono il maggior numero di relazioni nella rete — e in maniera crescente per identificare quelli meno connessi.

La Figura 3.1 mostra il risultato.

Top 10 Personaggi per Grado di Centralità:		
Character	Degree_Centrality	
Captain America	0.791411	
Wolverine ↳ Logan	0.674847	
Spider-man ↳ Peter Parker	0.668712	
Thing ↳ Benjamin J. Gr	0.656442	
Vision	0.653374	
Beast ↳ Henry & Hank& P	0.628834	
Thor ↳ Dr. Donald Blak	0.628834	
Iron Man ↳ Tony Stark	0.622699	
Mr. Fantastic ↳ Reed R	0.607362	
Cyclops ↳ Scott Summer	0.604294	

Bottom 10 Personaggi per Grado di Centralità:		
Character	Degree_Centrality	
Asp II ↳ Cleo	0.024540	
Black Mamba ↳ Tanya Se	0.024540	
Wing, Colleen	0.036810	
Deadpool ↳ Jack ↳ Wade W	0.039877	
Skin ↳ Angelo Espinosa	0.042945	
Diamond Lil ↳ Lillian	0.042945	
Dr. Spectrum ↳ Joseph	0.046012	
Whizzer II ↳ Stanley S	0.046012	
Cobra ↳ Klaus Vorhees	0.049080	
Husk ↳ Paige Guthrie	0.049080	

Figura 3.1: I 10 nodi con degree centrality più alta e più bassa

La *degree centrality* evidenzia con chiarezza quali personaggi occupano le posizioni più connesse nella rete Marvel. Nella Top 10 emergono figure estremamente centrali come Captain America, Wolverine, Spider-Man, Iron Man e Thor. Il loro alto valore di centralità riflette un ruolo narrativo trasversale: sono personaggi presenti in molte serie, team e crossover, e interagiscono con una porzione molto ampia dell'universo Marvel.

La Bottom 10, invece, include personaggi marginali, con pochissime connessioni e ruoli narrativi limitati a contesti specifici. Valori di degree centrality molto bassi indicano una presenza sporadica nella rete e una scarsa integrazione nelle trame principali.

L'istogramma, in Figura 3.2, rivela una distribuzione fortemente asimmetrica e concentrata verso valori bassi. La maggior parte dei personaggi presenta infatti un grado di centralità compreso tra 0.05 e 0.15, indicando che la rete è composta in larga parte da nodi con un numero relativamente ridotto di connessioni.

All'aumentare della centralità, la frequenza diminuisce rapidamente: solo pochi

personaggi raggiungono valori superiori a 0.30, e un gruppo estremamente limitato supera la soglia di 0.50. La lunga coda destra della distribuzione evidenzia la presenza di alcuni nodi molto connessi, veri e propri hub della rete, che contribuiscono in modo decisivo alla sua coesione strutturale.

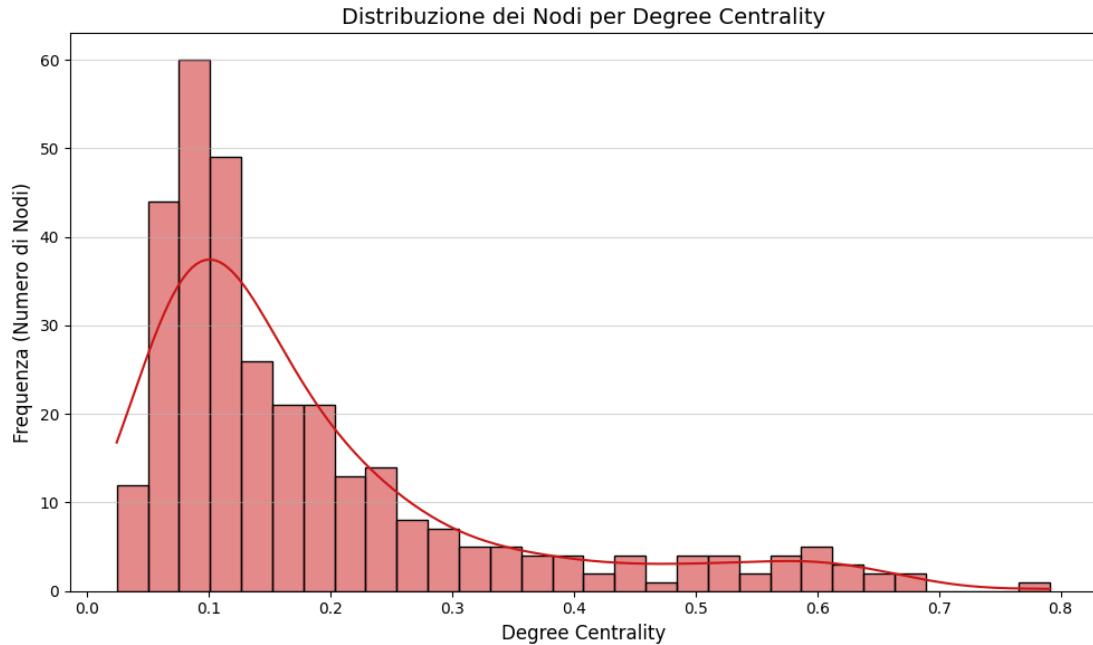


Figura 3.2: Distribuzione della degree centrality nella rete

Per rendere più intuitiva la distribuzione della *degree centrality*, è stata realizzata anche una rappresentazione grafica dell’intera rete, Figura 3.3, in cui ciascun nodo è colorato in base al proprio livello di centralità. I colori più chiari e tendenti al giallo indicano valori elevati, mentre quelli scuri rappresentano personaggi con poche connessioni.

La visualizzazione conferma quanto osservato nell’istogramma: i nodi maggiormente centrali si concentrano nell’area più densa del grafo, formando un nucleo compatto di personaggi altamente interconnessi. Intorno a questo centro si dirama una vasta area periferica composta da personaggi con pochi collegamenti, visualizzati con tonalità più scure.

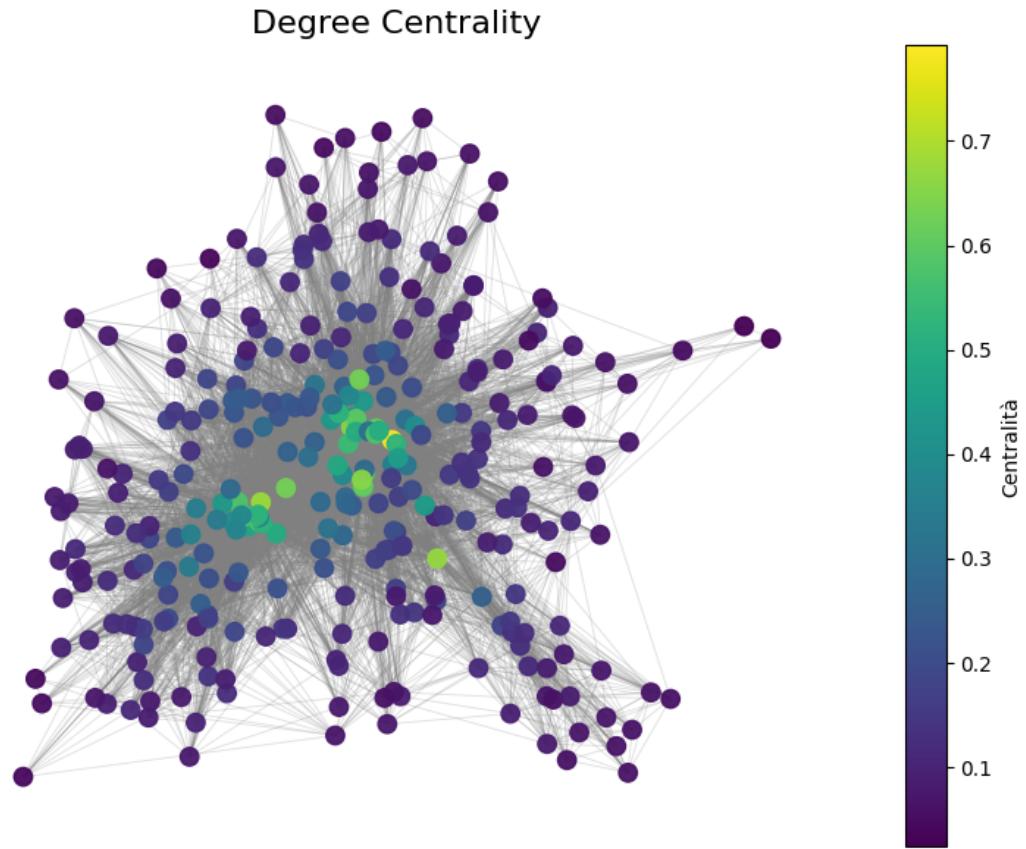


Figura 3.3: Rappresentazione della degree centrality nella rete

3.3 Betweenness Centrality

La *betweenness centrality* misura quanto un nodo funzioni da ponte in una rete, indicando quante volte quel nodo si trova sui cammini minimi che collegano coppie di altri nodi. Questo indice evidenzia il ruolo strategico di alcuni attori nel facilitare la connessione e la comunicazione tra diverse parti della rete: un personaggio con alta *betweenness centrality* può collegare aree che altrimenti sarebbero separate, anche se non ha molti collegamenti diretti.

Nelle reti sociali, chi possiede una *betweenness centrality* elevata è spesso fondamentale per la circolazione delle informazioni e può agire come mediatore tra comunità distanti. Questo ruolo di intermediazione permette ai nodi periferici di accedere a nodi centrali e, nei processi di diffusione, di controllare o modulare il flusso informativo tra gruppi differenti.

L'analisi di questa metrica è stata effettuata utilizzando la funzione di NetworkX `nx.betweenness_centrality(G, weight='distance')`, dove

$$distance = 1/Weight \quad (3.3.1)$$

Questa trasformazione è indispensabile perché, nel calcolo dei cammini più brevi, il legame più forte (peso maggiore) deve logicamente corrispondere alla distanza più breve (costo minore), assicurando che *betweenness centrality* individui i percorsi strategici che passano attraverso le co-appearance più frequenti significative della rete.

Per questa metrica, è stata adottata la stessa procedura utilizzata in precedenza ed il risultato è mostrato in Figura 3.4

Top 10 Personaggi per Betweenness Centrality:	
Character	Betweenness_Centrality
Captain America	0.531721
Beast \wedge Henry & Hank P	0.292841
Spider-man \wedge Peter Parker	0.200019
Wolverine \wedge Logan	0.186184
Cyclops \wedge Scott Summers	0.157628
Thor \wedge Dr. Donald Blak	0.121057
Thing \wedge Benjamin J. Gr	0.108350
Hulk \wedge Dr. Robert Bruce Banner	0.066748
Cannonball II \wedge Sam Gu	0.058575
Iron Man \wedge Tony Stark	0.055177

Bottom 10 Personaggi per B:	
Character	Betweenness_Centrality
Screaming Mimi \wedge Melis	0.0
Stankowicz, Fabian	0.0
Swordsman \wedge Jacques Du	0.0
O'brien, Michael	0.0
Diamondback II \wedge Rache	0.0
Aurora \wedge Jeanne-marie	0.0
Asp II \wedge Cleo	0.0
Hellstorm \wedge Daimon Hel	0.0
Bridge, George Washi	0.0
Hogun [Asgardian]	0.0

Figura 3.4: I 10 nodi con betweenness centrality più alta e più bassa

A differenza del caso precedente, l'ordine dei nodi risulta leggermente diverso. Nonostante ciò, Captain America (0.5317) domina nettamente la classifica, ricoprendo un ruolo strutturale cruciale: è il personaggio che più frequentemente si trova lungo i cammini minimi tra altre coppie di nodi. La distanza rispetto al secondo, Beast (0.2928), è considerevole, indicando un ruolo di leadership centrale assoluta nella connettività della rete. Seguono Spider-Man e Wolverine, anch'essi importanti snodi narrativi che collegano gruppi e storyline differenti.

La Bottom 10, invece, è composta esclusivamente da personaggi con betweenness pari a zero. Questi nodi non compaiono mai nei cammini minimi tra altri personaggi e ricoprono quindi ruoli totalmente periferici nella struttura globale della rete.

Successivamente, viene mostrata la distribuzione della *betweenness centrality* tramite l'istogramma in Figura 3.5, che rivela una distribuzione fortemente asimmetrica e con una concentrazione schiacciante verso valori minimi. Questa concentrazione estrema evidenzia che la maggior parte dei personaggi presenta un valore di centralità prossimo allo zero significando che, per la grande maggioranza dei nodi, l'informazione o l'interazione non deve necessariamente passare attraverso di loro per raggiungere un altro punto della rete; la comunicazione avviene tramite percorsi alternativi. Tuttavia, all'aumentare del valore, la frequenza crolla rapidamente. La lunga coda destra è l'elemento chiave, poiché rivela l'esistenza di un numero ristrettissimo di personaggi che raggiungono valori significativi e che si distinguono come veri e propri intermediari strategici (broker).

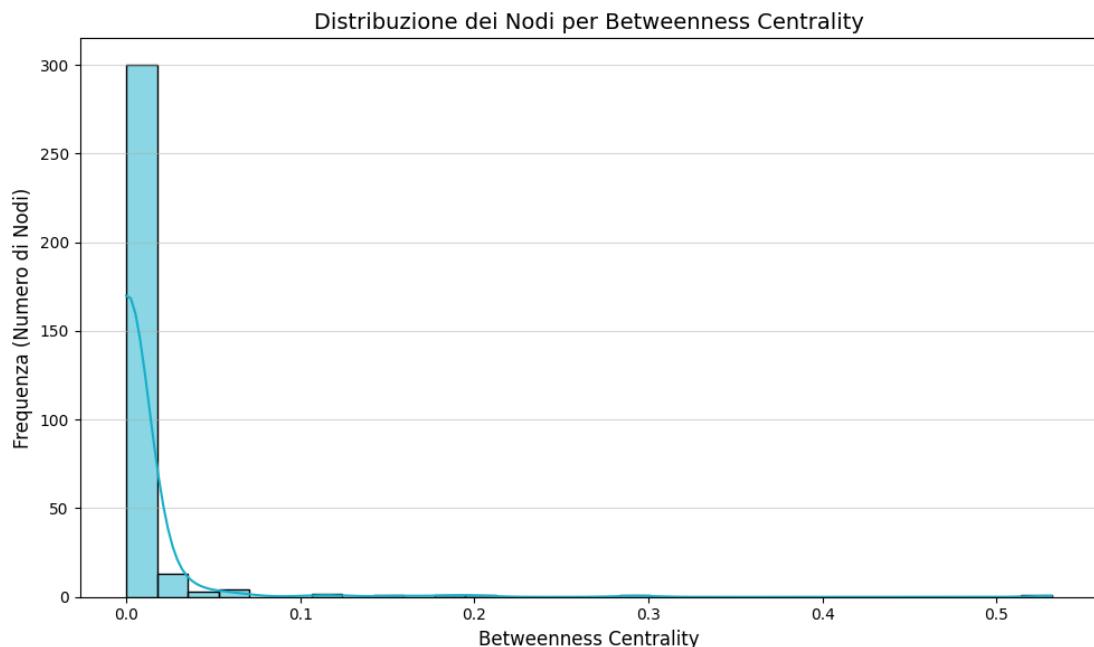


Figura 3.5: Distribuzione della betweenness centrality nella rete

Anche in questo caso, per rendere più intuitiva la distribuzione, è stata realizzata una rappresentazione grafica della rete, Figura 3.6, in cui ciascun nodo è colorato in base al proprio livello di centralità.

La visualizzazione conferma l'esistenza di un forte nucleo centrale (core), chiaramente identificato dai nodi di colore più chiaro (giallo/verde) che rappresentano i personaggi con la centralità più elevata. Questi hub si trovano strategicamente al centro della struttura, dominando l'organizzazione del grafo. Il rapido passaggio ai colori scuri (viola) evidenzia una vasta periferia che si connette al nucleo tramite percorsi limitati.

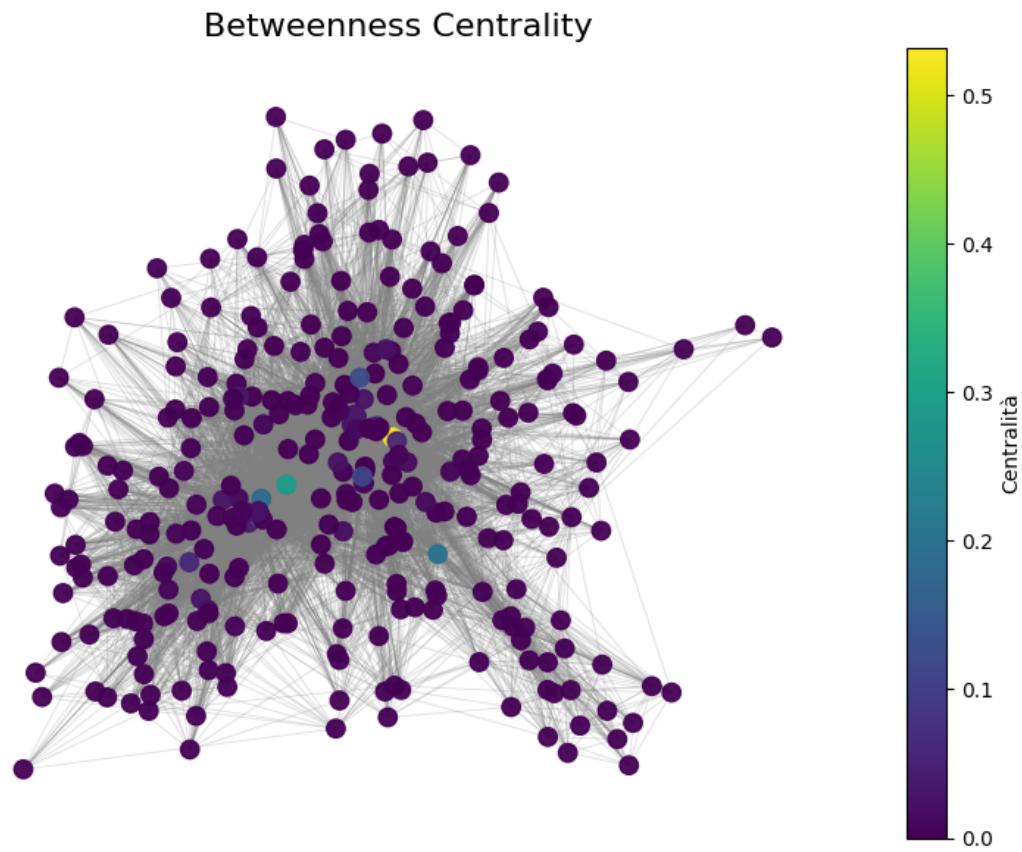


Figura 3.6: Rappresentazione della betweenness centrality nella rete

3.4 Closeness Centrality

La *closeness centrality* misura quanto rapidamente un nodo può raggiungere tutti gli altri nodi della rete. Formalmente, essa è definita come l'inverso della somma delle distanze minime dal nodo verso tutti gli altri: più un personaggio è "vicino" al resto della rete, più alta sarà la sua closeness.

In termini interpretativi, la *closeness centrality* identifica attori che possono diffondere informazioni, influenze o interazioni in modo rapido attraverso la rete, indipendentemente dal loro ruolo di snodo. A differenza della *betweenness centrality*, che evidenzia i punti di passaggio obbligati, la closeness cattura invece la posizione strategica complessiva di un nodo all'interno della struttura globale.

Anche per la *closeness* è stata applicata la stessa procedura, utilizzando la funzione `nx.closeness_centrality(G)`, ed il risultato è mostrato in Figura 3.7.

Top 10 Personaggi per Closeness Centrality:		
Character	Closeness_Centrality	
Captain America	0.827411	
Wolverine ↗ Logan	0.754630	
Spider-man ↗ Peter Parker	0.751152	
Thing ↗ Benjamin J. Gr Vision	0.744292 0.742597	
Thor ↗ Dr. Donald Blak	0.729306	
Beast ↗ Henry & Hank & P	0.729306	
Iron Man ↗ Tony Stark	0.726058	
Mr. Fantastic ↗ Reed R	0.718062	
Cyclops ↗ Scott Summer	0.716484	

Bottom 10 Personaggi per Closeness Centrality:		
Character	Closeness_Centrality	
Deadpool ↗ Jack ↗ Wade W Lubenski, Nate	0.401478 0.441734	
Mercado, Joy	0.442935	
Katzenberg, Nick	0.442935	
Robertson, Randy	0.443537	
Cushing, Kate	0.444748	
Grant, Gloria Glory Arranger ↗	0.447802 0.449036	
Skin ↗ Angelo Espinosa	0.449036	
Husk ↗ Paige Guthrie	0.449655	

Figura 3.7: I 10 nodi con closeness centrality più alta e più bassa

I valori ottenuti mostrano una struttura fortemente gerarchica nella rete Marvel. Nella Top 10 compaiono personaggi iconici come Captain America, Wolverine, Spider-Man, Thing, Vision e Iron Man. Questi nodi presentano i valori più elevati perché si trovano in posizioni particolarmente “centrali” nel grafo: la loro distanza media da tutti gli altri personaggi è ridotta, permettendo loro — in termini strutturali — di raggiungere rapidamente qualsiasi punto della rete.

Questo risultato riflette il loro ruolo narrativo: protagonisti che compaiono in molte storyline e che interagiscono con gruppi diversi, generando percorsi brevi verso un’ampia porzione dell’universo Marvel.

La Bottom 10, al contrario, include personaggi periferici o narrativamente meno integrati. I loro valori di closeness sono molto più bassi, indicando che si trovano in zone marginali della rete, lontane dal nucleo più denso.

Analogamente a quanto visto per la *betweenness centrality* viene mostrata la distribuzione della *closeness centrality* tramite l’istogramma in Figura 3.8.

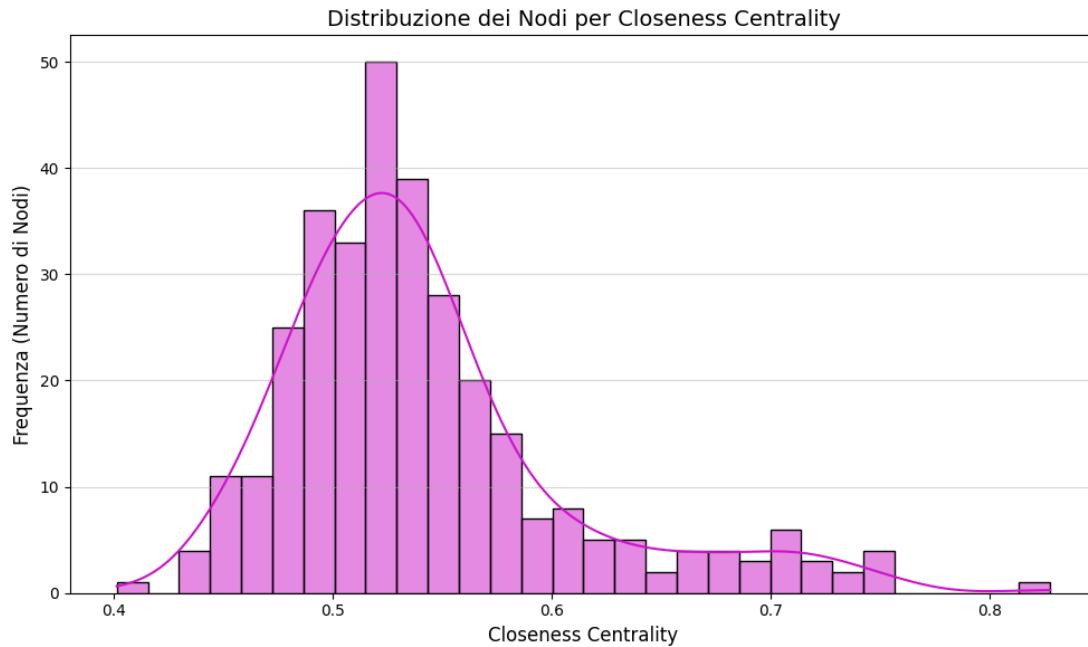


Figura 3.8: Distribuzione della closeness centrality nella rete

A differenza della *degree centrality*, la distribuzione della *closeness centrality* è meno sbilanciata e assume una forma più simile a una curva a campana, con un marcato picco centrale e una coda destra meno pronunciata. La maggior parte dei nodi si concentra su valori intermedi di closeness (circa tra 0.45 e 0.55), il che indica che molti nodi possono raggiungere gli altri con pochi passaggi, pur non essendo né estremamente centrali né del tutto periferici. Solo pochi nodi presentano valori molto elevati di *closeness*, visibili nella coda destra, e rappresentano personaggi particolarmente centrali, molto vicini al resto della rete.

Infine, la rappresentazione grafica della rete, Figura 3.9, conferma quanto osservato nella distribuzione: i nodi con valori medi di closeness centrality sono diffusi in tutto il grafo, mentre i pochi nodi con valori più elevati (in verde-giallo) si concentrano nelle zone centrali della struttura, da cui possono raggiungere rapidamente la maggior parte degli altri personaggi.

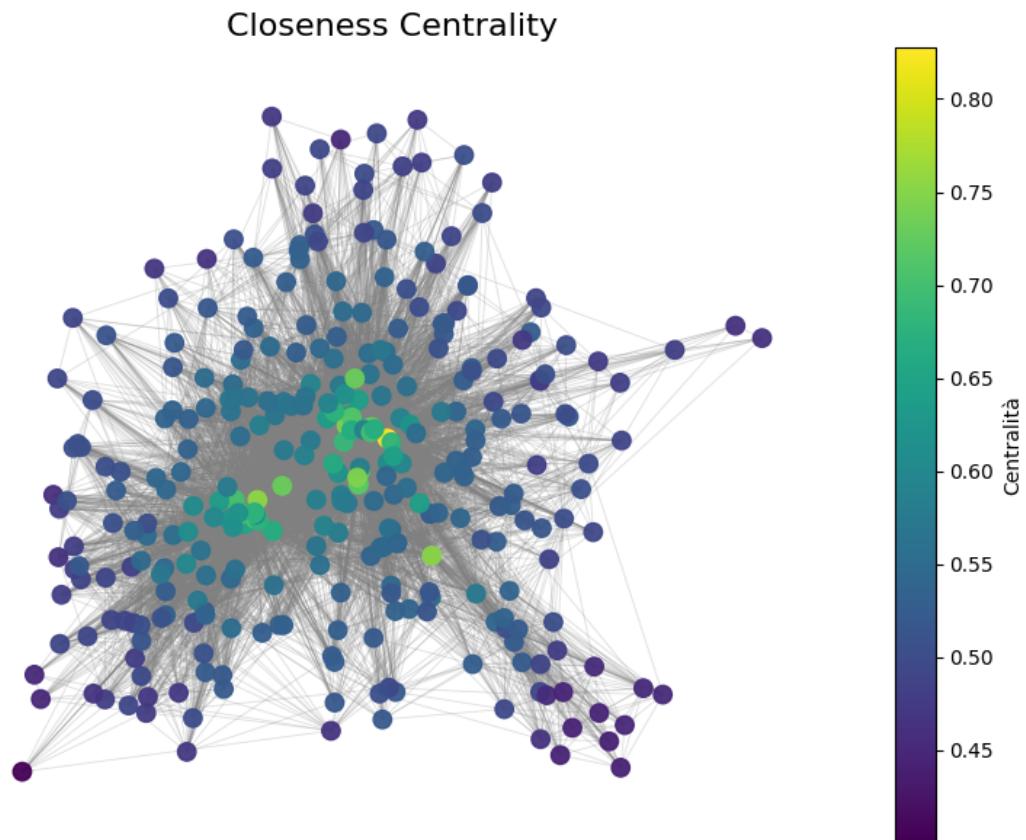


Figura 3.9: Rappresentazione della closeness centrality nella rete

3.5 Eigenvector Centrality

L'*eigenvector centrality* è una misura di centralità che valuta l'importanza di un nodo tenendo conto non solo del numero di collegamenti che possiede, ma anche dell'importanza dei nodi a cui è collegato. In altre parole, un nodo ottiene un punteggio elevato se è connesso a molti altri nodi che, a loro volta, hanno una centralità elevata.

Rispetto alla *degree centrality*, che conta semplicemente il numero di archi incidenti, l'*eigenvector centrality* premia la “qualità” dei vicini più che la loro quantità.

Top 10 Personaggi per Eigenvector Centrality:	
Character	Eigenvector_Centrality
Captain America	0.139418
Wolverine ↳ Logan	0.131218
Thing ↳ Benjamin J. Gr	0.129755
Vision	0.127247
Beast ↳ Henry & Hank P	0.126039
Spider-man ↳ Peter Parker	0.123751
Human Torch ↳ Johnny S	0.123180
Invisible Woman ↳ Sue	0.123150
Cyclops ↳ Scott Summers	0.123089
Scarlet Witch ↳ Wanda	0.122284

Bottom 10 Personaggi per Eigenvector Centrality:	
Character	Eigenvector_Centrality
Asp Ii ↳ Cleo	0.003236
Black Mamba ↳ Tanya Se	0.003236
Katzenberg, Nick	0.004394
Mercado, Joy	0.004586
Deadpool ↳ Jack ↳ Wade W	0.004735
Cushing, Kate	0.004794
Lubenski, Nate	0.004811
Robertson, Randy	0.005029
Arranger ↳	0.005348
Hammerhead	0.005622

Figura 3.10: I 10 nodi con eigenvector centrality più alta e più bassa

I risultati in Figura 3.10 mostrano che i personaggi con valori più alti di eigenvector centrality coincidono con quelli che avevano già i valori più elevati di degree centrality. Questo indica che gli attori più connessi della rete (come Captain America, Wolverine, Spider-Man, i membri dei Fantastici Quattro, ecc.) non solo possiedono molti collegamenti diretti, ma sono anche collegati tra loro e ad altri nodi a loro volta molto centrali, formando il nucleo più denso e “prestigioso” della rete. Inoltre, questa sovrapposizione tra le due misure suggerisce che, l’eigenvector centrality non modifica in modo sostanziale il ranking fornito dalla degree centrality, ma lo conferma e lo raffina.

La differenza tra le due metriche può essere osservata nella loro distribuzione. Infatti, mentre la degree centrality mostra una forte concentrazione su valori bassi e una lunga coda dovuta alla presenza di pochi hub molto connessi, nell’eigenvector centrality, Figura 3.11, la coda verso valori elevati è più graduale e distribuita. Questo significa che l’importanza non dipende solo dal numero di collegamenti di un nodo, ma anche dal fatto che esso sia connesso a nodi centrali; di conseguenza, alcuni nodi con grado medio possono comunque ottenere valori significativi di eigenvector centrality grazie alla qualità dei loro collegamenti. Tale risultato riflette la capacità di questa misura di

considerare sia la quantità sia il “prestigio” delle connessioni nella struttura della rete.

Infine, la rappresentazione grafica in Figura 3.12 mostra come i valori di *eigenvector centrality* si distribuiscono all’interno della rete: i nodi più centrali (in verde/giallo) si concentrano nel nucleo denso del grafo, dove si trovano anche gli hub individuati dalla *degree centrality*. I nodi periferici, visualizzati con colori più scuri, mantengono invece valori molto bassi, confermando che l’importanza si propaga soprattutto all’interno del cuore strutturale della rete, coerentemente con quanto osservato nella distribuzione dei valori di *eigenvector*.

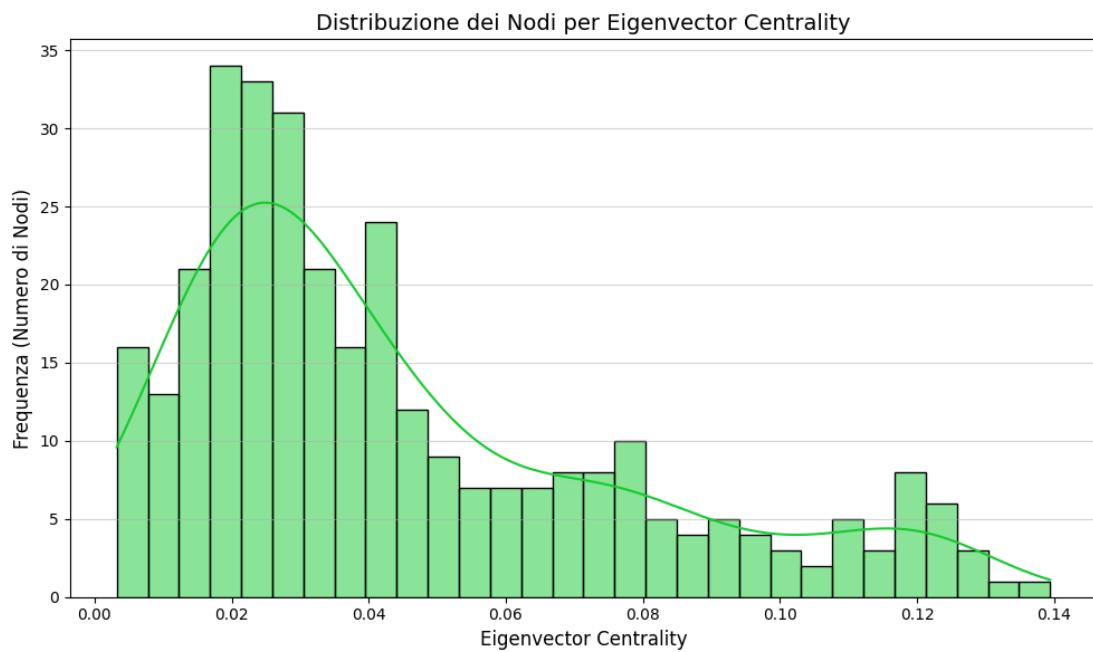


Figura 3.11: Distribuzione dell’eigenvector centrality nella rete

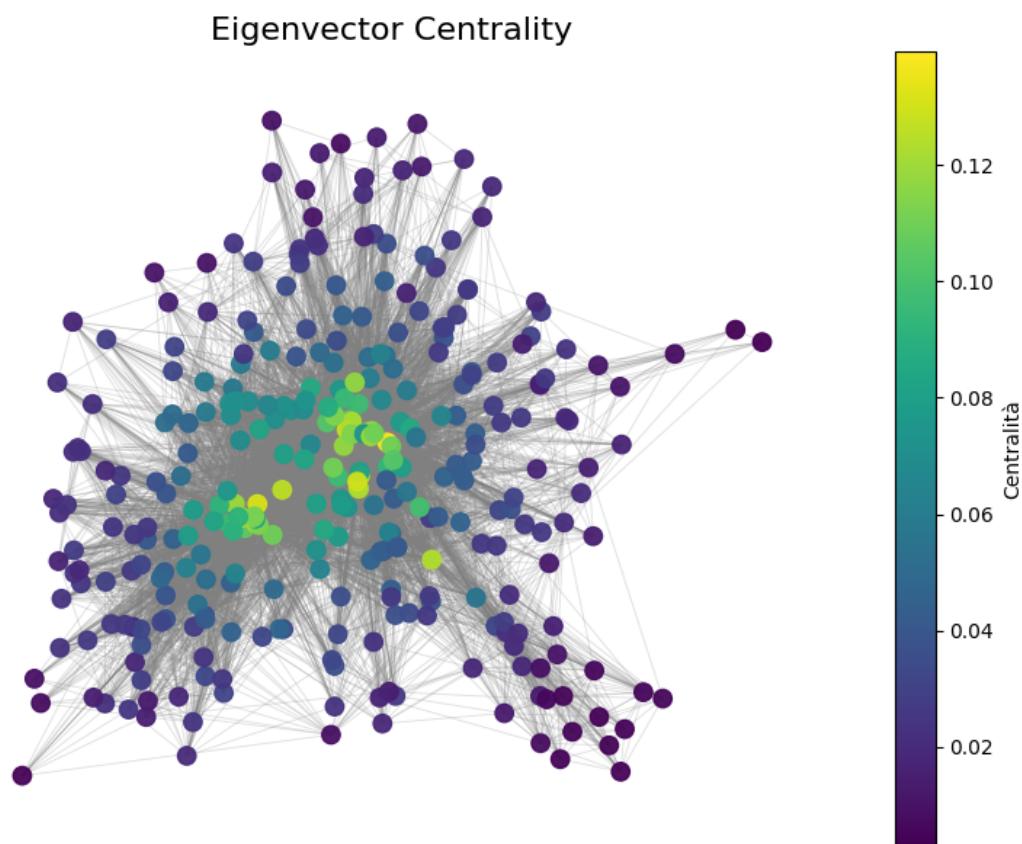


Figura 3.12: Rappresentazione dell'eigenvector centrality nella rete

CAPITOLO 4

Analisi delle strutture

In questo capitolo vengono approfondite le configurazioni e i pattern che caratterizzano la rete sociale a livelli più complessi rispetto ai singoli nodi e archi. L'analisi strutturale permette di scoprire come le relazioni tra gli attori si organizzano in sottogruppi, nuclei coesi e zone di frontiera, rivelando le dinamiche interne, i centri di aggregazione e le possibili linee di divisione nella rete.

4.1 Triadi

Una triade (o triangolo) è una configurazione formata da tre nodi tutti reciprocamente collegati. Si tratta della struttura coesiva più semplice e più significativa all'interno di una rete, perché rappresenta un piccolo gruppo in cui ogni attore è direttamente connesso agli altri due. Nelle reti sociali, la presenza di triadi indica un'elevata tendenza alla chiusura e alla collaborazione, mentre in un contesto narrativo come quello della Marvel suggerisce la co-occorrenza frequente di tre personaggi nelle stesse storyline o negli stessi team.

Studiare le triadi permette di comprendere la densità e la coesione locale della rete, di individuare piccoli nuclei fortemente interconnessi e di cogliere le dinamiche di aggregazione che non emergono dalle sole misure di centralità. Esse costituiscono dunque un indicatore micro-strutturale prezioso per interpretare il comportamento complessivo della rete.

Nel contesto del grafo Marvel, è stato innanzitutto calcolato il numero totale di triadi presenti nella rete, così da ottenere una misura diretta della coesione locale tra i personaggi. Il risultato è particolarmente significativo: sono state individuate 155.834 triadi, un valore estremamente elevato che conferma la forte tendenza dell'universo Marvel a strutturarsi in piccoli gruppi ricorrenti. Questa densità di triangoli riflette la

natura collaborativa del fumetto, in cui i personaggi compaiono spesso insieme, sia all'interno di team consolidati, sia lungo storyline condivise.

L'analisi è stata poi approfondita tenendo conto del peso degli archi: per ogni triade sono stati registrati i tre personaggi coinvolti, l'intensità delle loro connessioni e la somma complessiva dei pesi. L'obiettivo non era quello di analizzare singolarmente ogni triade, ma di individuare quelle potenzialmente "più forti", ossia caratterizzate da una maggiore frequenza di co-appearance.

Come si può osservare in Figura 4.1, un elemento degno di nota è la ricorrenza di alcuni personaggi nelle triadi più pesanti: ad esempio, Capitan America e Thing compaiono in numerose configurazioni significative. Ciò indica che questi personaggi fungono da nodi di snodo all'interno della rete, collegando gruppi diversi e agendo come punti di intersezione tra varie storyline o team. Questa osservazione conferma e amplia quanto già emerso dalle analisi di centralità, mostrando come certi attori rivestano un ruolo strutturale rilevante anche a livello micro-strutturale.

Numero Totale di Triadi (Triangoli) nel Grafo: 155834

```
• Triadi più pesanti:
Triade: ('Human Torch ↗ Johnny S', 'Mr. Fantastic ↗ Reed R', 'Thing ↗ Benjamin J. Gr') | Pesi: (713, 744, 708) | Peso totale: 2165
Triade: ('Human Torch ↗ Johnny S', 'Invisible Woman ↗ Sue', 'Mr. Fantastic ↗ Reed R') | Pesi: (713, 694, 701) | Peso totale: 2108
Triade: ('Human Torch ↗ Johnny S', 'Invisible Woman ↗ Sue', 'Thing ↗ Benjamin J. Gr') | Pesi: (744, 694, 668) | Peso totale: 2106
Triade: ('Invisible Woman ↗ Sue', 'Mr. Fantastic ↗ Reed R', 'Thing ↗ Benjamin J. Gr') | Pesi: (701, 708, 668) | Peso totale: 2077
Triade: ('Jameson, J. Jonah', 'Spider-man ↗ Peter Parker', 'Watson-parker, Mary') | Pesi: (616, 526, 235) | Peso totale: 1377
Triade: ('Captain America', 'Iron Man ↗ Tony Stark', 'Scarlet Witch ↗ Wanda') | Pesi: (372, 374, 446) | Peso totale: 1192
Triade: ('Parker, May', 'Spider-man ↗ Peter Parker', 'Watson-parker, Mary') | Pesi: (616, 380, 188) | Peso totale: 1184
Triade: ('Captain America', 'Scarlet Witch ↗ Wanda', 'Vision') | Pesi: (422, 385, 374) | Peso totale: 1181
Triade: ('Captain America', 'Iron Man ↗ Tony Stark', 'Thor ↗ Dr. Donald Blak') | Pesi: (344, 446, 386) | Peso totale: 1176
Triade: ('Jameson, J. Jonah', 'Robertson, Joe', 'Spider-man ↗ Peter Parker') | Pesi: (356, 526, 286) | Peso totale: 1168
```

Figura 4.1: Visualizzazione delle triadi più pesanti nella rete

4.2 Clique

Una clique è un gruppo di nodi in una rete in cui ogni nodo è collegato a tutti gli altri del gruppo, formando così un piccolo sottogruppo molto coeso. In una rete di personaggi, ad esempio, una clique rappresenta un insieme di personaggi che interagiscono frequentemente tra loro, comparendo insieme in varie situazioni o storie. Le triadi chiuse, o triangoli, sono un caso particolare di clique composta da tre nodi, ma le clique possono avere qualsiasi dimensione.

Le clique sono importanti perché permettono di individuare gruppi strettamente connessi all'interno di una rete più ampia, mostrando come alcuni nodi fungano da collegamento tra diversi sottogruppi e contribuendo a capire la struttura complessiva della rete.

Nella rete dei personaggi Marvel, Figura 4.2, sono state identificate 6396 clique, cioè gruppi di nodi in cui ogni personaggio è collegato a tutti gli altri del gruppo. Queste clique rappresentano piccoli sottogruppi molto coesi all'interno della rete.

```

Numero totale di clique trovate: 6396

Prime 10 clique (ordinate per dimensione):
Clique di dimensione 3: ['Hammerhead', 'Nomad Iii ↳ Jack Monro', 'Daredevil ↳ Matt Murdo']
Clique di dimensione 3: ['Gee ↳ Alex Power', 'Leech', 'Richards, Franklin B']
Clique di dimensione 3: ['Bridge, George Washi', 'Summers, Nathan Chri', 'Fury, Col. Nicholas']
Clique di dimensione 2: ['Deadpool ↳ Jack ↳ Wade W', 'Death']

Numero di clique di dimensione 5: 11
Clique di dimensione 5: ['Captain America', 'Kincaid, Dr. Jane Fo', 'Thor ↳ Dr. Donald Blak', 'Cobra ↳ Klaus Vorhees', 'Mister Hyde ↳ Calvin Z']
Clique di dimensione 5: ['Gee ↳ Alex Power', 'Beta Ray Bill', 'Mr. Fantastic ↳ Reed R', 'Human Torch ↳ Johnny S', 'Richards, Franklin B']
Clique di dimensione 5: ['Gee ↳ Alex Power', 'Spider-man ↳ Peter Parker', 'Spider-man Clone ↳ Ben', 'Firestar ↳ Angelica Jo', 'Justice Ii ↳ Vance Ast']
Clique di dimensione 5: ['Hobgoblin V ↳ Jason Ph', 'Spider-man ↳ Peter Parker', 'Ghost Rider Ii ↳ Johnn', 'Hulk ↳ Dr. Robert Bruce Banner', 'Ghost Rider Iii ↳ Dan']
Clique di dimensione 5: ['Hobgoblin V ↳ Jason Ph', 'Spider-man ↳ Peter Parker', 'Watson-parker, Mary', 'Ghost Rider Iii ↳ Dan', 'Hulk ↳ Dr. Robert Bruce Banner']
Clique di dimensione 5: ['Hobgoblin V ↳ Jason Ph', 'Spider-man ↳ Peter Parker', 'Watson-parker, Mary', 'Moon Knight ↳ Marc Spe', 'Hulk ↳ Dr. Robert Bruce Banner']
Clique di dimensione 5: ['Bridge, George Washi', 'Fixer Ii ↳ Paul Norber', 'Hawk', 'Fury, Col. Nicholas', 'Dugan, Timothy Aloyi']
Clique di dimensione 5: ['Bridge, George Washi', 'Summers, Nathan Chri', 'Phoenix Iii ↳ Rachel S', 'Cannonball Ii ↳ Sam Gu', 'Sunsplot ↳ Roberto Daco']
Clique di dimensione 5: ['Tombstone ↳ Lonnlie Tho', 'Spider-man ↳ Peter Parker', 'Darkhawk ↳ Chris Powel', 'Daredevil ↳ Matt Murdo', 'Punisher Ii ↳ Frank Ca']
Clique di dimensione 5: ['Ghost Rider Iii ↳ Dan', 'Spider-man ↳ Peter Parker', 'Punisher Ii ↳ Frank Ca', 'Watson-parker, Mary', 'Wolverine ↳ Logan']
Clique di dimensione 5: ['Spider-man Clone ↳ Ben', 'Spider-man ↳ Peter Parker', 'Watson-parker, Mary', 'Firestar ↳ Angelica Jo', 'Justice Ii ↳ Vance Ast']

```

Figura 4.2: Visualizzazione delle clique nella rete

Le clique di dimensione 2 o 3, che corrispondono rispettivamente a coppie e triangoli, sono molto frequenti e indicano interazioni dirette tra pochi personaggi. In totale, ci sono diverse decine di queste piccole clique, alcune delle quali includono personaggi noti come Spider-Man, Daredevil o Franklin Richards.

A scopo di esempio, sono state, inoltre, identificate 11 clique di dimensione 5, cioè gruppi di cinque personaggi tutti collegati tra loro. Questi sottogruppi più grandi mostrano team o storyline dove le interazioni sono particolarmente dense, con personaggi come Captain America, Thor, Hulk e Spider-Man presenti in più di un gruppo.

La massima clique della rete ha una dimensione di 46 nodi, Figura 4.3, comprendendo molti personaggi principali degli X-Men, degli Avengers e di altre squadre, come Captain America, Wolverine, Thor, Scarlet Witch e Thanos. Questo indica che, pur nella vastità della rete, esistono gruppi molto grandi di personaggi strettamente interconnessi, che giocano un ruolo centrale nella coesione complessiva della rete.

```

Dimensione della Massima Clique: 46

['Captain America', 'Vision', 'Scarlet Witch ↳ Wanda', 'Hudson, Heather', 'Beast ↳ Henry & Hank& P', 'Wonder Man ↳ Simon Wil']
['Thing ↳ Benjamin J. Gr', 'Human Torch ↳ Johnny S', 'She-hulk ↳ Jennifer Wa', 'Hercules [greek God]', 'Hulk ↳ Dr. Robert Bruce Banner', 'Sasquatch ↳ Walter Lan']
['Storm ↳ Ororo Munroe S', 'Wolverine ↳ Logan', 'Cyclops ↳ Scott Summer', 'Marvel Girl ↳ Jean Grey', 'Angel ↳ Warren Kenneth', 'Colossus Ii ↳ Peter Ra']
['Rogue ↳ ', 'Iceman ↳ Robert Bobby', 'Psylocke ↳ Elisabeth B', 'Havok ↳ Alex Summers', 'Invisible Woman ↳ Sue', 'Black Knight V ↳ Dane']
['Magus', 'Living Lightning ↳ Mig', 'Nova ↳ Richard Rider', 'Thunderstrike ↳ Eric K', 'Quasar Iii ↳ Wendell V', 'Black Widow ↳ Natasha']
['Strong Guy ↳ Guido Car', 'Professor X ↳ Charles', 'Dr. Strange ↳ Stephen', 'Moondragon ↳ Heather D', 'Harkness, Agatha', 'Silver Surfer ↳ Norrin']
['Galactus ↳ Galan', 'Dr. Druid ↳ Anthony Lu', 'Drax ↳ Arthur Douglas', 'Nova Ii ↳ Frankie Raye', 'Thanos', 'Warlock Ii ↳ Adam Warl']
['Pip ↳ Prince Gofern', 'Gamora', 'Thanos Doppelganger', 'Shaman ↳ Michael Twoyo']

```

Figura 4.3: Visualizzazione della clique massima

Infine, è possibili osservare la rappresentazione grafica di quest'ultima in Figura 4.4.

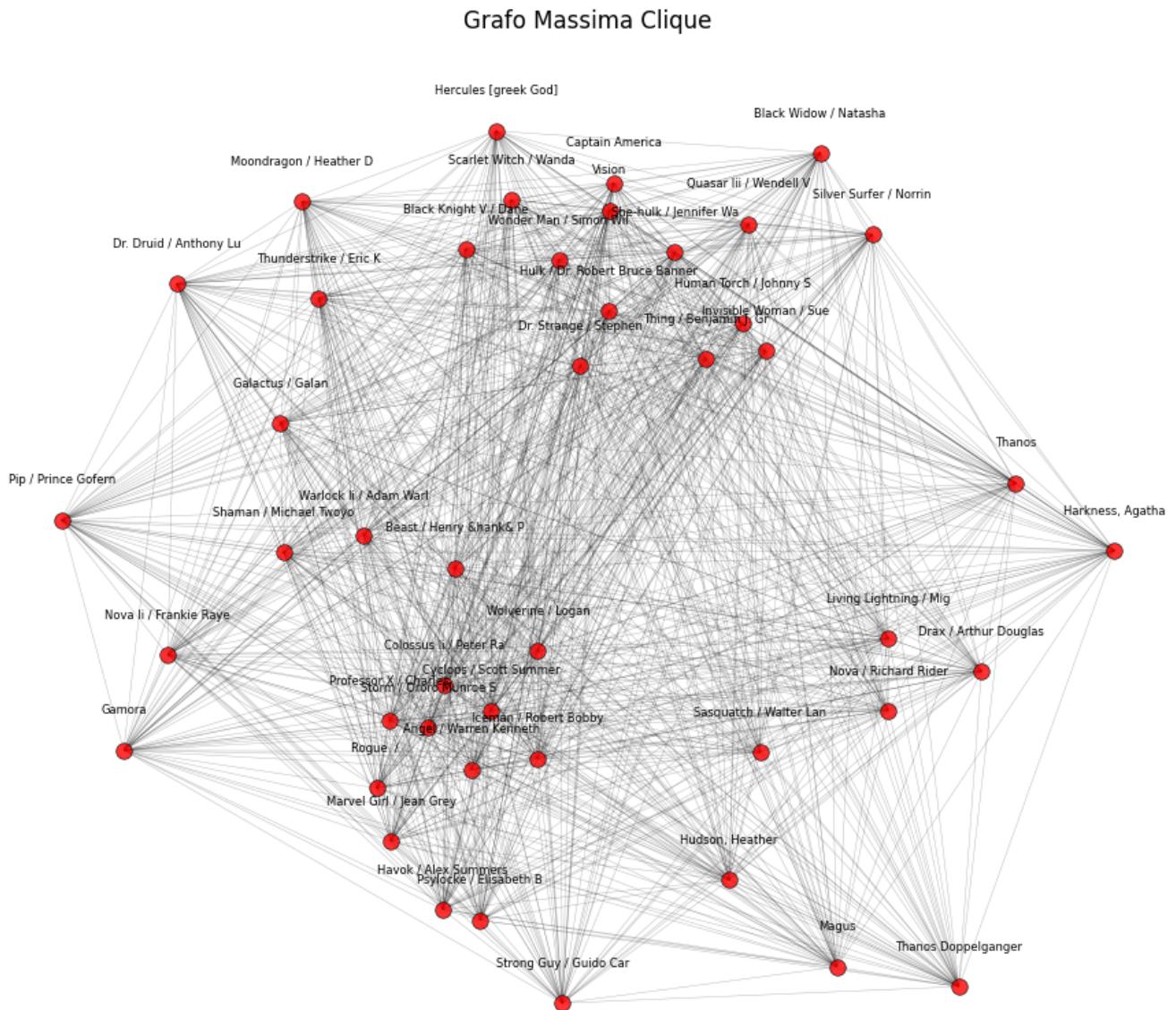


Figura 4.4: Rappresentazione grafica della clique massima

4.3 K-core

Il k-core è uno strumento utilizzato per capire quali sono le parti più solide e centrali di una rete. L’idea è molto semplice: un k-core è il gruppo di nodi che hanno almeno k connessioni con altri nodi dello stesso gruppo.

Per individuarlo, si “ripulisce” progressivamente la rete rimuovendo i nodi meno collegati: prima quelli con un solo collegamento, poi quelli con due, e così via. Dopo ogni rimozione, si ricontrollano i gradi dei nodi rimanenti, finché non resta un insieme stabile di nodi che soddisfa tutti la stessa soglia di connessioni (k).

Più è alto il valore di k , più il gruppo individuato è compatto, interconnesso e strutturalmente importante. A differenza del semplice grado di un nodo, il k-core mette in luce non solo quanti collegamenti ha un nodo, ma anche quanto è immerso nel cuore della rete, cioè quanto è sostenuto da una struttura di connessioni altrettanto forti.

Nel caso della rete Marvel considerata, Figura 4.5, il valore del k-core massimo è pari a 54. Questo significa che tutti i nodi appartenenti a questo sottografo possiedono almeno 54 connessioni con altri nodi dello stesso gruppo: una soglia molto alta, che indica una regione della rete estremamente coesa.

Il k-core massimo è composto da 68 nodi e 2148 archi, formando un sottografo molto denso e compatto. La presenza di un numero così elevato di connessioni suggerisce che questi personaggi rappresentano il vero cuore della rete: sono figure altamente ricorrenti, spesso coinvolte in molteplici storyline, team e interazioni trasversali.

K-Core Massimo Trovato (k): 54
Numero di nodi nel k-core massimo: 68
Numero di archi nel k-core massimo: 2148

Figura 4.5: Visualizzazione del k-core di grado massimo

La rappresentazione grafica del k-core massimo, in Figura 4.6, evidenzia proprio questa forte interconnessione: i nodi appaiono tutti strettamente legati tra loro, formando una sorta di "nucleo centrale" della rete Marvel. Questo nucleo è composto principalmente da personaggi principali appartenenti a gruppi storici come Avengers o X-Men.

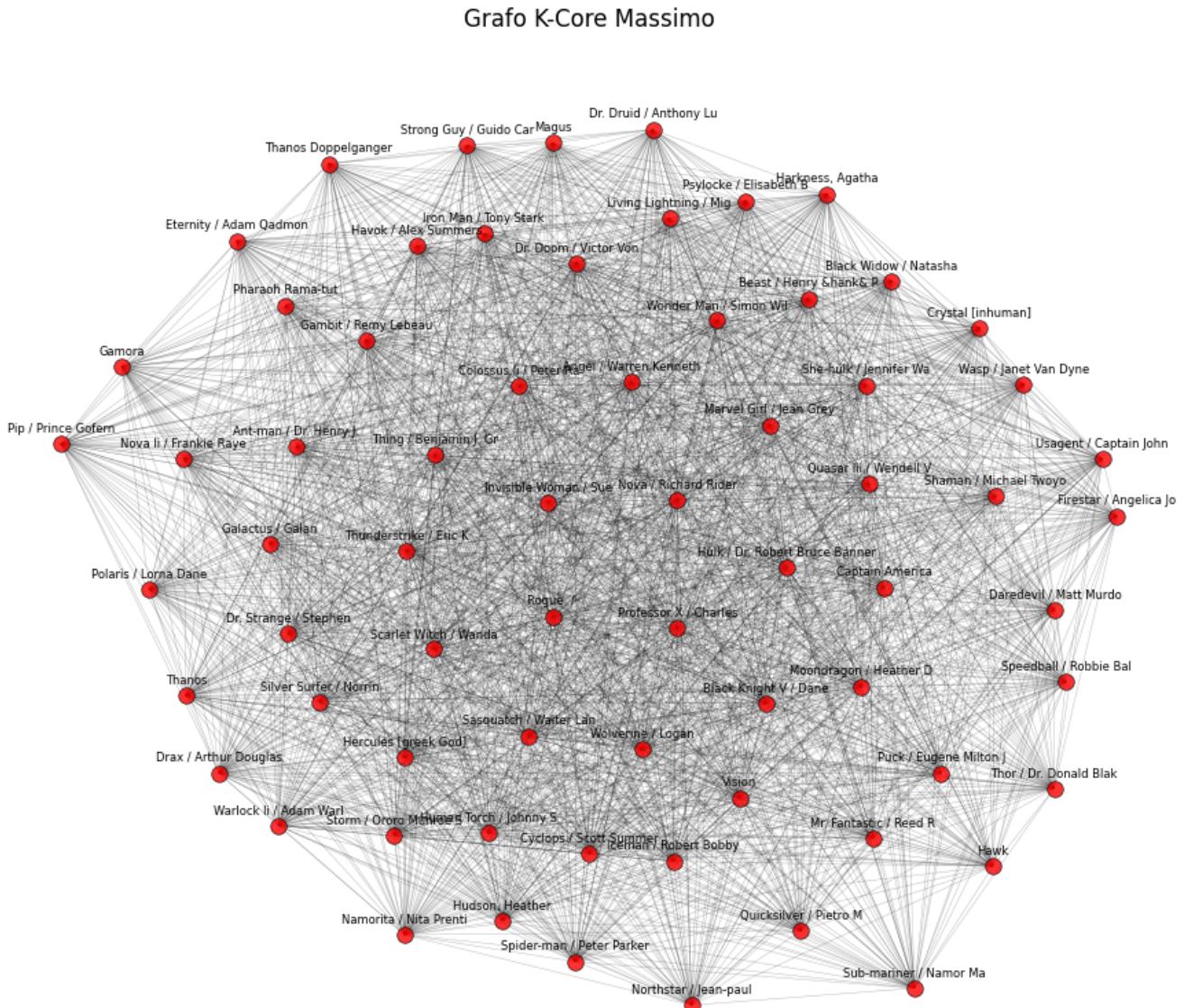


Figura 4.6: Rappresentazione grafica del k-core, con k=56

4.4 Ego network

La ego network è una sottostruttura della rete costruita intorno a un singolo nodo, chiamato ego. Si tratta della rete personale di un personaggio, composta da tutti i nodi con cui è direttamente collegato e dalle relazioni che esistono tra questi nodi. In questo

modo la ego network rappresenta l’ambiente locale di un personaggio e permette di osservare come si inserisce nel contesto narrativo complessivo.

Analizzare la ego network di un nodo significa quindi studiare il suo ruolo all’interno della rete: se è collegato a pochi o a molti altri personaggi, se questi personaggi formano tra loro gruppi coesi oppure sono poco connessi, e se l’ego funge da punto di collegamento tra mondi o storyline diverse. Questo tipo di analisi mette in evidenza non solo l’importanza di un personaggio in termini di numero di relazioni, ma soprattutto la struttura di queste relazioni.

Per l’analisi della ego network è stato selezionato come nodo centrale Captain America. Tale scelta deriva dal fatto che, in tutte le principali metriche di centralità considerate, il personaggio presenta sistematicamente i valori più elevati, risultando uno dei nodi strutturalmente più influenti dell’intera rete Marvel. La sua ego network costituisce quindi un caso particolarmente indicativo per osservare le dinamiche locali che caratterizzano il contesto relazionale di un nodo altamente centrale.

La rappresentazione grafica della ego network completa, Figura 4.7, mostra una struttura estremamente densa, con un numero molto elevato di nodi e connessioni. A causa di questa complessità, la visualizzazione risulta poco leggibile e tende a mascherare le relazioni più rilevanti.

Per ottenere una visione più chiara e interpretabile, è stata realizzata una seconda rappresentazione, Figura 4.8, limitata ai primi dieci nodi con il valore di degree centrality più alto tra i vicini di Captain America. Tale selezione consente di focalizzarsi sui personaggi che svolgono un ruolo più significativo nel suo intorno immediato, evidenziando le connessioni che contribuiscono maggiormente alla sua posizione centrale nella rete.

Dal grafico si nota, infatti, come Captain America sia fortemente connesso a molti personaggi chiave dell’universo Marvel, tra cui Iron Man, Spider-Man, Thor e Scarlet Witch. I nodi vicini sono tutti collegati tra loro, evidenziando una rete molto densa e coesa: questo conferma la sua posizione centrale e la sua importanza nelle interazioni tra i personaggi principali.

Ego Network di Captain America

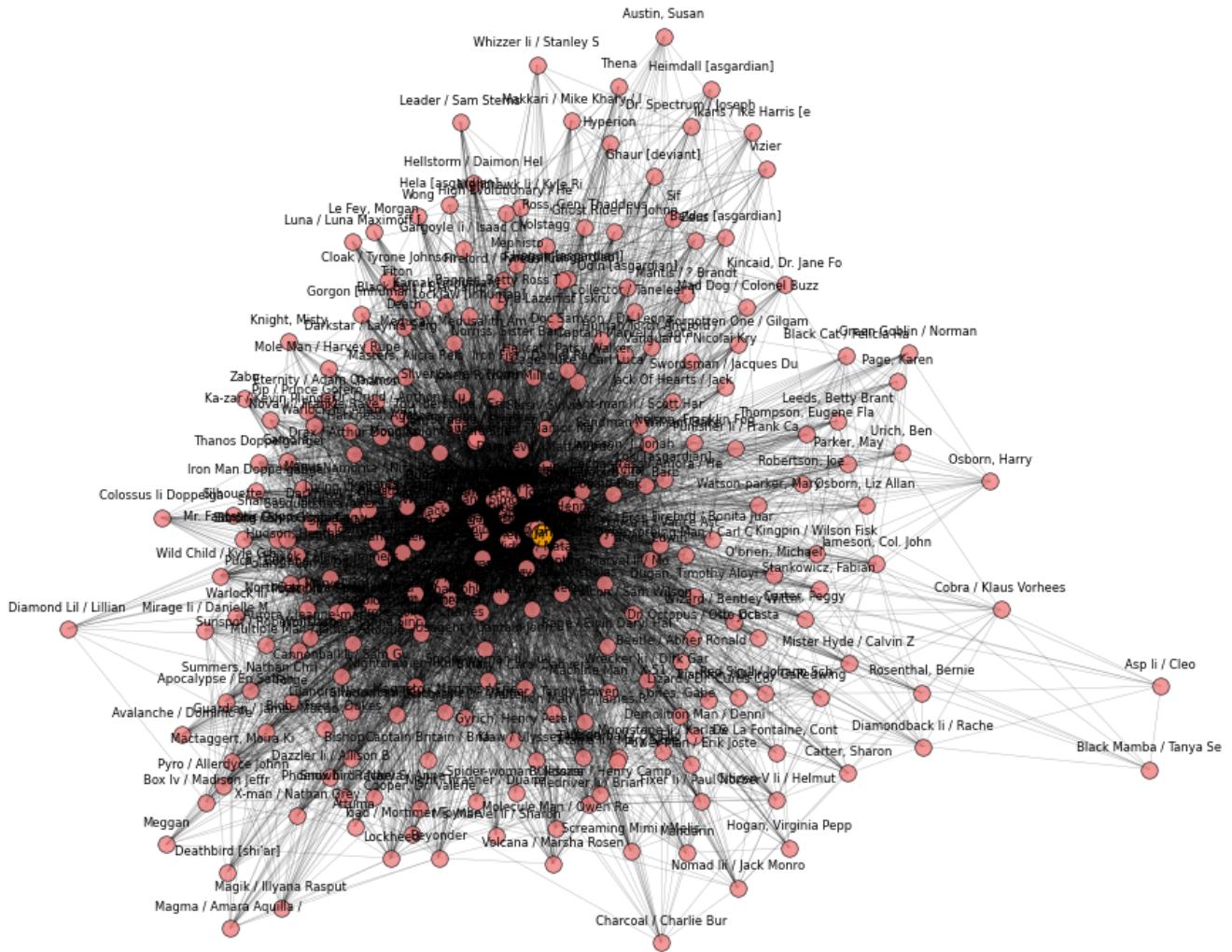


Figura 4.7: Rappresentazione grafica della ego-network di Captain America

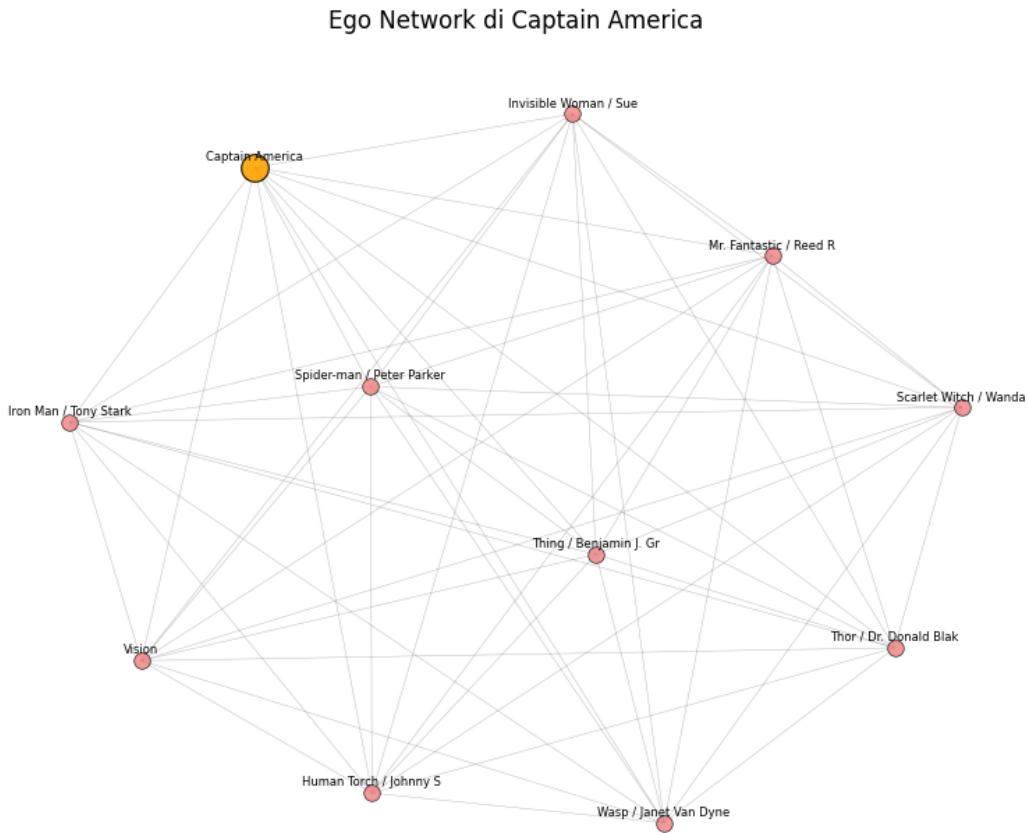


Figura 4.8: Grafo della ego-network dei 10 vicini con centralità maggiore

4.5 Communities

Una comunità è costituita da un insieme di nodi che presentano tra loro un numero di connessioni superiore rispetto a quelle che intrattengono con il resto della rete. In altre parole, all'interno di una comunità le relazioni risultano più dense e frequenti, mentre i legami con l'esterno tendono a essere meno numerosi. Questo tipo di analisi permette quindi di evidenziare la presenza di gruppi funzionali o narrativi, mettendo in luce affinità, collaborazioni o appartenenze tematiche tra i personaggi.

L'identificazione delle comunità consente di osservare la rete non come un insieme unico e indistinto, ma come una struttura articolata, caratterizzata da blocchi interni più coesi. In una rete complessa come quella dei personaggi Marvel, questo approccio permette di riconoscere team storici, gruppi narrativi ricorrenti o insiemi di personaggi che condividono trame e contesti simili.

Nell'analisi delle comunità sono stati adottati due diversi metodi disponibili in NetworkX: l'algoritmo di *Louvain* e quello basato sulla *greedy modularity*. Entrambi

mirano a individuare una suddivisione della rete in gruppi internamente più densi, massimizzando la modularità, ma lo fanno attraverso strategie differenti.

L'algoritmo di *Louvain* si caratterizza per un approccio gerarchico e particolarmente efficiente su reti di grandi dimensioni. Il procedimento si articola in due fasi iterative: nella prima, ogni nodo viene assegnato provvisoriamente alla comunità dei propri vicini qualora ciò comporti un aumento della modularità; nella seconda, le comunità appena formate vengono compattate in “super-nodi” e il processo viene ripetuto. Questa dinamica permette di ottenere una struttura comunitaria multilivello, in cui i gruppi individuati emergono progressivamente attraverso successive aggregazioni.

L'approccio *greedy modularity*, invece, segue una logica più semplice e diretta. Il metodo parte trattando ogni nodo come una comunità autonoma e procede combinando tra loro le coppie di gruppi la cui fusione produce il maggiore incremento di modularità. Il processo continua finché non è più possibile migliorare tale valore, generando così una partizione che massimizza la qualità della divisione secondo una strategia puramente incrementale.

4.5.1 Risultati con algoritmo Louvain

L'applicazione del metodo *Louvain* ha prodotto un totale di 7 comunità, Figura 4.9, evidenziando una suddivisione relativamente compatta della rete.

```
Total number of communities detected: 7

Top 5 largest communities by size:
Community_ID
2      94
6      86
0      41
3      32
4      29
```

Figura 4.9: Comunità rilevate tramite algoritmo di Louvain

Considerata la distribuzione delle dimensioni dei gruppi individuati, l'analisi successiva si concentra esclusivamente sulle prime due comunità. La scelta è motivata dal fatto che tali comunità rappresentano i nuclei più consistenti della rete e racchiudono la maggior parte delle relazioni significative, mentre i gruppi rimanenti presentano dimensioni molto più ridotte e risultano meno rilevanti ai fini interpretativi.

Per procedere all'interpretazione delle comunità individuate, ciascun gruppo è stato rappresentato graficamente, Figura 4.10 e Figura 4.11. Tuttavia, la loro elevata numerosità e densità rendeva difficile una lettura immediata della struttura interna.

Community 2 (size: 94)

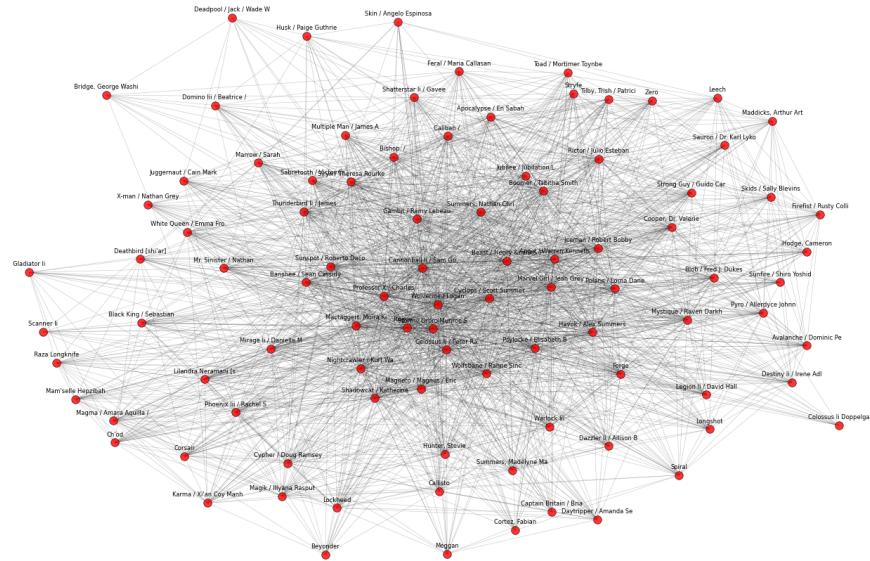


Figura 4.10: Rappresentazione grafica della comunità X-Men

Community 6 (size: 86)

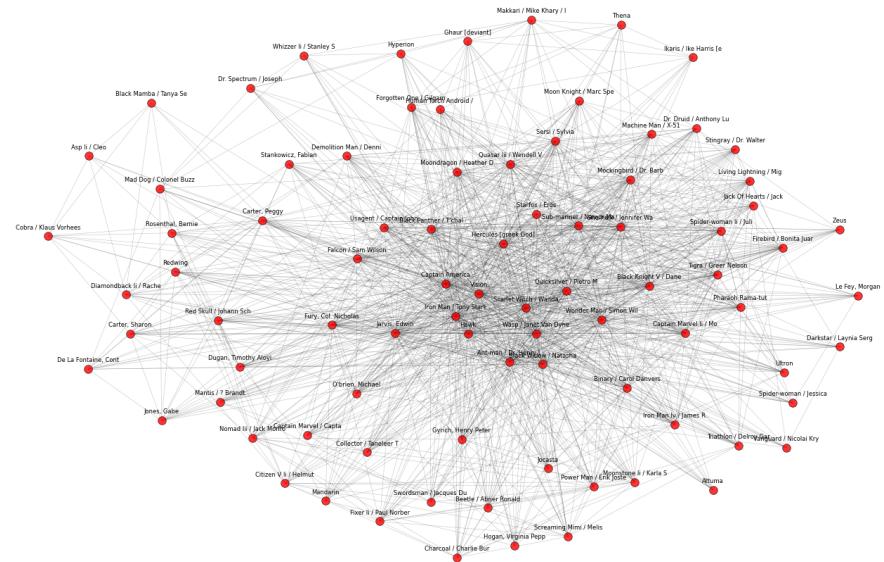


Figura 4.11: Rappresentazione grafica della comunità Avengers

Per ottenere una visualizzazione più chiara, è stata quindi generata una versione semplificata dei grafi, Figura 4.13 e Figura 4.14, includendo esclusivamente i 25 nodi con centralità più elevata all'interno di ogni comunità. Questa scelta consente di mettere in evidenza i personaggi più influenti e le connessioni strutturalmente più rilevanti, riducendo al contempo il rumore visivo prodotto dai nodi periferici.

Grazie alla combinazione tra la rappresentazione grafica e l'analisi dei valori di *betweenness centrality*, Figura 4.12, è stato possibile dedurre la natura delle comunità senza assumere a priori alcuna struttura narrativa. L'osservazione delle connessioni interne e delle posizioni centrali dei personaggi ha infatti messo in evidenza la presenza di gruppi tematici distinti. Nel primo caso, la concentrazione di personaggi quali Beast, Wolverine e Cyclops, unita alla loro elevata centralità, ha suggerito che la comunità fosse riconducibile al nucleo degli X-Men. Nel secondo caso, l'emergere di figure altamente centrali come Captain America e Iron Man ha permesso di identificare la comunità come gravitante attorno agli Avengers.

Character	Community_ID	Betweenness_Centrality
Beast ↗ Henry & Hank P	2	0.292841
Wolverine ↗ Logan	2	0.186184
Cyclops ↗ Scott Summer	2	0.157628
Cannonball ↗ Ii ↗ Sam Gu	2	0.058575
Marvel Girl ↗ Jean Grey	2	0.038900
Captain America	6	0.531721
Iron Man ↗ Tony Stark	6	0.055177
Hawk	6	0.055064
Scarlet Witch ↗ Wanda	6	0.040453
Vision	6	0.025512

Figura 4.12: Tabella con i primi 5 personaggi per centralità di ogni comunità

Questa classificazione, derivata unicamente dai dati e dalle strutture osservate, conferma la capacità dei metodi di rilevazione delle comunità di far emergere spontaneamente le principali suddivisioni tematiche dell'universo Marvel.

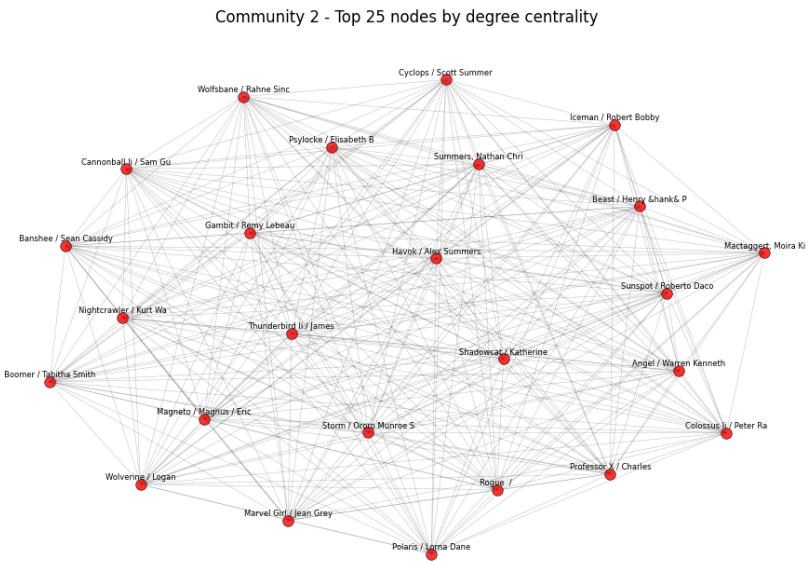


Figura 4.13: Rappresentazione grafica della comunità X-Men, con i top 25 nodi per centralità

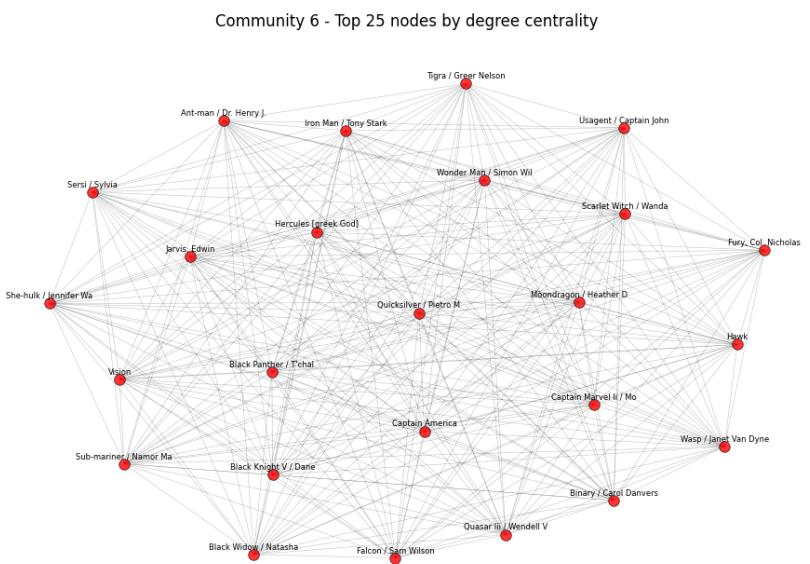


Figura 4.14: Rappresentazione grafica della comunità Avengers, con i top 25 nodi per centralità

4.5.2 Risultati con metodo Greedy Modularity

Utilizzando il metodo *greedy modularity* è stato rilevato lo stesso numero complessivo di comunità, pari a 7 (Figura 4.15).

Total number of communities detected: 7

Top 5 largest communities by size:

Community_ID

0	101
1	94
2	37
3	32
4	22

Figura 4.15: Comunità rilevate tramite metodo Greedy Modularity

In questo caso, è stato applicato lo stesso procedimento adottato per l’analisi basata sul metodo *Louvain*, limitando l’attenzione alle due comunità più estese e rappresentandole graficamente nella loro forma completa, Figura 4.16 e Figura 4.17, e poi ridotta ai 25 nodi più centrali, Figura 4.18 e Figura 4.19.

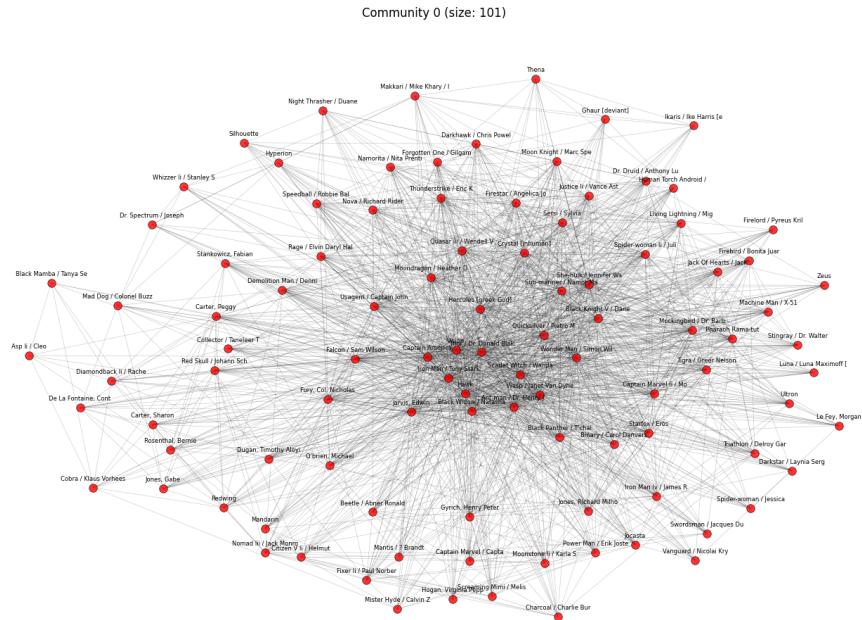


Figura 4.16: Rappresentazione grafica della comunità Avengers

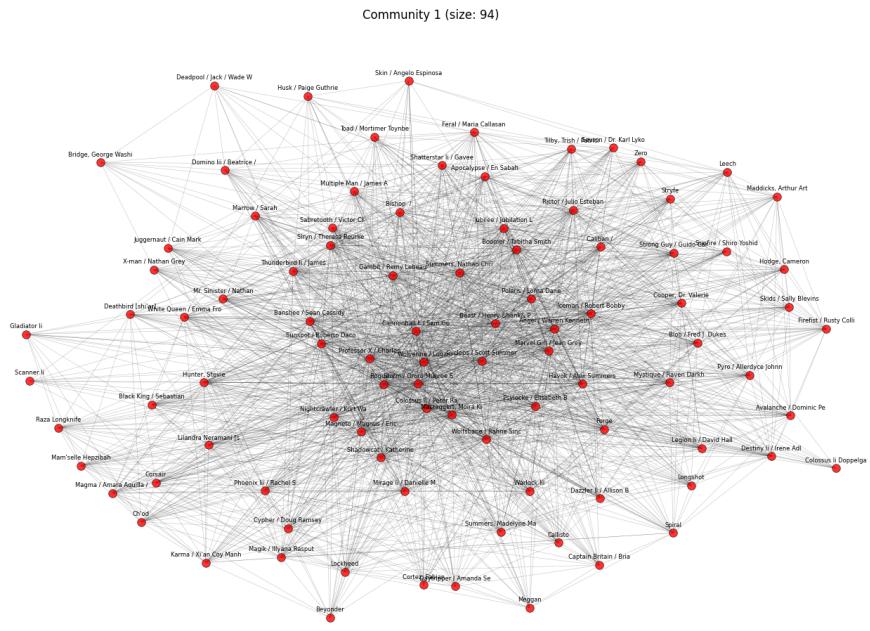


Figura 4.17: Rappresentazione grafica della comunità X-Men

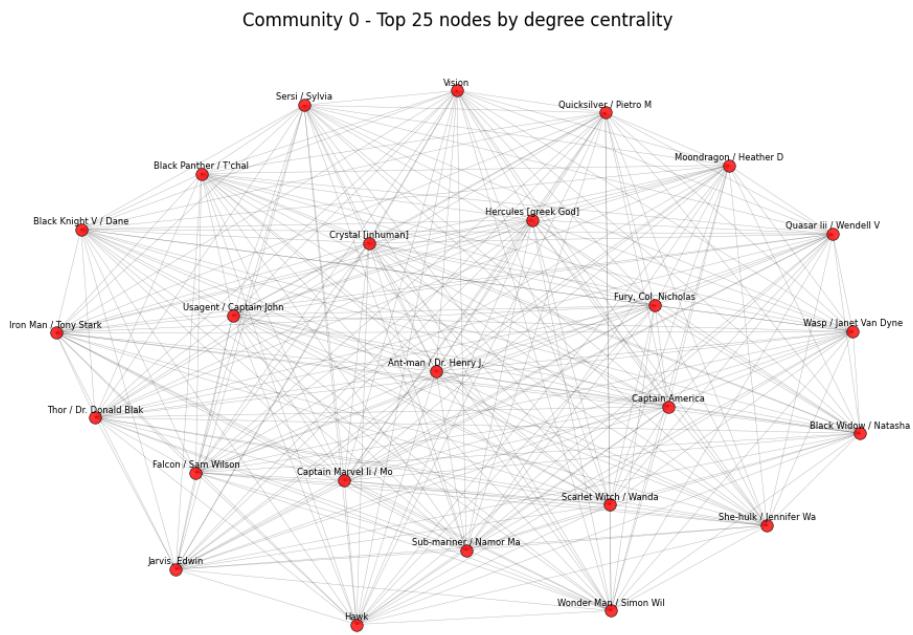


Figura 4.18: Rappresentazione grafica della comunità Avenger, con i top 25 nodi per centralità

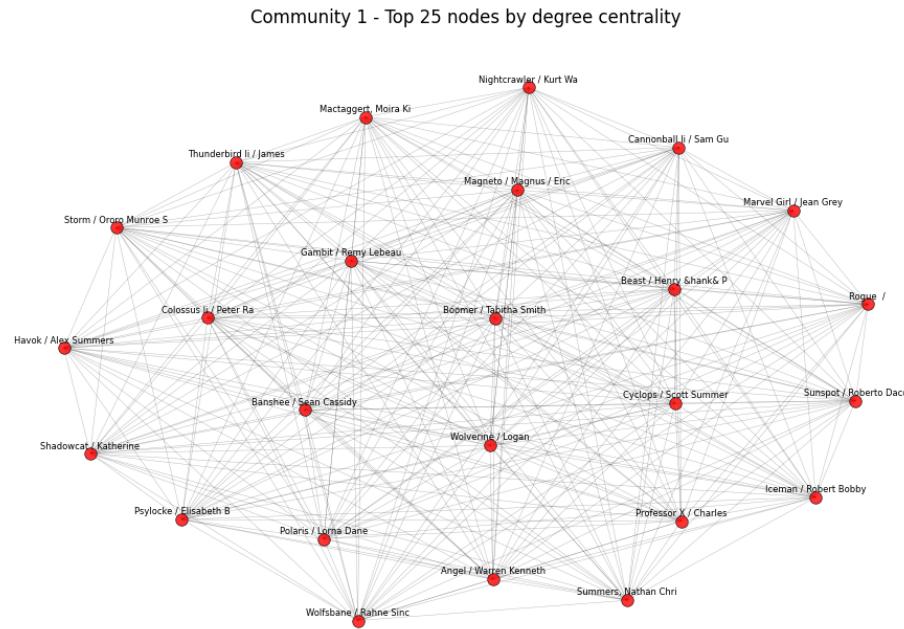


Figura 4.19: Rappresentazione grafica della comunità X-Men, con i top 25 nodi per centralità

Le differenze principali, rispetto al metodo precedente, riguardano la distribuzione delle dimensioni: con questo approccio, la comunità più ampia risulta essere quella con `id` pari a 0, riconducibile agli Avengers, sebbene con un margine relativamente contenuto rispetto alla successiva. Anche in questo scenario, tuttavia, le due comunità predominanti coincidono con i gruppi già osservati in precedenza, ovvero il blocco degli Avengers e quello degli X-Men, a conferma della stabilità strutturale di tali aggregazioni indipendentemente dal metodo utilizzato per la rilevazione.

--- Top 5 Personaggi per le due Comunità più Grandi (Greedy Modularity) ---			
Character	Community_ID	Betweenness_1/Peso	
Captain America	0	0.531721	
Thor ↗ Dr. Donald Blak	0	0.121057	
Iron Man ↗ Tony Stark	0	0.055177	
Hawk	0	0.055064	
Scarlet Witch ↗ Wanda	0	0.040453	
Beast ↗ Henry & Hank & P	1	0.292841	
Wolverine ↗ Logan	1	0.186184	
Cyclops ↗ Scott Summer	1	0.157628	
Cannonball II ↗ Sam Gu	1	0.058575	
Marvel Girl ↗ Jean Grey	1	0.038900	

Figura 4.20: Tabella con i primi 5 personaggi per centralità di ogni comunità

Come analisi finale, per ottenere una visione d'insieme della struttura della rete nell'universo Marvel, si è deciso di rappresentare il grafo suddiviso in comunità, evidenziando le principali aggregazioni di personaggi.

Grafo Marvel con tutte le comunità (Avengers e X-Men evidenziati)

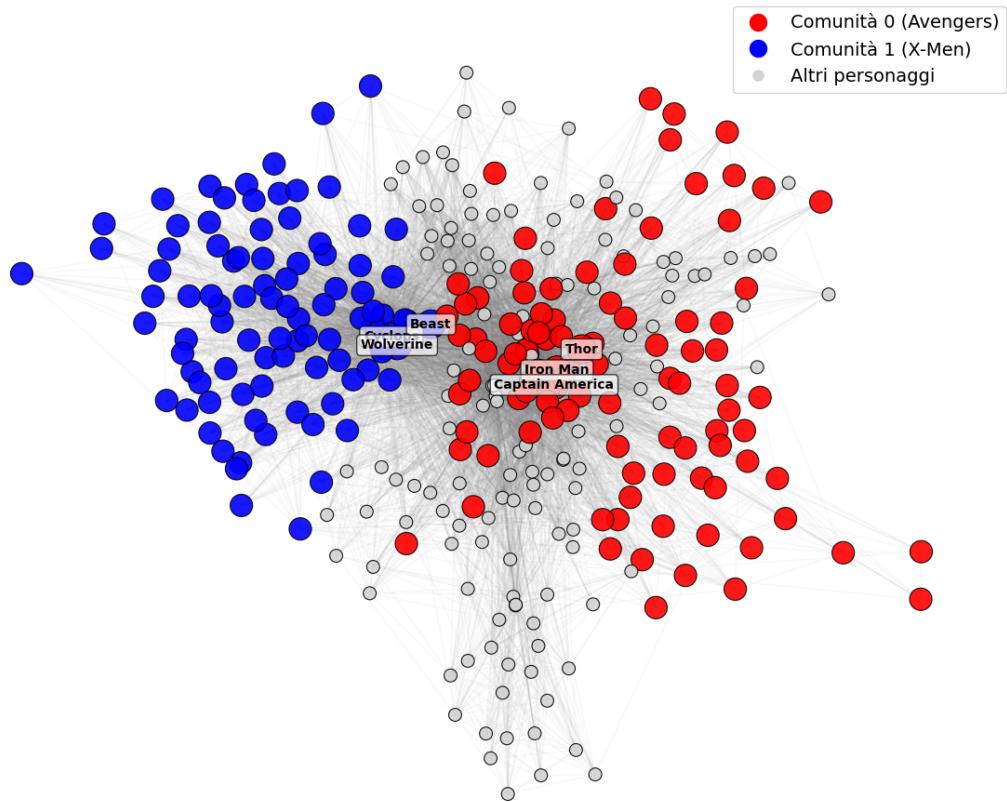


Figura 4.21

La Figura 4.21 mostra l'intero grafo dei personaggi Marvel, in cui sono rappresentate tutte le comunità individuate dall'algoritmo, mettendo però in evidenza soltanto i nodi appartenenti alle comunità degli Avengers e degli X-Men. I personaggi associati agli Avengers sono colorati in rosso, quelli legati agli X-Men in blu, mentre tutti gli altri nodi delle restanti comunità sono mostrati in grigio sullo sfondo, così da evidenziare il ruolo centrale e la posizione relativa dei due team principali rispetto al resto dell'universo narrativo.

In conclusione, questa rappresentazione dell'intero grafo mostra come le comunità degli Avengers e degli X-Men occupino posizioni centrali rispetto al resto dell'universo Marvel, fungendo da veri poli di aggregazione per molti altri personaggi. La presenza di numerosi collegamenti tra i nodi grigi e quelli rossi o blu evidenzia inoltre il ruolo

cruciale di questi due team nel tenere connessi i diversi filoni narrativi, confermando la natura fortemente integrata e corale del Marvel Universe.