

Supporting Literary Analysis Using Automatically Generated Social Network Graphs

Jocelyn M. Bayer
Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA 56267
bayer106@morris.umn.edu

ABSTRACT

In recent years, the availability of literary texts in electronic format has increased rapidly due to an interest in digital humanities. The digital format of these texts and the greater access to them allows texts to be analyzed using computational methods. The introduction of computing helps overcome one of the biggest obstacles of literary analysis: the amount of information found in a text. One particular computational approach to analyzing literary texts is to use social network analysis. In this approach, a social network is created from the relationships between characters in a text, which can then be analyzed and used to draw conclusions about characterization, plot, and genre. Because these properties of texts are difficult to track by hand, especially for entire literary periods, this automated approach allows past hypotheses to be tested for validity and new insights to be discovered about single texts as well as large bodies of texts. In this paper, we describe the creation of these social networks and some of the results of using these networks to do literary analysis.

Keywords

literary analysis, social network analysis, graph theory, digital humanities

1. INTRODUCTION

The ease with which printed texts can now be digitized has led to ever expanding digital libraries of literary texts. Texts that were once only easy to find in printed format are now available through archiving libraries found on the web. These libraries also make it easier to find and compare texts written in the same literary period. Because of the digital format of these texts and their availability to a greater audience, they can be examined in new ways. In this paper, we focus on the use of social network analysis to examine literary texts. In order to do so, social networks are created with the characters as nodes and the undirected edges as interactions between characters, which are refined by the length

and frequency of the interactions. The resulting social network can then be used to understand the existence, strength, and prevalence of relationships between characters using the weights of the edges. The overall shape and structure of this social network can also be used to draw conclusions about characterization, plot, and genre characteristics.

These conclusions drawn about the literary properties of characterization, plot, and genre can provide insight into areas that would have been difficult to prove only using manual literary analysis. Manual literary analysis tends to rely on specific occurrences of a phenomenon in a text rather than the overall text. This makes it difficult to deal with literary properties like plot and genre which tend to develop or be seen over the course of a novel. It is even more difficult to track the similarities and differences in genre and plot over an entire body of literature or literary period. As a result, hypotheses proposed by literary scholars about these properties are difficult to confirm using manual literary analysis. Using social network analysis on social networks created from character interactions in a text allows for literary analysis on overarching properties within a text. It also allows two or more texts to be compared based on the way characters are connected and the shape and structure of their corresponding networks. This can be helpful in drawing conclusions about a writer's body of literature or a specific literary period.

In this paper, we begin by discussing the past of literary analysis and how literary analysis using social networks differs from past techniques. Then, we explain the general algorithm that has been used in different experiments to create social networks from literary texts and how different approaches affect outcomes. Finally, we consider how the information collected using social network analysis can be used to test literary hypotheses and infer new information about literary texts.

2. BACKGROUND

In order to understand the insights that literary analysis using automatically generated social networks provides, it is necessary to understand how manual literary analysis works. Typically a text is read from beginning to end with special attention to trends within the text. Trends can be how a character changes through the course of the story, how the author treats a theme in the text, how the author does or does not adhere to a genre, among many other points of analysis. Once a specific area of analysis is chosen, short sections of the text that include material related to that area are examined further. A technique called close reading is often

This work is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/us/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

UMM CSci Senior Seminar Conference, Fall 2016 Morris, MN.

used in this examination. Close reading involves paying attention to diction, noting qualities such as the type of words used, when they are used, and any connotations they might have, along with other literary elements like symbolism and setting [1].

While this technique enables the analyzer to make claims about the text as a whole, the evidence concentrates on a few specific examples. It would not be plausible to include every example of a trend within a text as evidence using this technique. Creating a social network from the characters and their interactions enables a majority of these interactions to be used in analysis. Because character interaction tends to be what drives plot and the importance of certain characters and which characters interact effects genre, only character interactions need to be tracked in order to draw conclusions about these kinds of literary properties. A social network provides a good framework for this. As long as an interaction is between characters that play a significant role in the text, it will be factored into the weight of an edge between two characters. Once a network is constructed, social network analysis metrics can be used to draw conclusions about a text’s literary properties such as plot and genre. While close reading uses a micro-level approach, social network analysis analyzes from a macro-level. Instead of close-reading, there is almost an approach of “distant reading” [6]. Claims that are made using close reading that apply to overarching properties such as genre can then be tested for validity using social network analysis, to see if they hold up when applied to the entire text.

3. METHOD

In order to perform literary analysis using social network analysis, the graph representations of character interactions must first be built. Once this is completed, social networks analysis metrics can be calculated and humans can use the visual representation of the network to do literary analysis. This section will discuss the necessary pre-processing steps, the general process used by recent experiments, and the impacts of different choices in the process on different types of data.

3.1 Pre-processing

Even with the explosion of literary content on the web, especially public domain literature, there is still no easy way to immediately create a social network of characters from an entire text. To be able to create the representations of the relationships between characters, texts must first be processed. It would not be beneficial if a description of a landscape was included in the data or the city a novel is set in was treated like a character. Previous studies have shown that pre-processing allows for the removal of such unrelated data and as well as preventing unnecessary repetitions of characters.

The first step is to find all the character names in the text. In the experiments by Elson et al. (2010), as well as Ardanuy and Sporleder (2014), the Stanford Named Entity Recognizer (NER) was used for this purpose. NER is a software that labels instances of names. Because the NER is multi-purpose software, these experiments used the optional **Person** class provided by Stanford, as well as other augmentations, in order to tailor it to find characters in literary works [2].

However, literary works do not always use the same name

<u>Character Name</u>	<u>Occurrences</u>
Harry Potter	345
Harry	17827
The Boy Who Lived	13
The Chosen One	17
Voldemort	1106
You-Know-Who	201
He-Who-Must-Not-Be-Named	42
The Heir of Slytherin	29

Figure 1: An example list of characters that need to be resolved to proper character names or co-referents. Modeled after the list from [4].

to refer to a single character. A character may have one or many nicknames, or may be referred to by different defining characteristics at different points in a novel [4, 8]. These different names for a single character are called co-referents. A matching algorithm or other strategy may be used in order to find all the co-referents for each character [2, 3]. In order to create the networks, the co-referents are then clustered together or are replaced in the text by the proper name of the character [3, 5]. The algorithm will then be able to create a social network with each distinct character as a node, that includes all of their interactions no matter the co-referent used.

3.2 General Algorithm

Once the different characters are found, the network can be created. In these networks, characters are vertices or nodes and the interactions between characters are the edges. Depending on the approach, what constitutes an interaction may be different [5, 8]. This is considered in the next section. The edges are then weighted depending on the number and length of interactions between characters. These edges are undirected which means as long as the conditions for being an interaction are met, the weight between characters increases even if one character participates more in the interaction than the other [2].

Oftentimes, there are characters that do not appear more than a few times in a text and affect the plot very little. In comparison to the main characters, they do not have very many interactions. These characters are often pruned before the network is created or when the resulting network has too many leaves whose edges are not highly weighted [5, 8]. This is done in order to make the resulting network less cluttered and therefore easier to read.

3.3 Different Approaches

While most experiments use the same general approach to creating networks, there are some distinctive choices concerning what constitutes an interaction between characters and whether each novel is considered as a whole. This section outlines the choices made by recent experiments and the effects these choices have.

One of these choices is how interactions between characters are classified. In the experiment by Elson et al. (2010), interactions are classified by conversation. In order to be considered interacting, characters must be in the same place, know that the other person is there, and be speaking to each

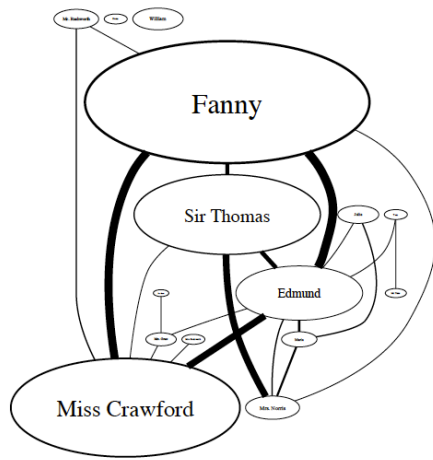


Figure 2: Social network created from the character interactions found in the novel *Mansfield Park* [5]

other. If two characters are linked by an edge, they have participated in conversation. The weight of an edge corresponds to the amount of conversation that exists between the two character nodes the edge connects [5].

While this may work well for dialogue heavy texts, dialogue is not always the most prominent way that characters communicate or interact. There are also texts that do not use the conventional form of marking dialogue consistently or are not dialogue heavy, which would need to be taken into account when deciding whether to use this approach [2]. An alternative to the focus on dialogue is to classify interactions based how closely the two characters are named in a text. An interaction is counted if the two characters are mentioned within a certain window of text [2, 8]. The weight of an interaction can be refined by calculating the distance between two character mentions.

Another choice made by recent experiments is whether one network is created from the entire text or whether separate networks are created from sections of the text. In the experiments such as those done by Ardanuy and Sporleder (2014), creating a single network allows for analysis on the novel as whole. The conclusions drawn can apply to overarching plot of the novel, centrality of the main characters, and characteristics of how certain authors set up their novels.

Sectioning texts off into distinct parts, as opposed to creating a network from the entire novel, yields its own interesting results. Because each network corresponds to a specific chunk of the text, changes can be seen over time. In the experiments done by Seo et al. (2014), this allowed conclusions to be drawn about shifts in the identity of the main character, importance of characters at different points in time, and how these characteristics play into the notion of genre.

4. DISCUSSION

Once a social network has been extracted from the characters and their interactions within a text, social network analysis can be used to draw conclusions about the text as well as the body of literature it comes from. This section discusses the insights into characterization, community structure, plot, and genre that can be extracted from lit-

erary social networks, some of the limitations of this type of analysis, and practical applications of the techniques in experiments.

4.1 Characters and Community

Some of the most basic insights that can be taken from a social network is information regarding characters and the relationships between them. For example, in Figure 3, it is very clear that the main character in Jane Austen’s *Pride and Prejudice* is Elizabeth. This is indicated by the centrality of the Elizabeth character node. Centrality is a measurement of the importance of a node in a network. For social networks, this is measured by the number of edges a node has. In this example, Elizabeth has the most connections and has had interactions with most, if not all, characters. The community portrayed in the novel is centered around her [2].

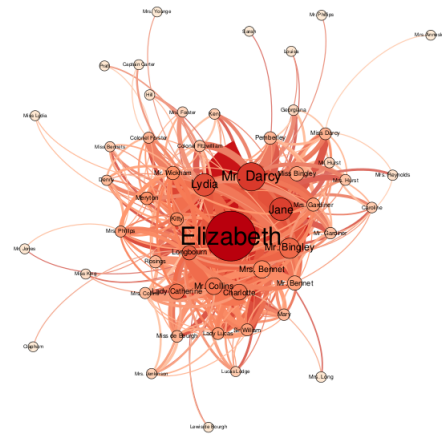


Figure 3: *Pride and Prejudice* [2].

In other works, no single node has a high measure of centrality. The network of William M. Thackeray’s *Vanity Fair* has no clear main character. There are nodes that have higher centrality, but no single node has significantly more edges than the others. From Figure 4, Rebecca, Amelia, and George are most likely the three characters most focused on in the text. They are important to the community of characters, but do not have as an important role in the text as

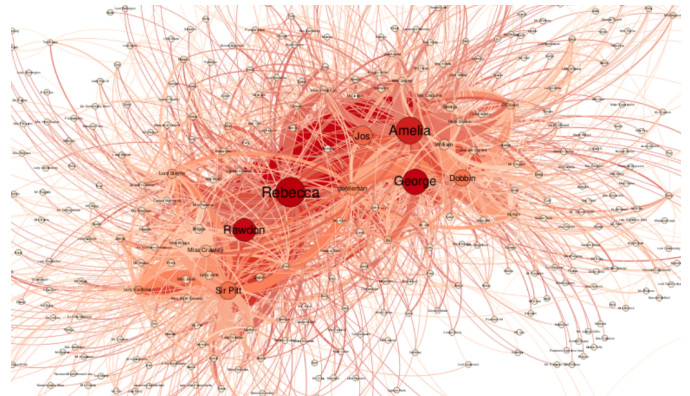


Figure 4: *Vanity Fair* [2].

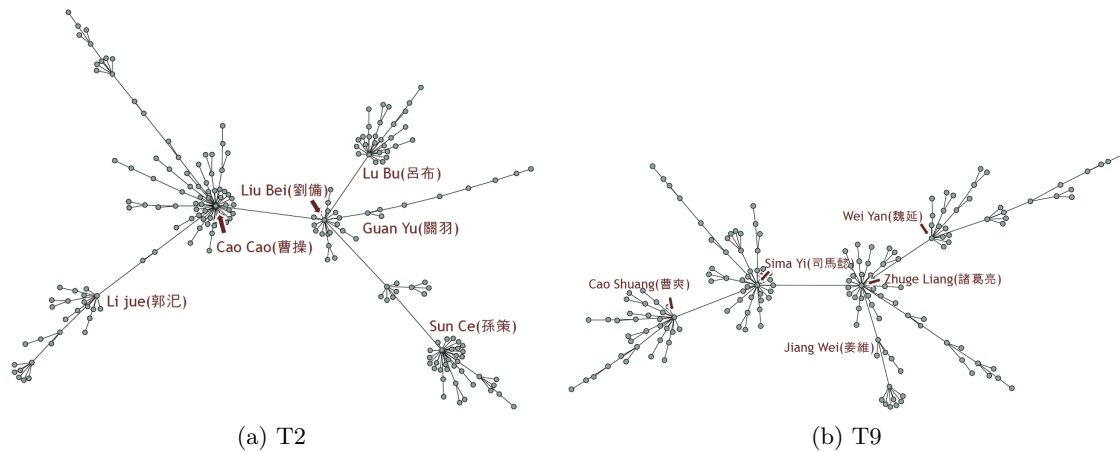


Figure 5: Dynamic social networks for *Romance of the Three Kingdoms* [8]

Elizabeth did in *Pride and Prejudice*. The community is also not as tight-knit. There is a larger number of character nodes and many that have fewer edges between them. There is also much more cohesion in the network created for *Pride and Prejudice* than *Vanity Fair*. Cohesion refers to how closely related characters are to each other. In Figure 3, most of the characters have interacted with most of the other characters. There are few characters that have not done so. Conversely, in Figure 4 there is much less cohesiveness. There are many characters that only interact with a small subset of the total characters in the text [2]. These differences in these social network shows the difference in community structure between the two novels.

4.2 Plot

A social network of characters can also give insight into plot. One way of doing this is using dynamic social networks. A dynamic social network does not take the entirety of a novel in mapping the interactions between characters. Instead, the text is divided off into sections and a social network is created from each of these sections. An example of this is shown in Figure 5, which shows the social networks for section 2 and 9 of the text *Romance of the Three Kingdoms*. It can be seen that through the course of the text, the main character changes from one section to the next due to the character node with the highest centrality never being the same. The main characters at T2 are Cao Cao and Liu Bei. The main characters at T9, however, are Zhuge Liang and Sima Yi. This is in opposition to Western texts like *Pride and Prejudice* or the *Harry Potter* novels by J. K. Rowling in Figure 7. The plot of these novels is centered around one main character that is the focus of the entire text [8, 2]. Instead of a plot focused on one character, the *Romance of the Three Kingdoms* follows a historical period in Chinese history. This means there is no single hero or heroine, but many.

Another method of analyzing for plot is looking at segmentation. Segmentation is a social network metric that measures how closely tied together a group is. Sometimes a social network might have a segment that is isolated from the rest of the network. The character nodes within this segment are usually well connected to other nodes in the group, forming a clique or social circle, but may not be connected

to the rest of the network with the exception of one or two nodes. An example of this can be seen in Figure 6. This may indicate a micro-narrative within the plot. A micro-narrative is small often self-contained plot within the main plot. It usually takes the form of a character's backstory or an anecdote told by a character during a conversation. An author may use a micro-narrative to indicate the speaker's personality in a way that does not effect the main plot of the narrative [6]. The visualization of the social network can help distinguish these micro-narrative character nodes, and distinguish between them and the characters nodes that are more important to the main plot of the text. [6].

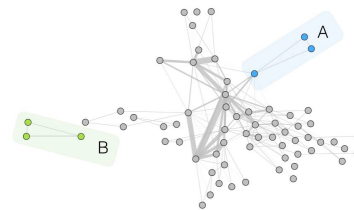


Figure 6: Micro-narratives in *Oliver Twist* [6]

4.3 Genre

The area of genre is the most difficult area of analysis, in part because it is one of the least clear-cut areas in literary analysis as a whole. There are constant debates in the literary community about whether a text falls into a certain genre or the more fundamental question of exactly what characteristics compromise a genre. This leads to genre being more difficult to understand through calculated social network analysis metrics, relying more on human interpretation of the networks. The metrics such as centrality and cohesiveness can be used to draw conclusions about the areas of character relationships and plot, but genre is more difficult if not impossible to directly quantify. However, these metrics can give literary scholars a unique perspective about a text's genre or genres when paired with the visual representation of a literary social networks.

One of the insights that literary social networks can provide is how their structure can affect the genre of a text. An example of this can be seen in the social network of the

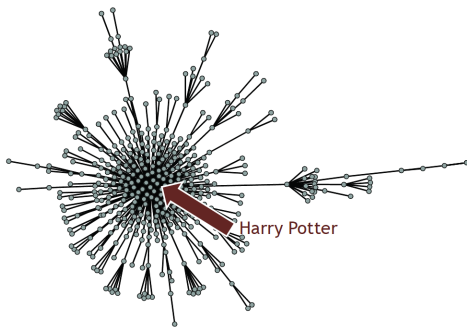


Figure 7: *Harry Potter* novels [8]

Romance of the Three Kingdoms seen in Figure 5 and the social network of the *Harry Potter* novels seen in Figure 7. According to literary theorists, the *Harry Potter* novels are classified as heroic novels, while the *Romance of the Three Kingdoms* is an epic novel. Heroic novels are usually focused on a single central character, their journey, and the relationships of the other characters with this central character [7, 8]. On the other hand, epic novels focus on a cast of connected characters rather than a single character [8]. These texts detail the various adventures of these characters, which often depict the history or legends of a country or culture. Within these texts, the title of main character does not belong to a single character, but often depends on the point of time or particular event being described [7]. Social network analysis metrics can provide evidence that can be combined with known literary knowledge to make claims about the genre of a text. The centrality of the Harry Potter node in the network for the series in Figure 7 suggests that Harry Potter is the main character. The nodes are also clustered tightly around this node, suggesting that the plot is centered around him and he is the main focus. This corresponds to the definition of a heroic novel. On the other hand, the node with the highest centrality for *Romance of the Three Kingdoms* changes over time and the main character for each section is only slightly more important in the network than the rest of the characters. This seems to indicate an epic novel rather than a heroic novel, where the different main characters are different important figures in the history or legends of a country, a characteristic of an epic [7, 8].

However, not all genre criteria can be inferred from social network metrics. Many genres overlap with each other, having many characteristics in common. These characteristics often cannot be inferred from character interactions alone. An example of this can be seen in the experiments by Ardanuy and Sporleder. One of the novels used in their experiments was *Pride and Prejudice*, whose social network can be seen in Figure 3. *Pride and Prejudice* is usually considered to be a bildungsroman, a genre characterized by a youthful main character trying to understand the world they live in and their place in it [7]. The plot is usually strongly concentrated on this character and there is a cohesive group of characters connected to them, which can be translated to the social network metrics of centrality, clustering, and cohesiveness. However, there are other texts that also have these traits, but are not included in the bildungsroman genre. The picaresque is an example of a novel genre that is also centered around a main character with

a cohesive group of characters strongly connected to them. The difference is that the picaresque is much less structured and usually episodic, while the bildungsroman is structured around the growth of the main character and has a plot of connected events [2, 7]. These traits cannot be inferred from a social network, but would need to be known by the analyst of the network.

In the experiments by Ardanuy and Sporleder, they attempted to use social network metrics such as graph density, average clustering coefficient, diameter, radius, proportion of eccentric, central and isolate nodes to classify literary social networks by genre and author. Due to the aforementioned difficulties with genre, the clustering by author was much more successful in comparison to clustering by genre. There were quite a few cases of text's networks being misclassified due to similarities in characteristics that can be partially quantified by social network analysis metrics. A picaresque would be classified as a bildungsroman, or vice versa. Authors, on the other hand, tend to use similar character relationships and community structures in their works, which correlates better with social network analysis metrics [2]. While social network analysis is currently effective in a descriptive sense, there is still room for improvement in its predictive capabilities.

4.4 Literary Theory

While the predictive power of literary social network graphs still requires improvement, the descriptive capabilities of literary social networks have a lot of potential for use in regards to literary theory. Many literary theorists, such as the famed Mikhail Bakhtin, were working at a time when it would be very difficult to read an author's entire corpus or the entire corpus for a literary period. As a result, their theories tended to rely on a specific collection of texts they were familiar with for evidence that their theory was applicable to a larger body of literature. As a result, many of their theories were taken to be applicable across the entire body of literature because the evidence provided seemed conclusive for the specific examples.

The speed with which literary social networks can be generated allows literary analysts to compare the characteristics and structures within an entire body of literature through their social network graphs and assess the validity of literary theories. In the experiments by Elson et al. (2010), literary social networks for 60 nineteenth century novels were generated and used to assess whether the setting of a novel affects the characters and the relationships between them. Bakhtin theorized that the character relationships in novels with urban settings would differ from those in novels with rural settings due to the "different social and emotional potentialities" [5]. Rural communities were thought to be more tightly bound with fewer characters and more conversation, while urban communities were loose-knit with many characters who do not engage in much conversation. However, there was very little difference in the number of characters, how close-knit the communities were, or the number of conversations in the experiments. Even with a smaller subset of nineteenth century novels, this kind of experiment calls into question a theory that had been taken as fact for the literary community for a very long time.

The ability to quickly generate literary social networks can also help in determining characteristics of an author's entire body of work as well as how they differ in style from other au-

thors. In the experiments by Ardanuy and Sporleder (2014), literary social networks were used to determine characteristics of different authors. The results show some of information that can be gleaned about an author’s style. Certain authors, such as Jane Austen, tend have novels with very high graph density whereas William Thackeray’s novels have significantly lower graph density. Sometimes the graph density changes over time, like in J. K. Rowling’s *Harry Potter* novels, where the graph density lessens over time. The significance of the main characters of an author can also be inferred. Austen and Rowling have novels where the protagonist plays a significant role in the text in contrast to novels written by Thackeray, who focuses less on the protagonist and spends more time exploring minor characters [2]. While much of this information aligns with current literary criticism regarding authorial style, it holds potential for ascertaining the validity of claims about less well-known and well-researched authors as well as more recent authors’ bodies of works.

5. CONCLUSION

The use of computational techniques in the digital humanities has opened up a whole new world of possibilities. Literary social networks are a powerful tool that allows computational techniques to be used to investigate and understand long-neglected areas of literature that had long evaded literary scholars. These areas were often neglected due to the prevalence of the close reading technique in literary analysis and the limitations of a human reader. In using social network analysis, characteristics from each text are based on the text as a whole or sections of a text rather than a limited number of examples. In making the generation of social networks automatic, it becomes much easier to produce literary networks for a large number of texts. This allows analysts to notice characteristics for a number of works of literature written by a certain author or written during a certain time period.

However, there are still drawbacks to this technique. It provides evidence for claims about the characteristics of a text, but there is still room for improvement in using social network analysis techniques on literary social networks for predictive purposes. One example would be genre, where it provides evidence of a text being of a certain genre, but is unable to accurately predict the genre based on social network analysis metrics alone. A human component to the process is still necessary in order to predict with any accuracy. Using social network analysis does not remove the human element. It instead provides evidence and support for human analysis that would be impractical for human analysts to gather on their own.

6. REFERENCES

- [1] H. Bort, M. Czarnik, and D. Brylow. Introducing computing concepts to non-majors: A case study in gothic novels. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education, SIGCSE '15*, pages 132–137, New York, NY, USA, 2015. ACM.
- [2] M. Coll Ardanuy and C. Sporleder. Structure-based clustering of novels. In *Proceedings of Computational Linguistics for Literature (CLFL)*, EACL 2014, pages 31–39, Gothenburg, Sweden, 2014.
- [3] M. Elsner. Character-based kernels for novelistic plot structure. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 634–644, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.
- [4] M. Elsner. Abstract representations of plot structure. *LiLT (Linguistic Issues in Language Technology)*, 12, 2015.
- [5] D. K. Elson, N. Dames, and K. R. McKeown. Extracting social networks from literary fiction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, ACL '10*, pages 138–147, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.
- [6] S. Grayson, K. Wade, G. Meaney, J. Rothwell, M. Mulvany, and D. Greene. Discovering structure in social networks of 19th century fiction. In *Proceedings of the 8th ACM Conference on Web Science, WebSci '16*, pages 325–326, New York, NY, USA, 2016. ACM.
- [7] C. H. Holman. *A Handbook to Literature*. Prentice Hall, 10 edition, 2005.
- [8] J.-K. Seo, S.-H. Kim, H.-S. Tak, and H.-G. Cho. A structural analysis of literary fictions with social network framework. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing, SAC '14*, pages 634–640, New York, NY, USA, 2014. ACM.