

PART II: BRACKENFELL COMMUNITY ANALYSIS

Community detection in network theory is a graph partitioning problem where a group of nodes is identified such that the connections a node has within a group are more than the connections a node has outside a group.

Data Source

#Brackenfell Reply Network - Largest Component

Data Transformation

Community detection techniques are typically applied to undirected networks. The direction in which a tie is sent or received is irrelevant; a connection is established as long as there is an interaction between users. The current network is therefore converted to an undirected network for the purposes of this analysis.

Method

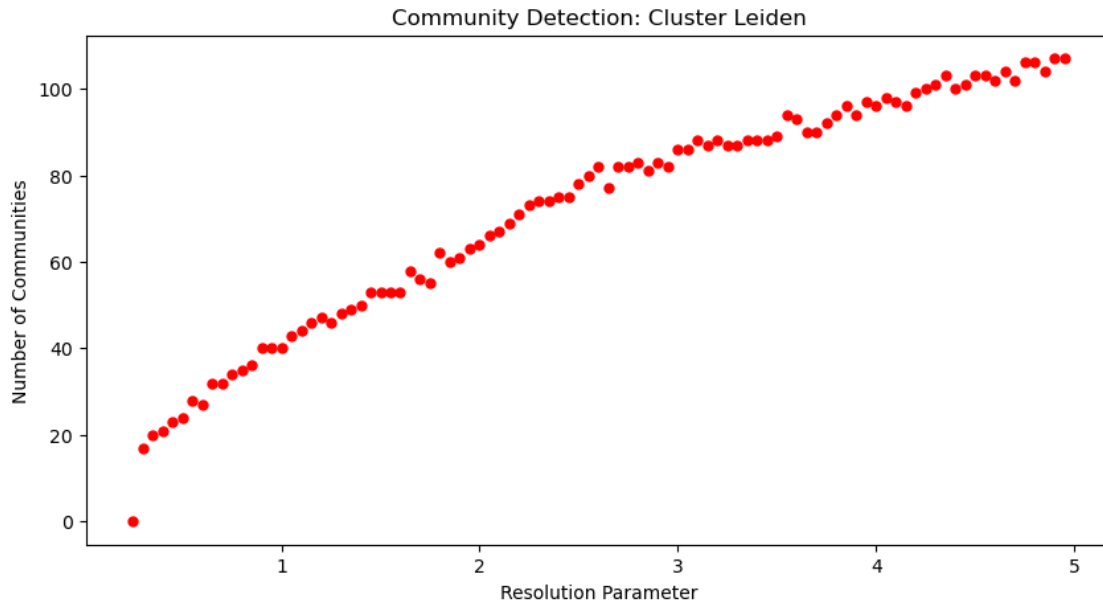
Using python's iGraph library, I apply the **Leiden** method to this community detection problem.

Undirected Graph:

IGRAPH U-W- 1578 1844

1 OPTIMISATION

To choose an optimal resolution parameter (gamma), I estimate the leiden algorithm and explore community outcomes across different values of gamma. Letting the data tell me which points are more optimal than others, I select gamma at the points in the graph where the number of communities detected reach a constant plateau or flatness.



2 DETECTION

Setting $\gamma = 0.8$, I detect 34 communities with a modularity of 0.84. The detected communities are visualised below along with descriptions of the top 3 largest communities (by weighted degree).

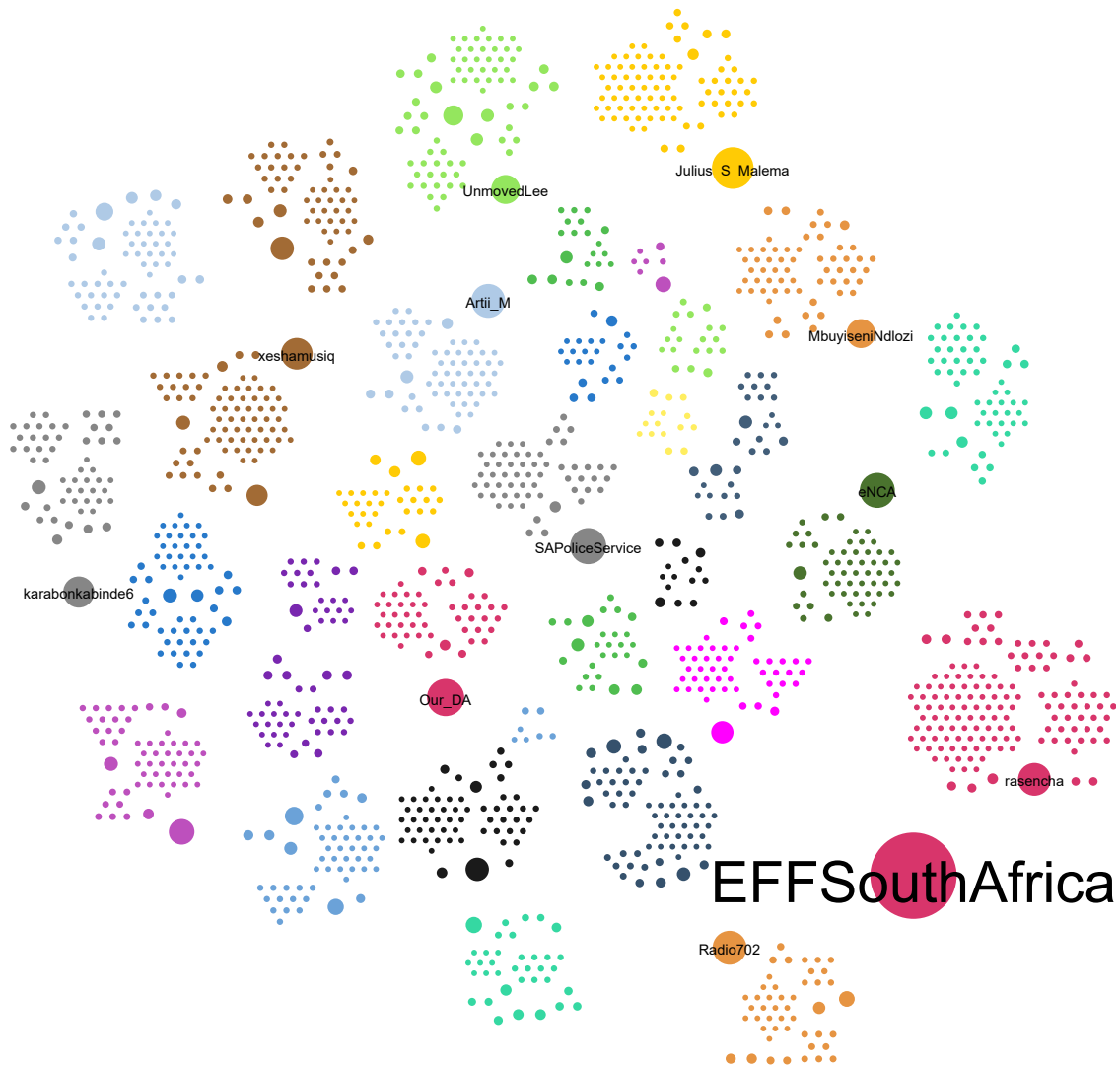
Communities:

34

Modularity:

0.8440284229710264

Brackenfell Communities:



Top 3 Weighted Degree Communities:

Community 1:

Size:
 141

Density:

0.016109422492401215

Average Degree:

2.26

Total Degree:

318

Members (10*) :

0 EFFSouthAfrica
1 rasencha
2 citrusramaphosa
3 Matewis_Dyches
4 IvorChalmers_
5 Clint_ZA
6 rubyjoms
7 Dumzie2
8 EffRonny
9 stormingnormanp

Community 2:

Size:

73

Density:

0.03538812785388128

Average Degree:

2.55

Total Degree:

186

Members (10*) :

0 UnmovedLee
1 ewnreporter
2 shroompopsicle
3 TeamNews24
4 wesleyfestersa
5 mjayijayi
6 Macfarlane123
7 HonourableHloni
8 Thandol76937181
9 SA_Mum

Community 3:

Size:

70

Density:

0.03850931677018633

Average Degree:

2.66

Total Degree:

186

Members (10*) :

0 lilanichlsn
1 RenaldoGouws
2 Mellow_Rocker
3 alfred_cabonena
4 ThuliMadonsela3
5 ramalokot
6 SimonPGrindrod
7 Om3rta_XIII
8 sick6_six
9 nash_cold

*Top 10 highest degree in community.

3 EVALUATION

Using various pair counting scores, I compare community outcomes of the leiden model at $\gamma=0.8$ and $\gamma=1.5$, and compare the leiden model at $\gamma=0.8$ and the louvain model.

3.1 Leiden ($\gamma=0.8$) vs Leiden ($\gamma=1.5$)

Leiden ($\gamma=1.5$)

Communities:

47

Modularity:
0.8420700374821725

Comparison Scores:

Variation of Information:
0.8523313862639712

Adjusted Rand:
0.6703353281906859

Normalized Mutual Information:
0.8523313862639712

3.2 Leiden (Gamma=0.8) vs Louvain

Louvain

Communities:
37

Modularity:
0.8431499313569365

Comparison Scores:

Variation of Information:
0.8854463301112254

Adjusted Rand:
0.7728835393099504

Normalized Mutual Information:
0.8854463301112254

4 SUMMARY

In this section, I use community detection methods to identify densely connected groups and communities present in the #Senekal twitter reply network.

When searching for optimum modularity, I find the resolution parameter at $\gamma=0.8$ and $\gamma=1.5$ robust. I detect 34 and 47 communities with a modularity score of 0.844 and 0.842 at these points, respectively. Similarly, using the louvain method, 37 communities are detected with a modularity score of 0.843. The coincidence scores from the evaluation and the modularity scores, which are above the 0.7 threshold, are all a sign of a well partitioned network. However, given the large number of communities detected, a reasonable or meaningful interpretation is difficult. In other words, attempting to characterise all 34 communities or predict each of their behaviour and influence in the discourse based on their memberships will unlikely result in an intuitive outcome.

To atleast get some sense of the communities detected, I examine the sizes, densities, and degrees of the top 3 largest communities as well as the top 10 members (sorted by *weighted degree*) assigned to each of the 3 communities. Going by the users and tweet content, **Community 1** and **Community 2** appear balanced in terms of political or other affiliations as well as members in support or opposition of the protest, while users in **Community 3** seem to share or echo a similar rhetoric and language in opposition of the movement. Evidence of this is not conclusive, however.

In the next section, I classify users by the positions or roles they play within the discourse based on their patterns of interactions and other network features.