THE RELATIONSHIP BETWEEN TWITTER RESPONSES TO ALLEGATIONS AGAINST MP CHRISTIAN PORTER: A SOCIAL NETWORK ANALYSIS

BUSA2020 Advanced Analytics Techniques
Assignment 1

45345473

Amritha Jeyarathan

TABLE OF CONTENTS

I. PURPOSE

This report visualises and analyses a social network to support social phenomena including group formation, social roles and personal influence.

II. METHODOLOGY

II A. DATA COLLECTION

Netlytics was employed to extract a dataset based on Twitter responses following allegations of historical sexual assault against Attorney-General Christian Porter. The search query included terms **Christian Porter** and **sexual assault/rape**. For further specification, tweets were filtered, only including those with replies and a minimum of five retweets (figure 1). A name network analysis was then downloaded to be analysed in *Gephi* (figure 2).

Figure 1. (left) Search query performed on Netlytics to extract dataset from Twitter.

Figure 2. (bottom) A visualisation of the name network analysis, (based on who mentions whom in a tweet) on Netlytics.
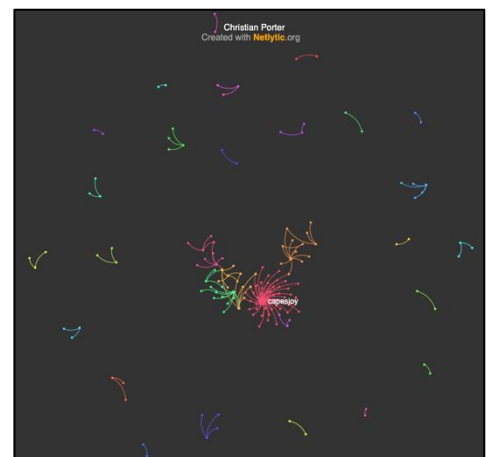
Nodes: 163
Edges: 142
Nodes: 95 (58.28% visible)
Edges: 103 (72.54% visible)

Figure 3. (above) The number of nodes and edges prior to and following a filtration of the dataset based on 'Giant Component'. The filter removes noise and only includes the 'main island' of the network, reducing the number of nodes by 41.72% and the number of edges by 27.46%.

II B. DATA VISUALISATION

The *Giant Component* filter facilitated in removing smaller detached groups, seen in figure 2, that don't contribute significantly to the analysis (figure 3). The *Force Atlas 2* layout was utilised to effectively communicate and maintain a clear division between different clusters.

II C. ANALYSIS A: MEASURING DEGREE AND MODULARITY CLASS

Degree was measured to communicate individuals possessing the most direct connections in the network and potentially, a greater 'reach'. The size of each node (individual) was determined by their degree ranking. Also considered, was modularity class to distinguish various subgroups that form and cohere. Each node colour represents a different modularity class (figure 4).
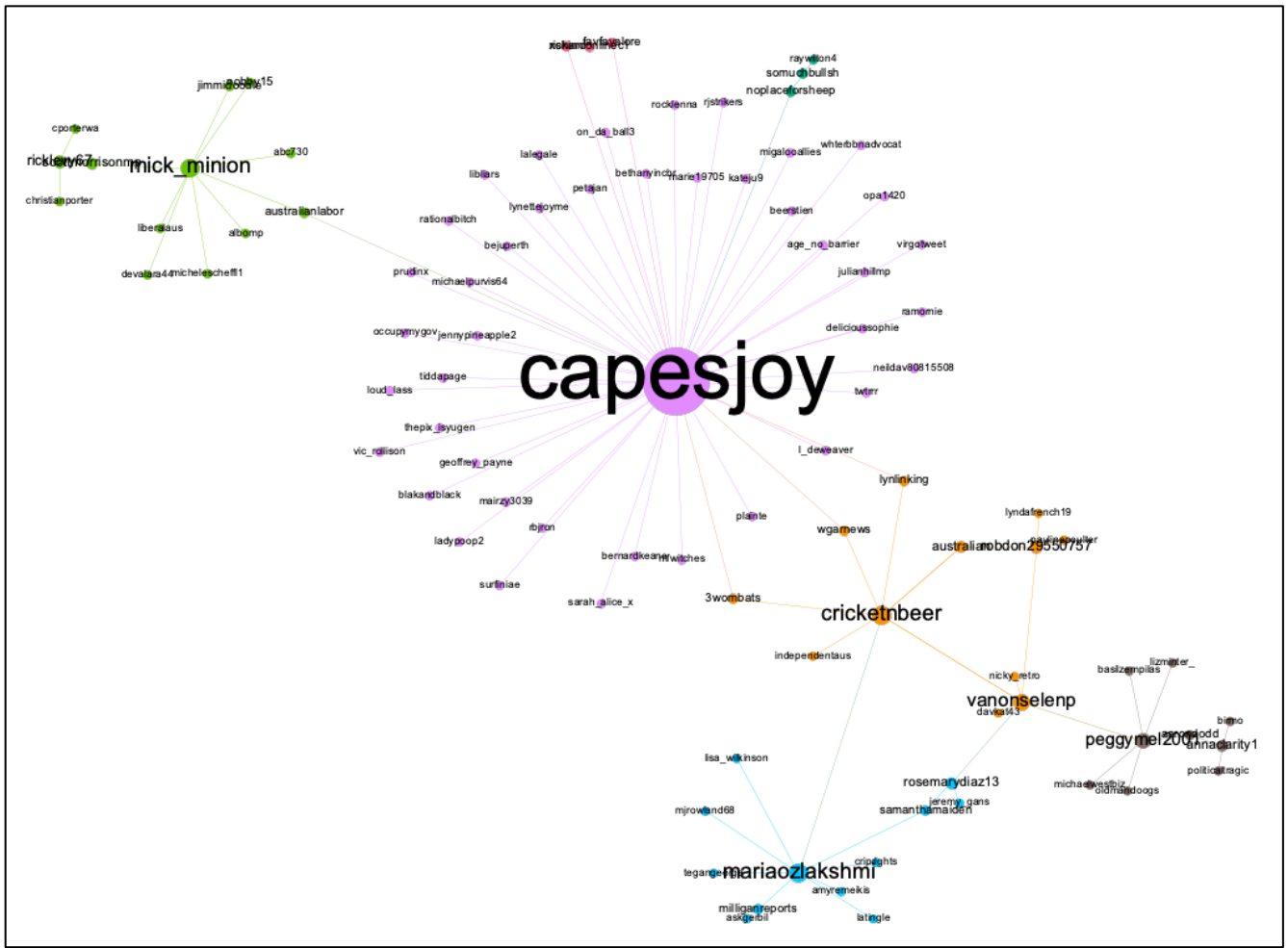
Figure 4. (above) Analysis A reveals nodes (or individuals) with varying degrees, based on size i.e. the larger the node, the more direct connections and thus potential 'reach' the individual has while smaller nodes denote with less connections and hence less contact. Modularity classes are based on colour, where each class is represented by a colour.

II D. ANALYSIS B: MEASURING BETWEENESS CENTRALITY AND EIGENVECTOR CENTRALITY

Betweenness centrality identifies network bridges – individuals connecting otherwise unconnected clusters, greatly influencing the flow of information between different subgroups. Node size was determined by their betweenness ranking (figure 5). Also measured was eigenvector centrality, indicating individuals with greater access to influential users. Node colour was ranked by its eigenvector score (figure 5).

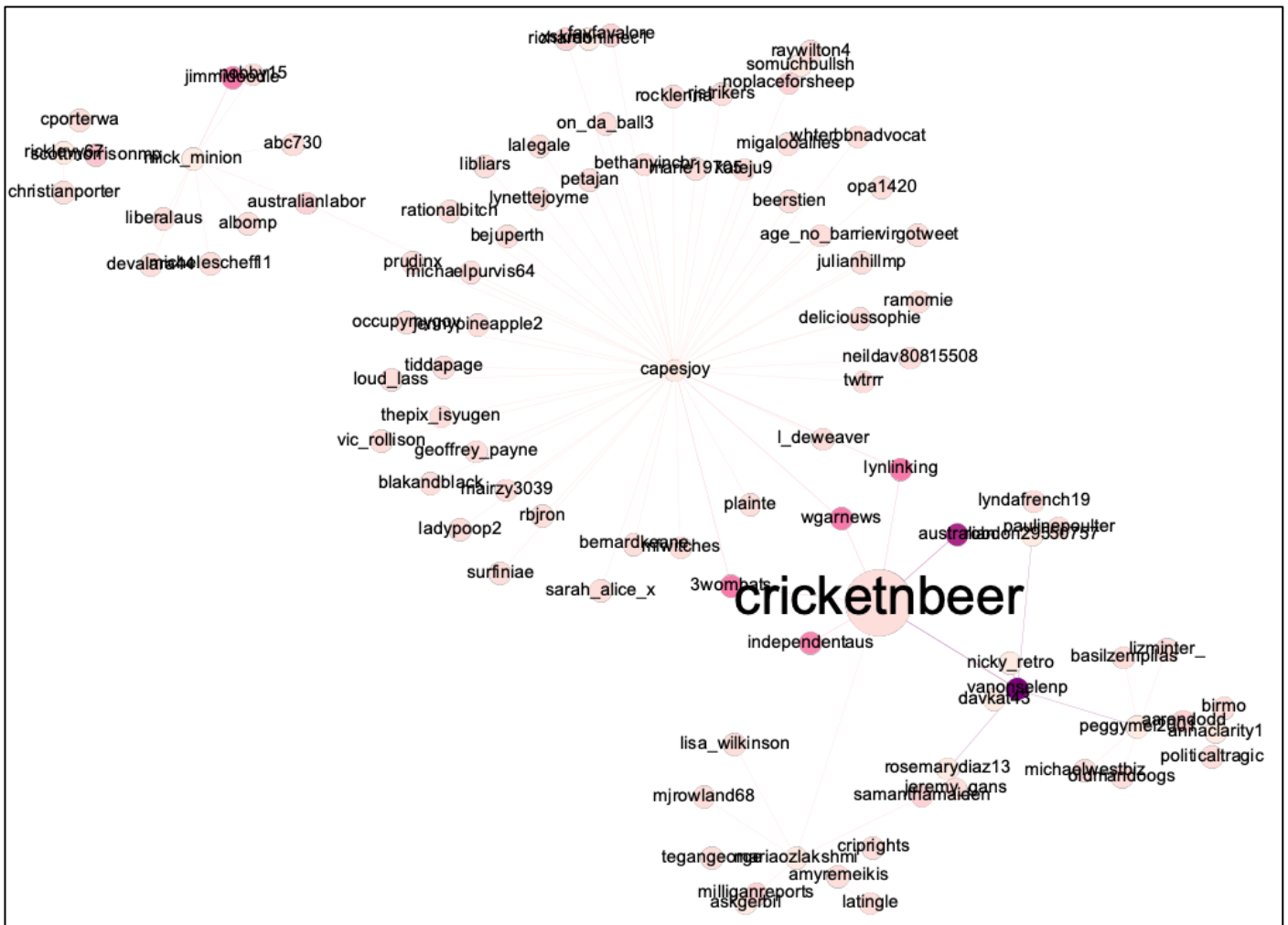*Figure 5. (above) Analysis B conveys nodes based off their betweenness centrality i.e. the larger the node, the higher betweenness score and thus, a greater ability for the individual represented to connect different clusters and; eigenvector centrality i.e. the darker the node, the higher eigenvector score and consequentially, the greater the access that individual has to influential Twitter users.*

III. CONCLUSIONS

III A. CONCLUSIONS FROM ANALYSIS A

With the average individual having 1.084 connections, *capesjoy* sits highest at 49 (figure 6); inferring they potentially have the greatest 'reach'. A 0.639 modularity implies sturdy divisions between the seven subgroups. From external knowledge, this may be true as communities containing political figures (green) and journalists (blue) are segregated and don't directly interact (figure 4).

Id	Label	Interval	isposter	user_statuses_count	user_friends_count	user_followers_count	In-Degree	Out-Degree	Degree
n37	capesjoy		☑	21874	654	1282	0	49	49

Figure 6. (above) Capesjoy has the highest degree of 49, with an in-degree of zero and an out-degree of 49. Although they prove to have the greatest number of connections, it does not necessarily imply greater social capital. Note that their overall degree is solely their out-degree, suggesting that out of all individuals, they have mentioned the largest number of people. Hence, it can be implied that they have the largest 'reach' as more people are likely to see their post and gain exposure to the conversation as a result of being mentioned.

III B. CONCLUSIONS FROM ANALYSIS B

Cricketnbeer, boasting the highest betweenness value of six, has the vital role of connecting isolated subgroups, yet possesses a lower eigenvector score. *Vanonselenp* has the highest eigenvector score, holding power for being connected to other powerful individuals, but conversely scores lower on betweenness; indicating the two measures are independent (figure 7).

4

Eigenvector Centrality Distribution — **Betweenness Centrality Distribution**
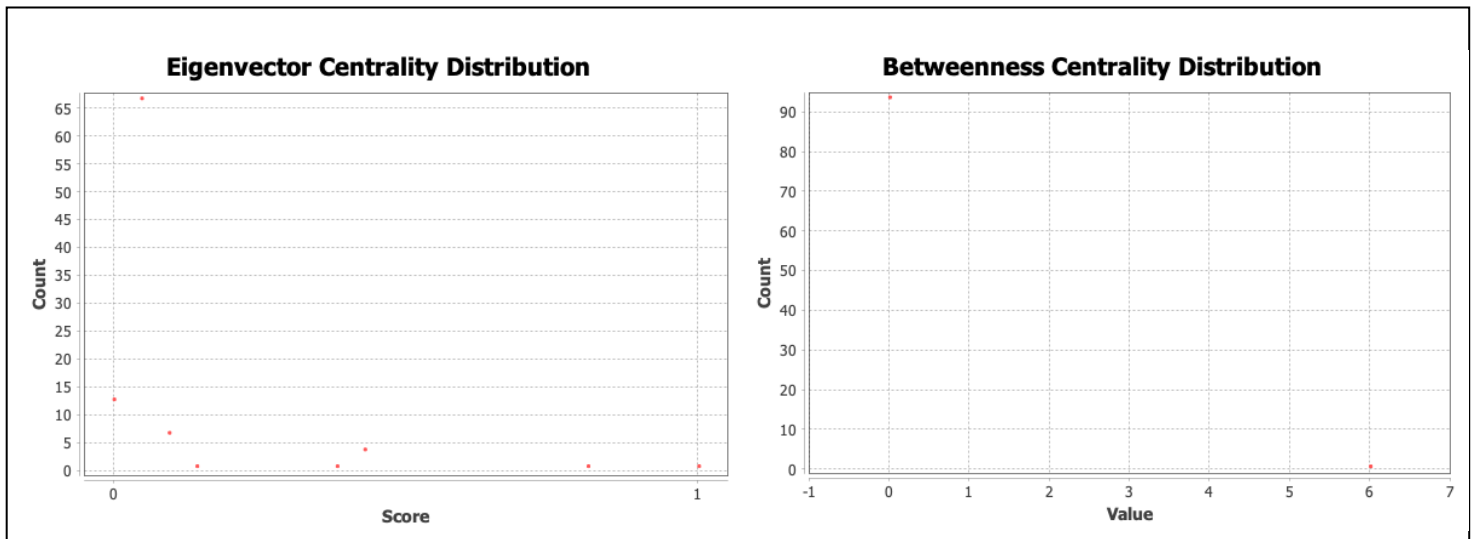
Figure 7. (above) The distributions of eigenvector centrality and betweenness centrality produced by Gephi. Both measures follow a similar distribution i.e. most of the individuals tend to score relatively low for both measures only one to two individuals tend to score high. Individuals however, can score high on betweenness but low on eigenvector, as they are still distanced from the centres of power in the network. Similarly, individuals can score high on eigenvector but low on betweenness, if they aren't on the shortest path between two clusters (figure 5).

III C. CONCLUSIONS FROM ANALYSES A AND B

Despite *capesjoy* having the greatest 'connections', their associations are likely less powerful individuals, explaining their low eigenvector score. These high-degree individuals also scored lower in betweenness; inferring that more connections don't equate to a more influential or wider reach. In analysis B, those initially seen 'inconsequential' (*cricketnbeer* and *valonselenp*), have greater roles in transferring information or reaching influential audiences. (figure 8 and 9).
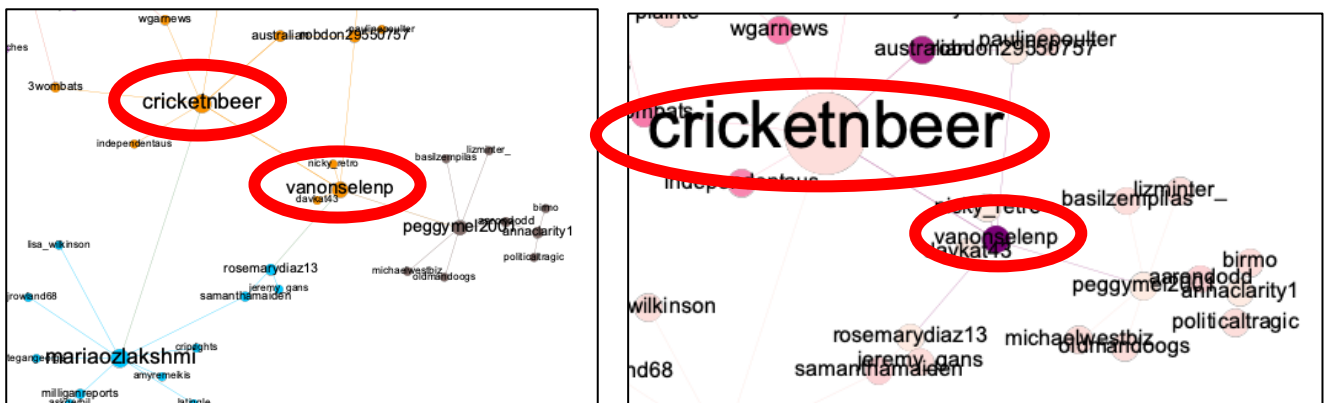

*Figure 8. (above) Analysis A and B convey two different truths, individuals circled are seen as insignificant in A and their social roles and importance aren't fully realised until B. Conclusions from **III C.** is plausible as journalist Peter van Onselen (valonselenp) is more likely to be well-connected (proven by his high eigenvector score) despite having a lower degree.*

Id	Label	Interval	isposter	In-Degree	Out-Degree	Degree	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality
n148	scottmorrisonmp		☑	3	0	3	10.0	0.207048	0.255957	273.0
n57	australianlabor		☑	2	0	2	8.0	0.317568	0.399747	984.0
n69	cporterwa		☑	1	0	1	12.0	0.147799	0.172223	0.0
n53	albomp		☑	1	0	1	10.0	0.204348	0.243545	0.0
n101	liberalaus		☑	1	0	1	10.0	0.204348	0.243545	0.0

Figure 9. (above) The dataset of well-known figures (top to bottom) Scott Morrison, the Australian Labor Party, Christian Porter, Anthony Albanese and the Australian Liberal Party. While being influential in the media, these users scored reasonably low in degree, betweenness and eigenvector, indicating that the dataset has its limitations and social networks do not always represent the greater community off Twitter.

5