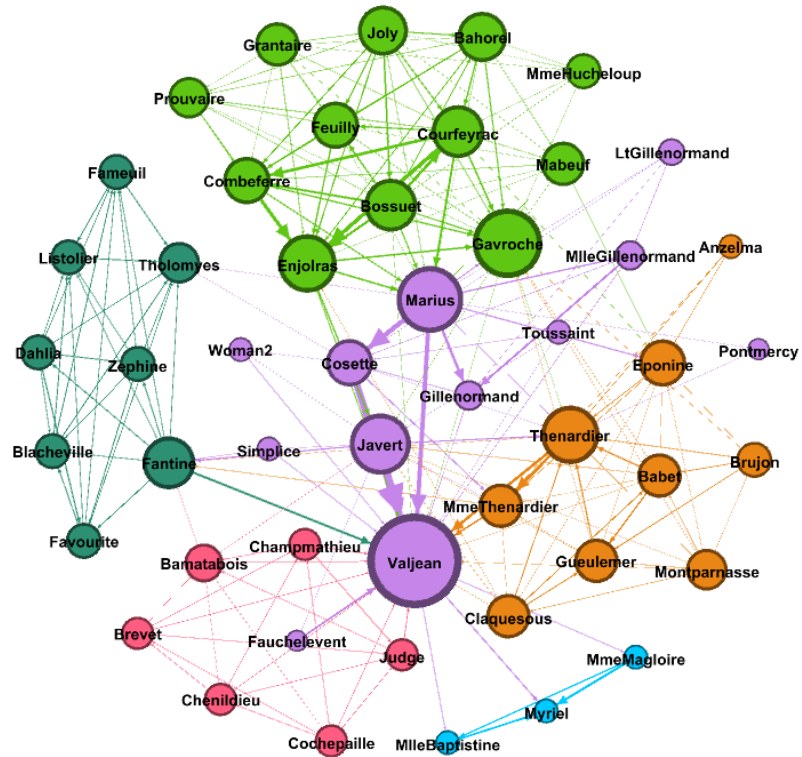


Social Network Analysis of Les Misérables



Data Description

The network analysis is related to the connection between the character in the novel 'Les Misérables', which is a French historical novel by Victor Hugo, first published in 1862. The dataset is downloaded from the Victor Hugo from Gephi. This analysis aims to analyze the connectivity between characters throughout the novel. And find out who has the most connection with other. There are 77 nodes and 254 edges in the dataset.

Method and Result

Basic Analysis	Result		
Graph Density	0.043		
Average Degree	3.299		
Average Clustering Coefficient	0.287		
Modularity	0.565, No. of communities: 6		
Degree Distribution	1: 17	2: 10	17: 1
	3: 6	4: 3	19: 1
	6: 5	7: 10	22: 1
	8: 1	9: 3	36: 1
	10: 5	11: 6	
	12: 2	13: 2	

	15: 2	16: 1	
Average Path Length	2.4		

The node size is determined by the degree distribution, which is the sum of out-degree and in-degree. This reflect the number of connections between characters. The larger the node, the more the connection with other characters. The node color is divided by the modularity class. Moreover, the degree range is set from 2 to the maximum (36), so make the graph more efficient. “Fruchterman Reingold” layout is used, then “Noverlap” layout was used for preventing the node overlap and “Label Adjust” is run to prevent label overlapping.

Result and interpretation

There are six modularity classes given in different colors (light green, dirty green, orange, blue, purple, and red colors). The 0.287 average clustering coefficient indicate that the familiarity between characters are low. The 0.565 modularity value indicates that the interactions between characters occur predominantly within the group.

The graph shows that Valjean has the largest node, which means he has the most connectivity with other characters throughout the novel. It is possible because he is the main character of the novel. Also, from the purple class, Cosette, Javet, Marius are the main characters in their community as their node size is large compare with others in the same community. In the blue group, Myriel may be a significant character in the community. Then in the orange group, Thanardier is a significant character in his community. Moreover, in the light green group, Enjolres and Gavroche are also clearly the main characters in their community. Besides, among the dirty green groups, Fantine is the main character in the community. Overall, Valjean has the most connection with other characters in the novel, followed by Gavroche, Marius, and Javet.

Additionally, in the co-appearance perspective, the edge thickness between nodes is determined by the weight. The thicker the edge, the more co appearance of the characters throughout the novels. The edge between Valjean and Javert is the thickness, which indicates they always co-appear throughout the novel. This may because Javert is an antithesis of Valjean and obsessed with tracking down and re-imprisoning Valjean. Besides, Valjean and Cossette's edge is thick, indicating that they always co-appear throughout the novel. It is possible because Cosette is the adopted child of Valjean. Moreover, the edge between Cosette and Marius is thick. It may occur because of the loving relationship between Cosette and Marius in the novel. Furthermore, it is clear to see that although Javert and Marius have a large node, they never co-appear in the novel. Apart from this, it is clear to see that Valjean had connected with others in other community groups. However, some

characters are far apart from the center and with small node sizes, indicating that they are not the novel's main characters.

To conclude, Valjean is the main character of the novel. And other main characters and supporting characters include Gavroche, Marius, Javet, Fantine, Enjolras, Fantine, and Thenardier. Furthermore, the most co-appearance character pair is Valjean and Javert. This social network analysis is helpful to find out the relationship between the characters. However, this analysis's primary focus is to discover the most frequency connected with other characters and most co-appearance among the characters throughout the novel. Thus I would only explain the result related to the purpose.