Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo'

Alessandra Failla, Digital Humanities and Digital Knowledge, 0001004725

1. Introduction

This study deals with the network analysis of Alexandre Dumas' novel 'The Count of Monte Cristo', published between 1844 and 1846 in serialized format. The novel, spanning 1815 to 1839, follows the life of Edmond Dantès, a young sailor falsely accused of treason on the eve of his wedding and imprisoned in the Château d'If for fourteen years. During his incarceration, he meets a fellow prisoner who reveals the location of a treasure, setting Edmond on a path of vengeance against the characters responsible for his misfortune.

The network analysis presented in this report is based on the interactions among characters within the novel. Network analysis applied to literary works represents a powerful tool to analyze the role of characters within a novel through character interactions.

After discussing problems and motivations, the study addresses data gathering and dataset preparation. Particular focus is placed on the definition of interactions between characters to develop a network based on characters' interactions using Python 3.10.9 (Van Rossum & Drake Jr, Python reference manual, 1995).

After addressing the validity and reliability of the results, graph visualizations are provided using Gephi (Bastian, Heymann, & Jacomy, 2009), followed by computation and comparison of centrality measures. The final sections address the obtained results in the light of the initial research questions, with suggestions on the possible alternative methodologies.

2. Problem and Motivation

This report explores 'The Count of Monte Cristo' core themes of revenge and redemption by observing how they are reflected in the character networks. To this extent, it aims to show how the social network of Dumas' novel favors Dantès' plan of revenge, uncovering the most relevant characters and the evolution of their centrality compared to their relationship with Edmond throughout the novel. In particular, the analysis focuses on the evolution of the centrality of Dantès' enemies and the evolution of the relationship with the protagonist across the volumes. Dantès' enemies include Fernand, a former friend of his fiancée who betrayed him for personal gain; Danglars, a scheming and untrustworthy colleague who orchestrated his downfall; Gerard de Villefort, a corrupt prosecutor who wrongfully imprisoned him; and Caderousse, Dantès coward neighbor that was present during the conspiracy but did nothing to avoid Dantès arrest.

Throughout the novel, Dantès carries out his revenge plan against these characters by manipulating the fabric of the Parisian society for his interests. His actions lead Fernand to commit suicide in the fourth volume. Danglars, who is strongly attached to wealth, faces financial ruin caused by the protagonist. Villefort's reputation crumbles as his dark secrets are exposed, leading to tragic consequences for his family. Meanwhile, Caderousse falls victim to his greed, is manipulated by Dantès and ultimately killed by one of his affiliates.

Understanding character relationships contributes to a deeper comprehension of the novel, offering insights into themes, plot development, and character motivations. Dynamic network analysis is conducted to unveil patterns and information that could be overlooked or skewed by static network analysis methods. Moreover, dynamic analysis enables capturing the temporal dimension of the novel, facilitating the observation of evolving relationships among characters (Agarwal, Corvalan, Jensen, & Rainbow, 2012). The findings of the study on the evolution of character connections will show that the novel explores beyond the pervasive theme of revenge to delve into themes of redemption as well.

3. Datasets

The data required for the analysis was extracted from the text of the novel 'The Count of Monte Cristo'. The complete text of the novel was made available by Project Gutenberg in its 1998 version and is freely accessible to the public (Dumas & Maquet, 1998). To ensure the accuracy of results, chapters were organized into segments ranging from 5 to 15 chapters, consistent with the division into five volumes as presented by the data source.
Subsequent steps involved text cleaning to remove headings, indexes, and other non-essential elements, and converting the text into Python string format.
A comprehensive list of 46 characters was compiled based on the novel's Wikipedia page (Wikipedia contributors, s.d.), supplemented by a list of all name variations for each character encompassing first names, surnames, nicknames, and honorifics, to ensure precise character identification. Special focus was dedicated to the main character, Edmond Dantès, also known as the Count of Monte Cristo, who adopts multiple identities throughout the narrative. A manual validation step was performed to ensure that all variations were encompassed.
A set of regular expressions (RE) was developed from the name lists, using the 're' Python module for RE operations (Van Rossum, The Python Library Reference, release 3.8.2, 2020). Subsequently, the named entities were replaced with unique identifiers for each character. Four separate RE files were employed to minimize potential errors in name recognition, each developed with minor variations corresponding to specific sections of the novel.
The Python library Natural Language Toolkit (nltk) (Bird, Klein, & Loper, 2009) was used to perform text tokenization, with punctuation removed from the token list. This step was essential, as this network analysis relies on the detection of character interactions based on the distance of character mentions in the text. This methodology, commonly employed in literary network analysis, was preferred over identifying dialogues and direct communication, a strategy more suitable for plays. Novels often feature interactions that do not exclusively involve verbal communication (Coll Ardanuy & Sporleder, 2014). Moreover, the manual annotation of interactions within the text is prone to human errors and would reduce the reliability of the work.
The approach to define the narrative window for detecting interactions employed by Grayson, Wade, Meaney, & Greene (2016) was adopted, exploring different window sizes obtained through collinear, coplanar, and combined strategies. The collinear approach identifies consecutive character mentions linearly, recording interactions between characters appearing sequentially. In contrast, the coplanar method considers all character mentions within a designated window of text, even if they are not consecutive, resulting in a more expansive recording of interactions.
The process of defining the narrative window using the collinear method consisted of the construction of an interaction network for each window size ranging from 10 to 500 words for every volume of the novel. These networks were created using the Python library NetworkX (Hagberg, Swart, & S Chult) based on interactions dictionaries. These dictionaries were formed by identifying character co-occurrences within each window size and consist of character IDs as keys and nested dictionaries containing interactions with other characters as values. Character pairs with less than three interactions were excluded to prioritize more relevant character relationships and avoid the inclusion of non-existent interactions. The edge density of each network for each window size was computed and the resulting values were

plotted. The resulting curve was analyzed to determine the point at which it flattened, indicating the optimal window size. As the methodology to obtain the plateau value for each volume was not explicitly detailed in the dedicated paper section by Grayson, Wade, Meaney, & Greene (2016), the adopted approach was developed autonomously. The plateau point was defined as the start of a set minimum number of iterations that do not present variations in edge density. The chosen value for this condition was 100, based on empirical observation. Coplanar connections require a different approach to derive narrative window sizes due to the continuous increase in edge density with window size increments. Instead, the number of tokens between characters, referred to as "gaps", is examined, treating these gaps as boundaries of character interaction occurrences to generate window sizes based on statistically derived upper limits. Specifically, the interquartile range (ICR = $Q_3 - Q_1$) is used to define the probable upper limits, with any elements outside these limits considered as suspected outliers and left out (Grayson, Wade, Meaney, & Greene, 2016). Table 1 below shows window sizes obtained through collinear and coplanar approaches on the whole novel as well as on the single volumes.

Input text	Collinear window	Coplanar window (Q_3)
Full text	385	183
Volume 1	186	141.5
Volume 2	306	178
Volume 3	298	231
Volume 4	196	175.5
Volume 5	147	163
Average value (vol. 1-5)	**226.6**	**177.8**
Discrete value	**227**	**178**

Table 1 Window sizes from collinear and coplanar approaches on the whole novel and single volumes.

The data suggests that the collinear window sizes tend to be larger than the coplanar window sizes. This trend is observed both for values related to the entire novel, as well as across the volumes individually. The collinear approach results in an average window size of 226.6, while the coplanar approach yields an average window size of 177.8.

4. Validity and Reliability

The English version of 'The Count of Monte Cristo' used for this analysis is readily accessible on the Project Gutenberg website and enables replication of the presented study. All material developed during the study has been made publicly available via a dedicated GitHub repository (Failla, Graphing Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo', 2024).

To ensure the validity of results, a consistent approach for detecting character mentions was designed. A consistent window size was adopted to identify interactions and construct character networks, employing a distance-based interaction modeling approach. To avoid omissions, RE were carefully designed and refined to address name ambiguity. Additionally, the novel was segmented into multiple files based on character appearances, considering its narrative spanning several years and featuring characters from different generations within the same family.

Reliability of results is inherent in the study's design, as only automated modifications were applied to the text, reproducible using the developed RE files, which ensure consistent retrieval of character mentions. The results of this study are reproducible through the application of the procedure outlined in the dedicated file (Failla, Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo', 2024).

5. Measures and Results

After evaluating the results obtained through collinear and coplanar strategies, a window size of $Q_3 = 178$ was adopted for this analysis, representing the average of the coplanar window sizes across the five volumes. This approach was considered more suitable for finding interactions within groups, which happen frequently in 'The Count of Monte Cristo', generally in the context of the salons of Parisian society. Through the application of a collinear method several interactions would be overlooked or skewed. Results of the collinear strategy were instead used to integrate the results to account for possible missing interactions detected within a larger window size. The collinear window size $w = 227$ was selected for the integration, representing the mean of the collinear window sizes from each volume (Table 1).

Gephi was used to analyze and edit the resulting five graphs, available in the dedicated GitHub repository (Failla, Graphing Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo', 2024). The character interactions networks for each volume are represented as undirected graphs; characters are depicted as nodes and their interactions as edges. The graphs are unweighted since interaction weights were not considered in this study.

For the scope of this study, *k*-core decomposition was applied to extract the most relevant characters from the graphs developed for each volume, minimizing the risk of spurious inclusion of interactions. The *k*-core of a graph is the maximal subgraph that contains nodes with a minimum degree of *k* (Bickle, 2010); *k*-core detection was performed using NetworkX without specifying the degree of the core so that the core with the largest degree was obtained (k_core, s.d.). Within each *k*-core graph, communities were detected to uncover possible clusters of characters who share stronger and more frequent interactions, each representing a subset of nodes within the graph such that connections between the nodes are denser than connections with the rest of the network (Radicchi, Castellano, Cecconi, Loreto, & Parisi, 2004). The project focused on the evolution of graph density for each *k*-core graph to gain an insight on the intensity of character relationships in the network. Graph density represents the ratio between the number of edges and the number of nodes. The number of edges and their variations across volumes were also compared to obtain information on character relationships based on interactions.

Degree centrality and eigenvector centrality were computed and stored using Python libraries NetworkX and pandas (McKinney & others, 2010) and compared to assess character relevance. The degree centrality of a node in an undirected unweighted graph is the count of edges connected to it (Powell & Hopkins, 2015). Values of degree centrality were normalized by dividing each degree centrality by the maximum possible degree in a simple graph of $n-1$ nodes, where n is the number of nodes in the graph (degree_centrality, s.d.).

Eigenvector centrality measures a node's importance while considering the importance of its neighbors and is often used to measure the influence of a node within the overall network (Golbeck, 2013). This centrality measure was chosen due to its ability to summarize both direct connections and indirect influence of a node within the network; it considers the overall network structure and enables the identification of characters with widespread influence, crucial to understanding the protagonist's manipulation of the novel's social landscape.

The dynamic analysis of *k*-cores across the five volumes revealed shifts in character relevance over narrative progression. The number of nodes within the k-cores subgraphs increases from 9 nodes in the first volume to a peak in the fourth volume with 20 nodes. The number of edges follows a similar pattern, with 35 in Volume 1 and peaking with 145 in Volume 4. While most characters from Volume 1 also are relevant characters in the subsequent volumes, new characters of notable relevance are introduced starting from Volume 2.

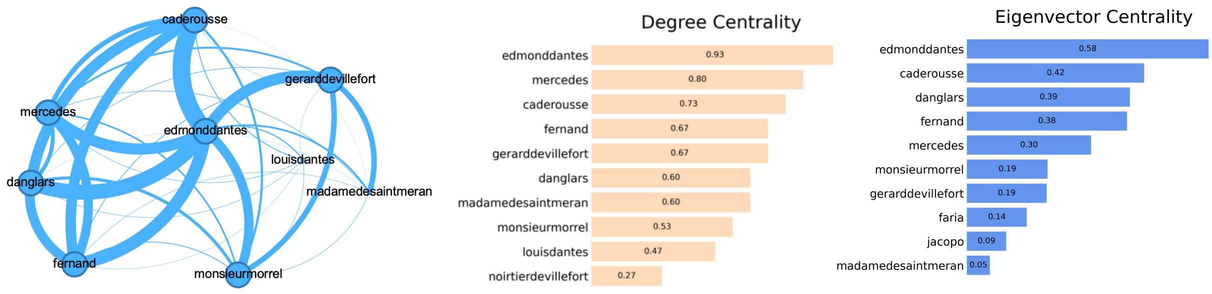

Figure 1 Volume 1: *k*-core subgraph and centrality measures.

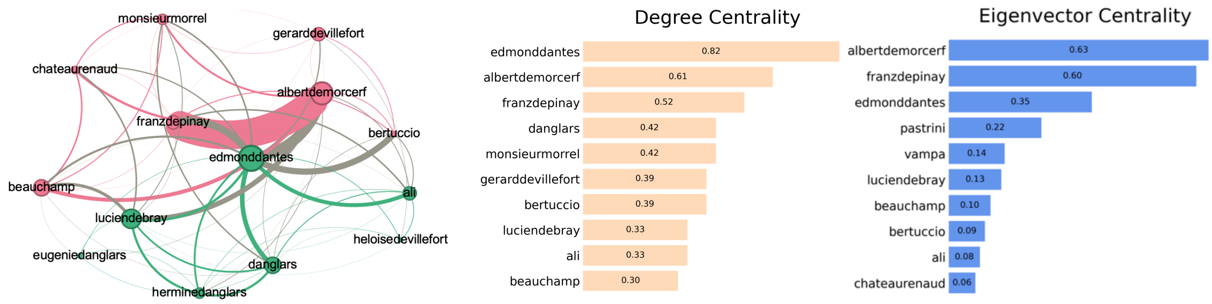

Figure 2 Volume 2: *k*-core subgraph and centrality measures.

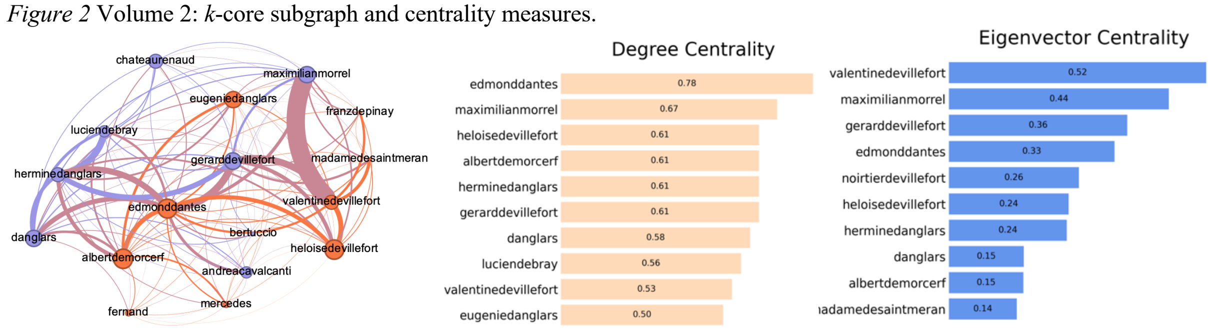

Figure 3 Volume 3: *k*-core subgraph and centrality measures.

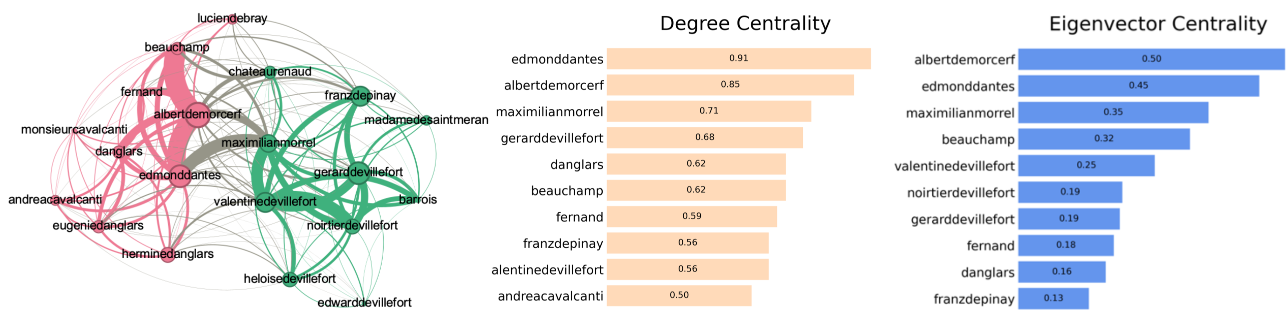

Figure 4 Volume 4: k-core subgraph and centrality measures.

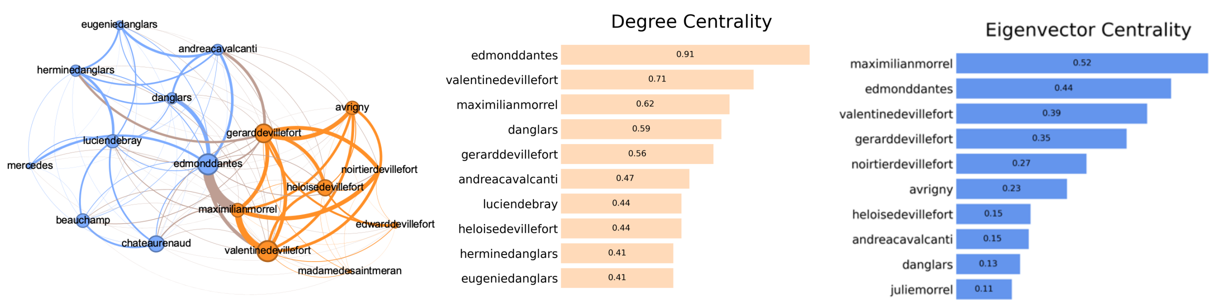

Figure 5 Volume 5: *k*-core subgraph and centrality measures.

Communities within the *k*-core networks were detected and highlighted with different colors using Gephi; the graphs also show the degree of each character through their node size (Figure 1 to Figure 5 above). While the first graph does not present separate groups, the remaining graphs consistently present two communities, with a progressive intensification of connections. Community members are not consistent throughout the novel and each community does not strictly identify single-family units, as these are highly interconnected. Dantès keeps significant relationships with all communities throughout the novel.
By comparing variation of graph density and number of edges in Table 2 below, we can gain an insight of the evolution of the dynamics between characters throughout the novel.

Volume	Density	N. of edges	Change in n. of edges (%)	Change in density (%)
1	0.972	35	*Not applicable*	*Not applicable*
2	0.681	62	77.14	-29.88
3	0.838	114	83.87	22.90
4	0.763	145	27.19	-8.98
5	0.779	106	-26.90	2.21

Table 2 Graph density and number of edges and their variations throughout the novel

Volume 1 shows a high graph density of 0.972, with 35 edges, indicating a densely connected network. Volume 2 shows a decreased graph density of 0.681 and 62 edges, meaning a wider but less dense network compared to Volume 1. Volume 3 shows increased interactions, with a higher graph density of 0.838 and 114 edges. Similarly, Volume 4 maintains relatively high graph density at 0.763, with 145 edges, suggesting a significant number of interactions at a slightly lower density. Even in Volume 5, with a smaller number of edges (106), the high graph density of 0.779 indicates significant character interactions.
Edmond Dantès remains one of the most central and interconnected characters throughout the whole novel, as he is one of the few characters that appear consistently in every *k*-core graph. This is confirmed by his highest normalized degree centrality across all volumes, ranging from 0.78 to 0.93. Besides Dantès, 'The Count of Monte Cristo' features several characters with a high degree centrality. Figure 6 below tracks the degree centralities of Dantès enemies – Fernand, Danglars, Villefort, and Caderousse – across the five volumes; notably, they show a high value in the first volume, followed by a significant decline in the second volume.

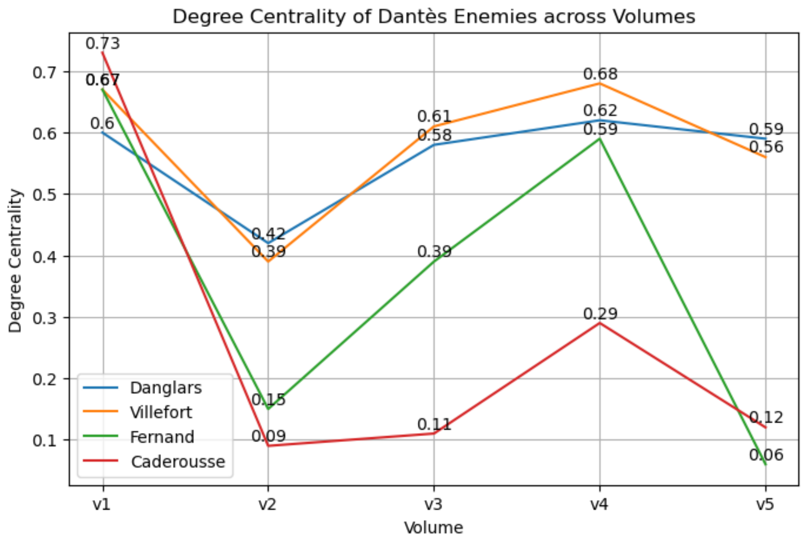

Figure 6 Degree centralities of Dantès' enemies across the volumes.

Caderousse holds a degree centrality of 0.73 in Volume 1 (Figure 1), however, in subsequent volumes he loses prominence in degree centrality, until a new increase in the fourth volume, where his degree centrality slightly increases, reaching 0.29 (Figure 2). Villefort, Fernand, and Danglars follow a similar trend but with higher degree centralities in the central volumes.

Their values rise in the third and fourth volumes, with a slight decrease in the fifth volume for Danglars and Villefort, while Fernand experiences a dramatic drop. Villefort and Danglars consistently maintain a high degree centrality, frequently ranking among the top ten throughout the novel, whereas Fernand appears less frequently.
Other notable characters include Albert the Morcerf, son of Fernand, who shows one of the highest degree centralities, ranging from 0.61 to 0.85 from Volume 2 to Volume 4, and presents the highest eigenvector centrality in Volumes 2 and 4, respectively 0.63 and 0.50 (Figure 2 and Figure 4).
By observing centrality values of the last three volumes, Maximilian Morrel shows high degree centrality values ranging from 0.62 to 0.71 (Figure 3 to Figure 5), and fluctuates among the third, second, and first position in the rankings for eigenvector centrality, with values oscillating between 0.35 and 0.44 (Figure 3 to Figure 5). Valentine de Villefort also shows significant variations in centrality, as her degree centrality increases from 0.53 in Volume 3 to 0.71 in Volume 5, ranking second after Edmond Dantès (0.91), and followed by Maximilian Morrel (0.62), as shown in Figure 5.
Table 3 below shows Dantès' interactions with his enemies and their families starting from Volume 2, which marks the start of his revenge plan. Examining the number of edges between character pairs reveals that Dantès shares a significant number of edges with his enemies, such as 747 with Danglars and 512 with Villefort. The data also illustrates the significant connections between Dantès and his enemies' families, with notable totals such as 2104 with the de Morcerf family and 1466 with the Danglars family. Moreover, Dantès shares a total of 1333 edges with Albert de Morcerf. It is noticeable that Edmond and Albert share 699 interactions in Volume 4. It is possible to put these results into perspective by comparing them with the data in Table 4 below, representing character pairs sharing the highest number of interactions across Volumes 2 to 5.

Source	Target	Edges Vol. 2	Edges Vol. 3	Edges Vol. 4	Edges Vol. 5	Total
Edmond Dantès	Fernand de Morcerf	19*	19	205	8	251
Edmond Dantès	Mercedes de Morcerf	19	59	337*	105	520
Edmond Dantès	Albert de Morcerf	392	222	699	20	1333
Edmond Dantès	**De Morcerf Family**	**430**	**300**	**1241**	**133**	**2104**
Edmond Dantès	Baron Danglars	169	232	167	179	747
Edmond Dantès	Hermine Danglars	68	290	103	38	499
Edmond Dantès	Eugenie Danglars	14	72	106	28	220
Edmond Dantès	**Danglars Family**	**251**	**594**	**376**	**245**	**1466**
Edmond Dantès	Gerard de Villefort	43	346	53	70	512
Edmond Dantès	Heloise de Villefort	22	192	5	24	243
Edmond Dantès	Valentine de Villefort	0	104	48	159	311
Edmond Dantès	**De Villefort Family**	**65**	**642**	**106**	**253**	**1066**
Edmond Dantès	Caderousse	61	0	206	6	273
Edmond Dantès	La Carconte	21	0	7	0	28
Edmond Dantès	**Caderousse Family**	**82**	**0**	**213**	**6**	**301**
* Does not appear in the volume's *k*-core subgraph						

Table 3 Number of edges between Edmond Dantès and his enemies, and with two additional family members of each enemy across Volumes 2-5.

Source	Target	Edges Vol. 2
Edmond Dantès	Albert de Morcerf	392
Edmond Dantès	Franz d'Epinay	267
Edmond Dantès	Bertuccio	229
Albert de Morcerf	Lucien Debray	189
Edmond Dantès	Danglars	169

Source	Target	Edges Vol. 3
Valentine de Villefort	Maximilian Morrel	884
Edmond Dantès	Gerard de Villefort	346
Hermine Danglars	Lucien Debray	329
Valentine de Villefort	Gerard de Villefort	324
Edmond Dantès	Hermine Danglars	290

Source	Target	Edges Vol. 4
Albert de Morcerf	Beauchamp	774
Edmond Dantès	Albert de Morcerf	699
Valentine de Villefort	Noirtier de Villefort	485
Edmond Dantès	Maximilian Morrel	481
Valentine de Villefort	Maximilian Morrel	439

Source	Target	Edges Vol. 5
Edmond Dantès	Maximilian Morrel	465
Valentine de Villefort	Maximilian Morrel	291
Noirtier de Villefort	Maximilian Morrel	188
Gerard de Villefort	Avrigny	187
Edmond Dantès	Danglars	179

Table 4 Top 5 of the highest number of interactions across Volumes 2, 3, 4, and 5, represented through the number of edges.

Albert de Morcerf shows the highest number of interactions with Dantès in Volume 2 (392) and the second highest number in Volume 4 (699) but does not show a significant number of interactions in the third and fifth volumes. In Volumes 3 and 4, Dantès shows the highest number of interactions with Villefort (346) and Maximilian Morrel (465). It is noteworthy that the only character that appears consistently among the five most frequent interactions across the considered volumes is Edmond Dantès. By further examining the data it is possible to notice that Maximilian Morrel and Valentine de Villefort are the only secondary characters that appear consistently among the most frequent interactions in at least three of the four considered volumes. In Volume 3, they share the highest number of interactions (884). In the fourth volume, Valentine shares her highest number of interactions with his grandfather Noirtier (485), followed by her interactions with Maximilian (439), while Maximilian and Edmond present 481 interactions. Maximilian's interactions with Edmond decrease to 465 in Volume 5, which however represents the highest number of interactions in this volume. This significantly increasing number of interactions among these three characters aligns with their increasing centrality across the last three volumes, presented in Figure 3 to Figure 5.

6. Conclusion

This study explores how the themes of revenge and redemption are reflected in the social network of 'The Count of Monte Cristo'. The construction of *k*-cores graphs for the five volumes facilitated the identification of the most relevant characters. Figure 1 to Figure 5 showcased a gradual increase in the number of nodes, peaking at 20 nodes in Volume 4. This trend mirrors the novel's narrative arc, where Dantès, after enduring fourteen years of imprisonment, dedicates the second and third volumes to building his social network in preparation for revenge, facilitated by the treasure entrusted to him by his fellow prisoner Abbè Faria. Despite his role, Faria's absence from the k-core subgraphs is notable due to his sole significant connection with Dantès during their years in prison. In Volumes 3 and 4, Dantès' revenge unfolds, also revealing unexpected consequences that reshape the course of the narrative. In Volume 5, he is forced to confront the true nature of his pursuit and its impact on those around him.

Volume 4 represents a narrative climax as the number of interactions intensifies significantly (Figure 4). The high graph densities related to the central volumes provide further proof of the increasing dense and interconnected relationships in Volumes 3 and 4, mirroring the repercussions of Dantès' plans on the Parisian elite families (Table 2). Moreover, the identification of highly interconnected and dense communities within the *k*-graphs of the central volumes revealed a tightly interconnected society, providing a fertile ground for Edmond's schemes, as he infiltrates the tightly knit Parisian elite and builds a robust network of relationships. The graphs show that the families of Edmond Dantès' enemies are highly

interconnected and often part of the same community. Dantès is also tightly connected to his enemies and their relatives, showing how his thirst for revenge drives him to cultivate relationships with his adversaries and characters close to them, taking advantage of their societal and familial ties for his plan.
The prominence of Dantès' enemies' degree centralities throughout the novel (Figure 1 to Figure 5) further highlights the importance of their role and thereby of the role of revenge in the narrative. Their high degree and eigenvector centralities in Volume 1 highlight their pivotal role in the earliest events of the novel, where their actions led to Dantès' wrongful imprisonment. Figure 6 shows that Dantès' enemies' degree centralities follow the same pattern and culminate in Volume 4, as they face the consequences of the Count's revenge.
Villefort and Danglars maintain higher values throughout the volumes, followed by Fernand and Caderousse; his trajectory diverges from the others as his story develops outside the Parisian high society, as evidenced by his absence from the *k*-core graphs from Volume 2 onwards. Figure 6 also shows a dramatical drop in Fernand's degree centrality between Volume 4 and Volume 5, reflecting the event of his suicide after his incrimination and consequent abandonment by his family. The notable reduction of Albert de Morcerf's centrality and interactions in the last volume is further proof of this event.
Albert's high centrality measures show his pivotal role in the novel; his high eigenvector centrality in Volumes 2 and 4 provides information on his pervasive influence on the overall network (Figure 1 and Figure 4). He establishes a strong bond with Edmond throughout the novel, evidenced by their high number of interaction (Table 3). Dantès connection with this influential character is thus crucial for his vengeance since it allows him to manipulate the novel's social dynamics. The data in Table 3 also shows that Dantès effectively establishes strong relationships with several secondary characters linked to his enemies. This emphasizes his propensity to form bonds primarily with characters instrumental to his goals.
The results present Maximilian Morrel and Valentine de Villefort as further central figures in the last three volumes of the novel. As shown in Figure 3 to Figure 5, their centralities as well as the number of interactions between them and with other characters increase across the last three volumes (Table 4). While Maximilian Morrel is absent from previous volumes' *k*-core networks, it is noticeable that his father, Monsieur Morrel, is a relevant character in Volumes 1 and 2. He is Dantès benefactor from earlier volumes, who tried to rescue the protagonist from imprisonment. Volumes 4 and 5 see Edmond Dantès forming a notable bond with Maximilian Morrel, as evidenced by the high number of interactions between them (Table 4). This relationship is notable since most of Dantès' relationships serve his grand scheme for vengeance, yet Maximilian Morrel emerges as a crucial figure in the last stages of the novel, despite having no part in Dantès' conspiracy plan. The emergence of one of the few selfless relationships of the protagonist is a signal for the increasing relevance of a further theme beyond revenge: Edmond Dantès' redemption.
In Volume 3, Maximilian Morrel and Valentine de Villefort share the highest number of edges overall in the whole network (Table 4). The high number of interactions between them is evidence of the nature of their relationship: they are clandestine lovers, and their relationship is exposed to readers in the third volume of the novel. Results show that, while Maximilian shares a strong connection with Edmond, Valentine, and Edmond do not present a high number of interactions, as shown in the results of Table 4. This is motivated by the fact that Dantès, unaware of his friend's romantic involvement, was discreetly planning Valentine's death as part of his revenge against Villefort and did not have any significant relation with the corrupt prosecutor's daughter. Upon discovering the truth in Volume 5, however, he intervenes to save Valentine and facilitates her secret reunion with Morrel. The relevance of these three characters and their relationship in the final volume is exhibited by their highest degree centralities compared to other characters (Figure 5); additionally, both Edmond and Valentine

share a high number of edges with Maximilian (Table 4), who appears to have become a pivotal character in this volume. This transformation in character relationships and roles across the volumes mirrors Dantès' changing motivations. Initially driven by revenge and the pursuit of private justice, he later shifts towards redemption and love, as he attempts to rectify the consequences of his earlier vengeful actions, refraining from fully executing his plan.

7. Critique

This study effectively demonstrates the manifestation of the themes of revenge and redemption within the social networks of Dumas' novel. Initially, it illustrates how Dantès strategically plants the seeds of his vendetta in the fertile ground of the interconnected Parisian society, a characteristic that persists throughout the volumes of 'The Count of Monte Cristo'. The analysis investigates further the dynamics between Dantès, his adversaries, and their families, as well as their evolving roles throughout the novel to finally address the increasing relevance of relationships that represent a signal of the protagonist's redemption.

Some alternative centrality measures could have been considered to address certain challenges in the analysis. While eigenvector centrality was used to highlight characters' influence, particularly those closely connected to the Count of Monte Cristo, other measures like betweenness centrality, identifying "bridge" nodes within the network, and closeness centrality, assessing proximity to all nodes and identifying nodes that could influence the entire network most quickly, could have provided additional insights into the network dynamics.

The detection of interactions was based on a RE-based approach to ensure accuracy and avoid errors in graph construction, including in the erroneous inclusion or exclusion of nodes. An exclusively manual process for the retrieval of interactions was avoided to ensure the high reliability of results. An approach based on REs, instead, provides consistent results and ensures the reproducibility of the study. Alternatively, researchers could opt to implement Named Entity Recognition (NER) algorithms to detect character mentions within the text for potentially increased accuracy and efficiency.

A further topic of discussion in literary network analysis is the detection of character interactions. A manual approach based on markup of interactions is a possible option but is prone to human errors, besides being highly time-consuming. A further methodology, adopted in this study, consists of the definition of the narrative window, where characters' co-occurrence constitutes an interaction. Researchers interested in exploring alternative window sizes could adopt different values identified through coplanar, collinear, or combined approaches. However, a mere collinear approach was not considered suitable for the scope of this study due to its potential to skew group interactions, which, however, are at the core of this novel.

While the identification of a narrative window for automatic interaction detection is a viable strategy, it may misinterpret co-occurrence as actual interactions. This study prioritized the analysis of relationships among key characters identified through k-cores isolation to minimize the risk of inaccurate interactions. While this represents a possible method for identifying central characters, it is important to note that it fails to capture the full complexity of character relationships. An alternative approach based on Large Language Models (LLM) to model character relationships, however, could lead to more accurate results, given their ability to capture linguistic patterns and context.

References

Agarwal, A., Corvalan, A., Jensen, J., & Rainbow, O. (2012). Social network analysis of Alice in Wonderland. In D. Elson, A. Kazantseva, R. Mihalcea, & S. Szpak (Eds.), Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature (pp. 88-96). Association for Computational Linguistics.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. Third international AAAI conference on weblogs and social media.

Bickle, A. (2010). The K-Cores of a Graph. Dissertation. Western Michigan University. Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.

Coll Ardanuy, M., & Sporleder, C. (2014). Structure-based clustering of novels. Proceedings of the First Workshop on Computational Models of Narrative (pp. 31-39). Association for Computational Linguistics.

degree_centrality. (n.d.). Retrieved 2024, from NetworkX: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.degree_centrality.html

Dumas, A., & Maquet, A. (1998). The Count of Monte Cristo by Alexandre Dumas and Auguste Maquet. Retrieved 2024, from Urbana, Illinois: Project Gutenberg: https://www.gutenberg.org/ebooks/1184

Failla, A. (2024). Graphing Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo'. Retrieved from GitHub: https://github.com/AlessandraFa/montecristo-network.git

Failla, A. (2024). Graphing Revenge: A Network Analysis of 'The Count of Monte Cristo'. Retrieved from GitHub: https://github.com/AlessandraFa/montecristo-network/blob/e4bef19788fb8fd33647daeebd84d9372ef449c6/Notebook_CountofMonteCristo_NetworkAnalysis.ipynb

Golbeck, J. (2013). Network Structure and Measures. In J. Golbeck, Analyzing the Social Web. Morgan Kaufmann.

Grayson, S., Wade, K., Meaney, G., & Greene, D. (2016). The sense and sensibility of different sliding windows in constructing co-occurrence networks from literature. In J. Shen, H. Tong, J. Tang, & X. Zhou (Eds.), Social Informatics: 8th International Conference, SocInfo 2016. Proceedings, Part I (pp. 65-77). Springer.

Hagberg, A., Swart, P., & S Chult, D. (n.d.). Exploring network structure, dynamics, and function using NetworkX. Exploring network structure, dynamics, and function using NetworkX. Los Alamos, NM, United States: Los Alamos National Lab (LANL).

k_core. (n.d.). Retrieved 2024, from NetworkX: https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.core.k_core.html

McKinney, W., & others. (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, 445, pp. 51-56. Austin, TX.

Powell, J., & Hopkins, M. (2015). Graph analytics techniques. In J. Powell, & M. & Hopkins, A Librarian's Guide to Graphs, Data and the Semantic Web (pp. 167-174). Chandos Publishing.

Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. Proceedings of the National Academy of Sciences, 101(9), pp. 2658-2663.

Van Rossum, G. (2020). The Python Library Reference, release 3.8.2. Python Software Foundation.

Van Rossum, G., & Drake Jr, F. L. (1995). Python reference manual. Amsterdam: Centrum voor Wiskunde en Informatica.

Wikipedia contributors. (n.d.). The Count of Monte Cristo. Retrieved 2024, from Wikipedia, The Free Encyclopedia: https://en.wikipedia.org/w/index.php?title=The_Count_of_Monte_Cristo&oldid=1215274087