

PART I: BRACKENFELL

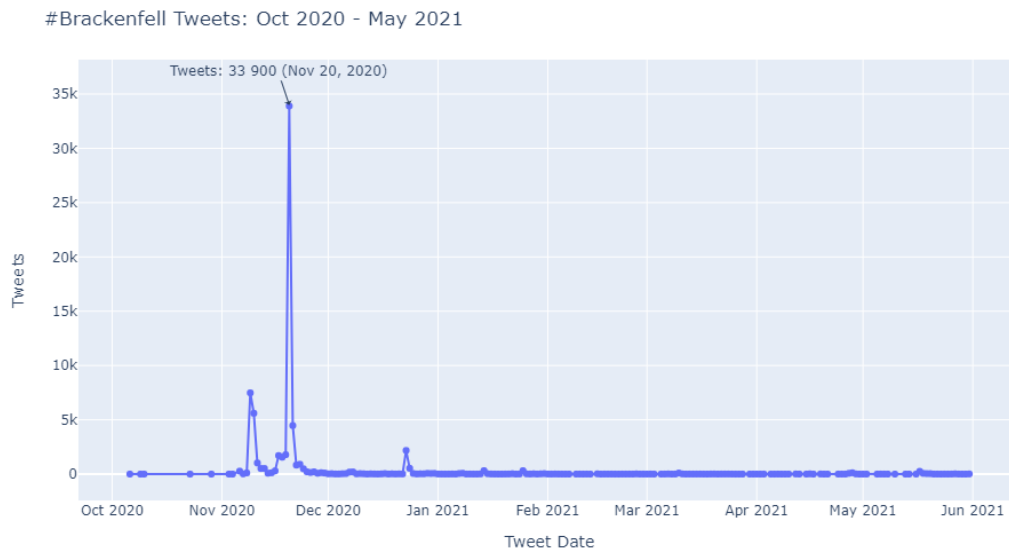
DESCRIPTIVE ANALYSIS

The #Brackenfell protests occurred in November 2020 in the northern suburbs of Brackenfell, Cape Town, South Africa. These protests, organised by both the EFF and the PAC, mobilised around allegations that a parent-organised function for Brackenfell High School students on 17 October 2020 was only open to whites at the school, that the presence of two teachers at this function was an indication of school support for the event, and that this was symptomatic of systemic racism at the school (and by implication within the broader community). These protests led to clashes with the mainly white parents of children at the school as well counter-protests.

Data Source

Data is extracted from Twitter's API matching the following search query: “#EFFinBrackenfell” and “EFF Brackenfell”, for the period: 1 October 2020 to 31 May 2021.

Number of Daily Tweets in #Brackenfell Discourse:



Data Transformation

The extracted data set consists of tweets, retweets and replies to tweets related to this movement, however for the purposes of this analysis, I focus purely on direct engagements and interactions between users (i.e. replies to tweets). To construct this type of network, I only include data where a user tweets a reply (sender) to a user whose tweet was replied to (receiver).

To ensure that the transformed data still reflects the full discourse and no significant information is lost as

a result of the conversion from full network to reply network, I derive aggregate statistics of the excluded tweet and retweet data. In this way, the properties of the #Brackenfell full network are still described in the #Brackenfell reply network. See below for a list of derived variables before transformation.

- **Derived node variables:**

- *#tweeted* -> A measure of the total number of tweets by a user in the #Brackenfell network.
- *#retweeted* -> A measure of the total number of tweets by a user that are retweeted by other users in the #Brackenfell network.
- *#liked* -> A measure of the total number of tweets by a user that are liked by other users in the #Brackenfell network.

- **Derived edge variables**

- *weight* -> This captures the total number of replies from a sender to a receiver in the #Brackenfell network. For example, replying to a user only once gives a weight = 1, but replying to a user's tweets multiple times gives a weight = number of replies to that user.

Once variables are derived, the extracted dataset is converted to a directed network, which I call the #Brackenfell Full Network.

- **#Brackenfell Full Network**

The #Brackenfell full network consists of 68746 tweets (incl. retweets, and replies to tweets) with 21796 active users behind those tweets.

- Number of users: 21796
- Number of tweets: 68746

After dropping all tweets and retweets that do not have a reply, the #Brackenfell Full Network is then converted to the #Brackenfell Reply Network.

- **#Brackenfell Reply Network**

The #Brackenfell reply network includes data on users replying to tweets and users whose tweets are being replied to. This is a weighted network where weights measure the frequency of interaction, i.e. the number of times a user replies to the same users tweets.

- Number of users: 2369
- Number of replies: 3025

Next, I exclude all disconnected components and focus my analysis on the #Brackenfell Reply Network - Largest Connected Component. This sub-network is found by extracting the component with the largest number of connected nodes in the network. To explain, a connected component is a network or sub-network where every node is atleast connected to one other node, i.e. there must exist a path such that any two nodes can be connected starting from point *A* to point *B*.

- **#Brackenfell Reply Network - Largest Connected Component**

The #Brackenfell reply network has 356 connected components. In this network, the largest connected component has 1578 nodes and 2116 edges, and the second largest connected component has 26 nodes and 28 edges.

- Number of users: 1578
- Number of replies: 2376

Method

Using python's iGraph library, I describe the #Brackenfell Reply Network - Largest Connected Component in terms of its **centrality** and **connectivity and cohesion** properties.

Directed Network:

IGRAPH DNW- 1578 2116

1 CENTRALITY

1.1 Node Centrality

To identify key or “important” node’s, I determine the centrality of a node relative to other nodes in a network.

Node Centrality: Summary Statistics

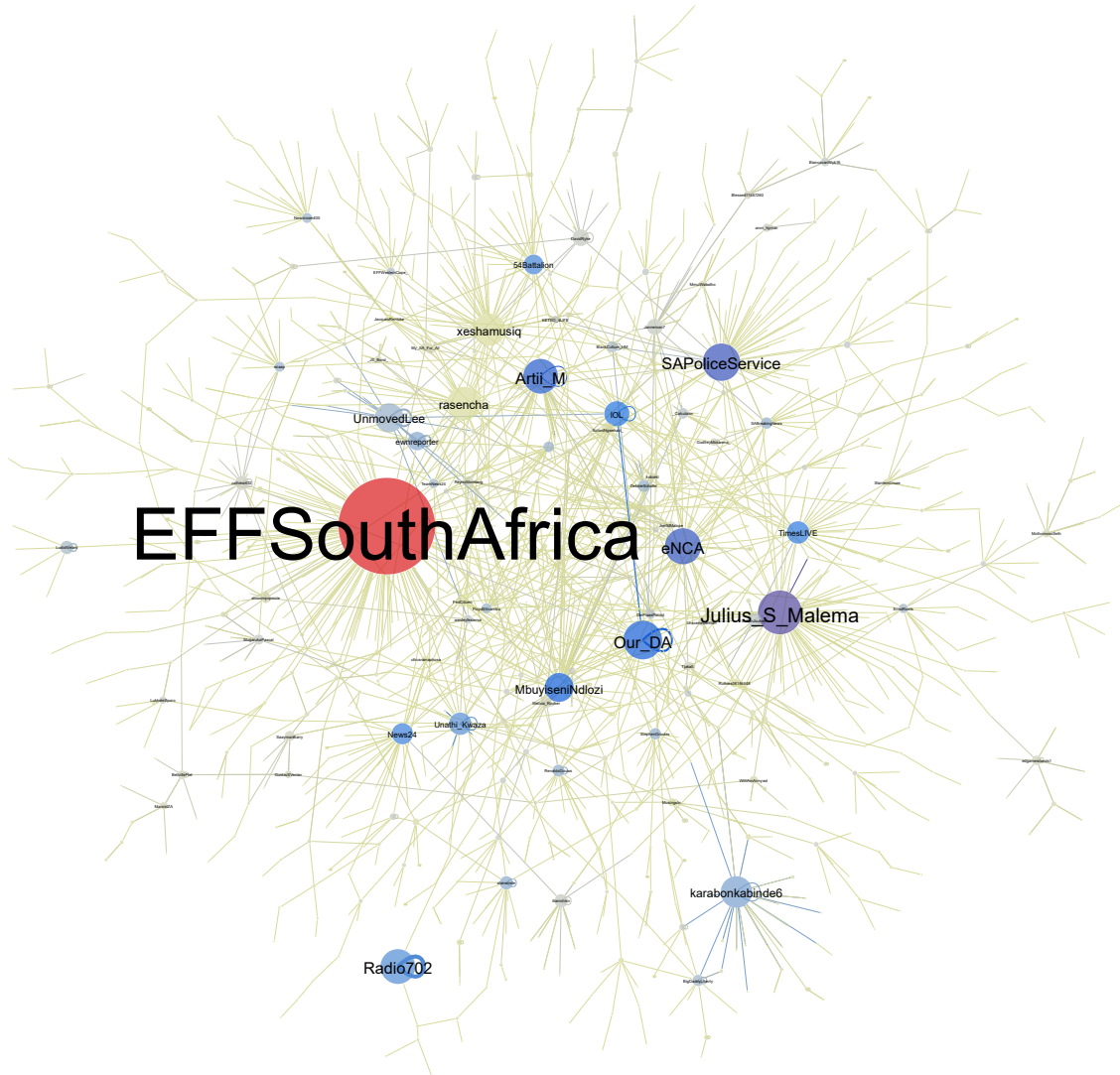
	indegree	outdegree	closeness	betweenness	pagerank_
↔ \					
count	1578.000000	1578.000000	1578.000000	1578.000000	1578.000000
mean	1.505703	1.505703	0.161173	8.799676	0.000634
std	5.555647	2.630060	0.025760	56.537619	0.002159
min	0.000000	0.000000	0.080959	0.000000	0.000223
25%	0.000000	1.000000	0.142583	0.000000	0.000223
50%	1.000000	1.000000	0.162243	0.000000	0.000245
75%	1.000000	2.000000	0.184731	0.000000	0.000455
max	147.000000	50.000000	0.240360	1019.000000	0.060501

	eccentricity
count	1578.000000
mean	11.406844
std	1.170014
min	9.000000

25%	11.000000
50%	11.000000
75%	12.000000
max	16.000000

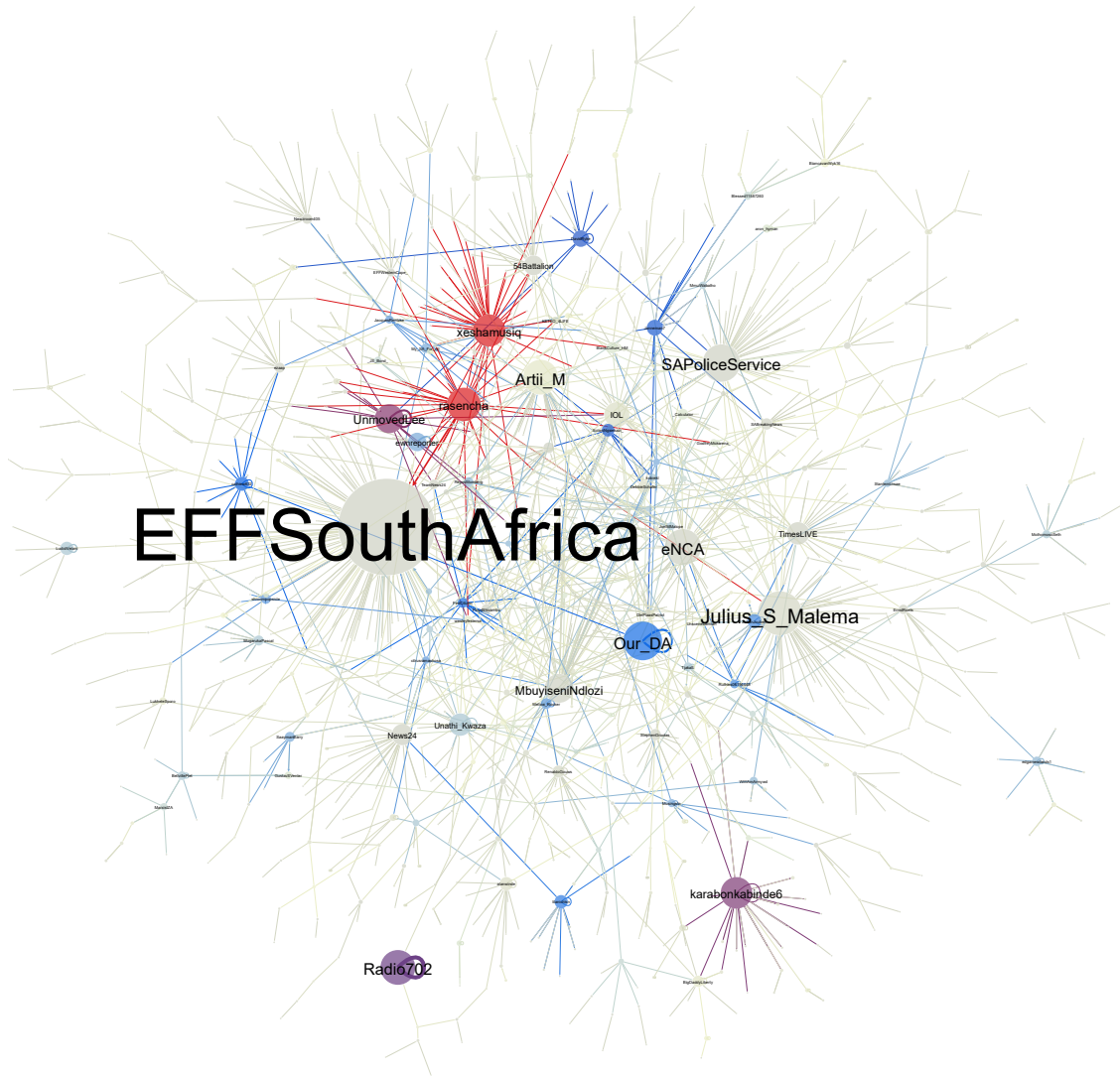
Maximum Indegree:

EFFSouthAfrica



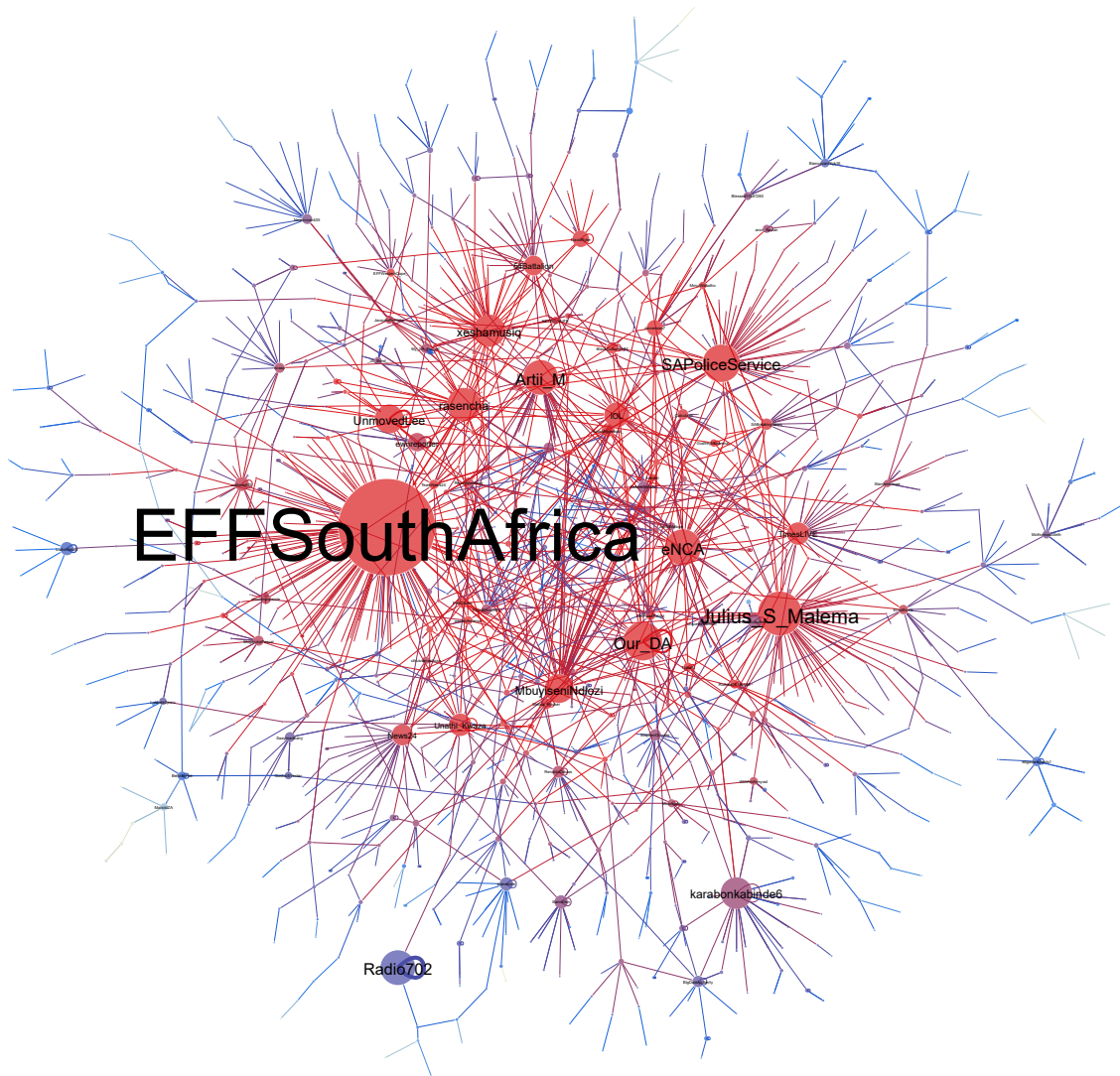
Maximum Outdegree:

rasencha



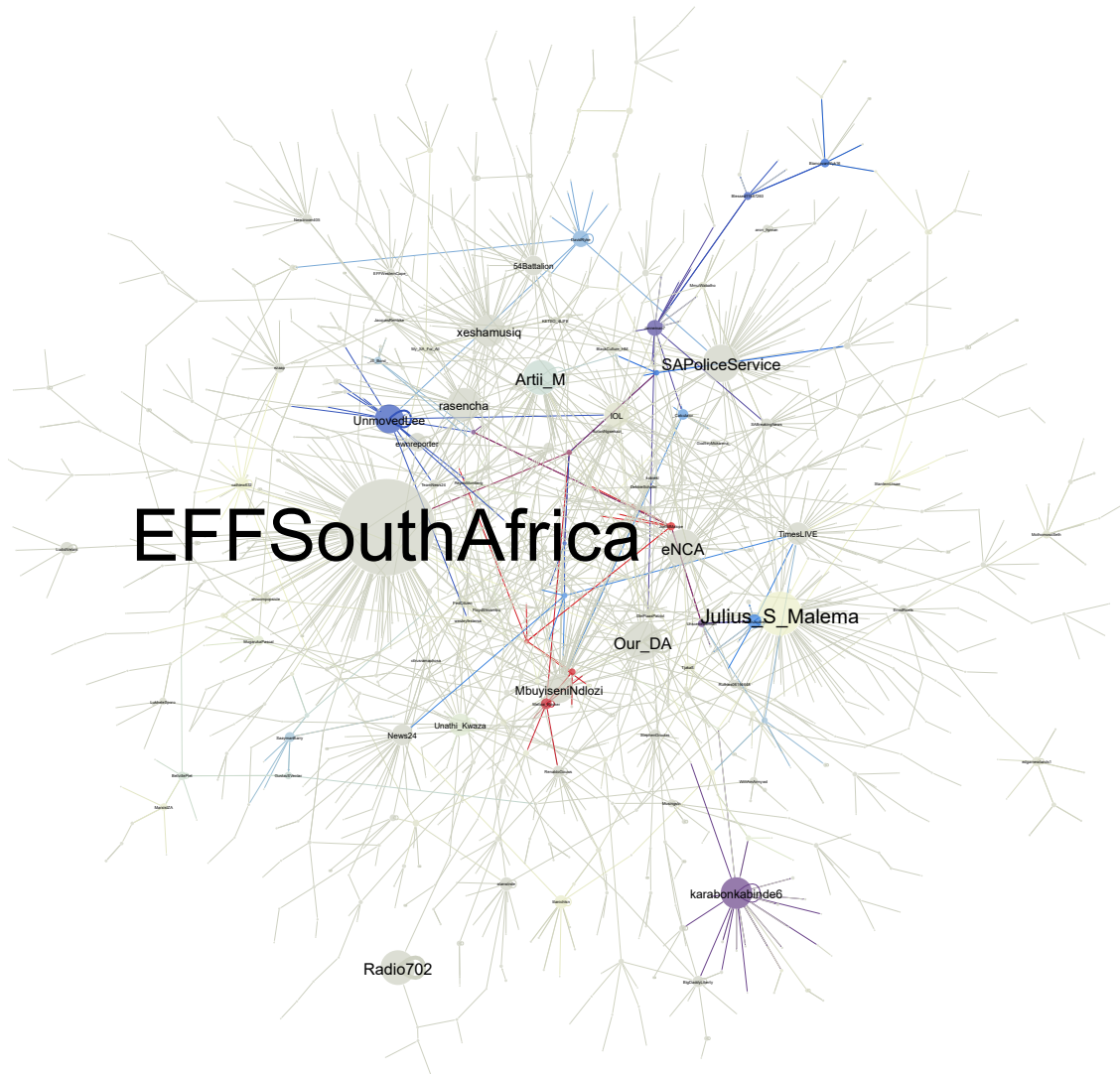
Maximum Closeness:

EFFSouthAfrica



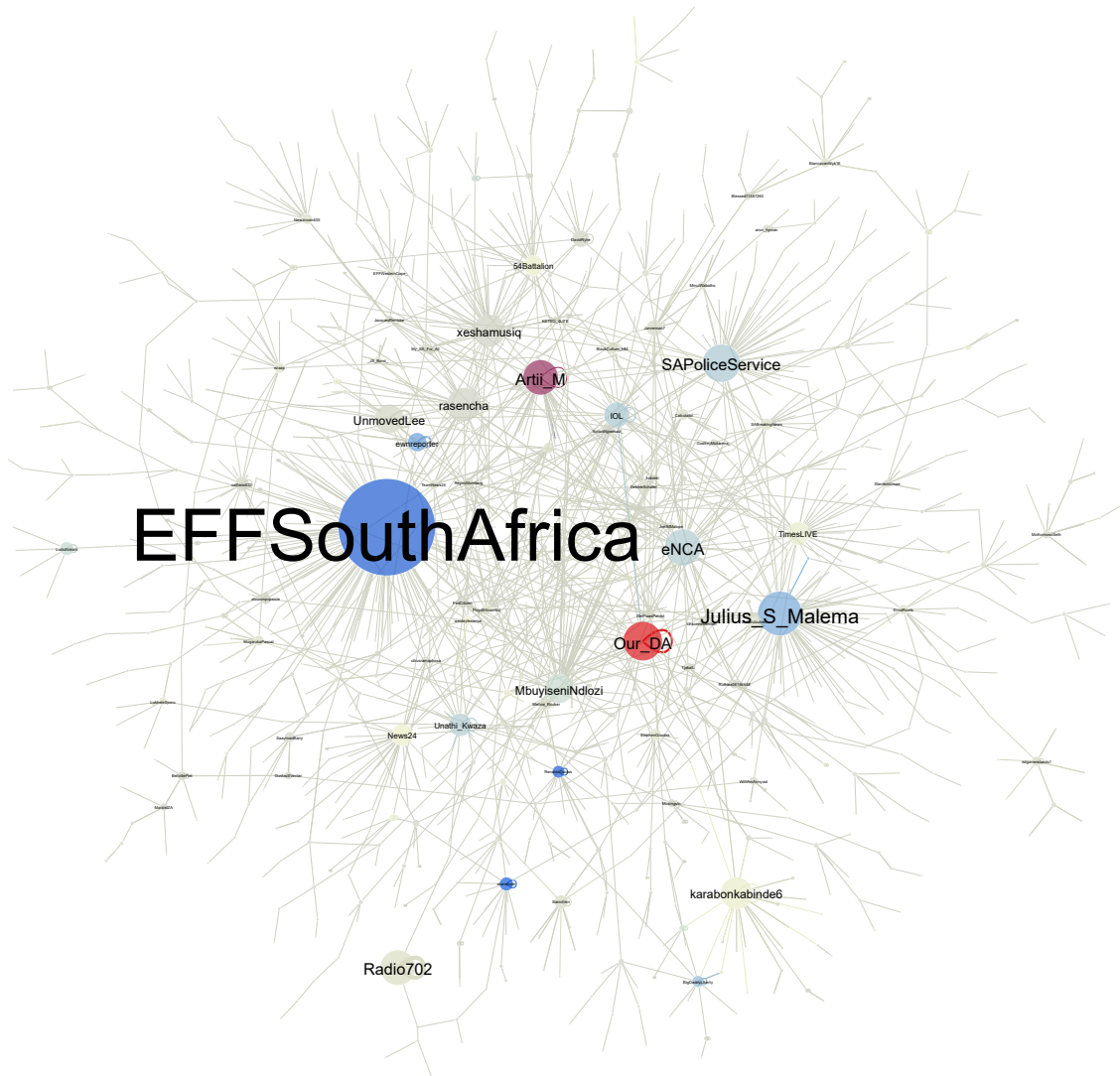
Maximum Betweeness:

JoelMMalope



Maximum PageRank:

Our_DA



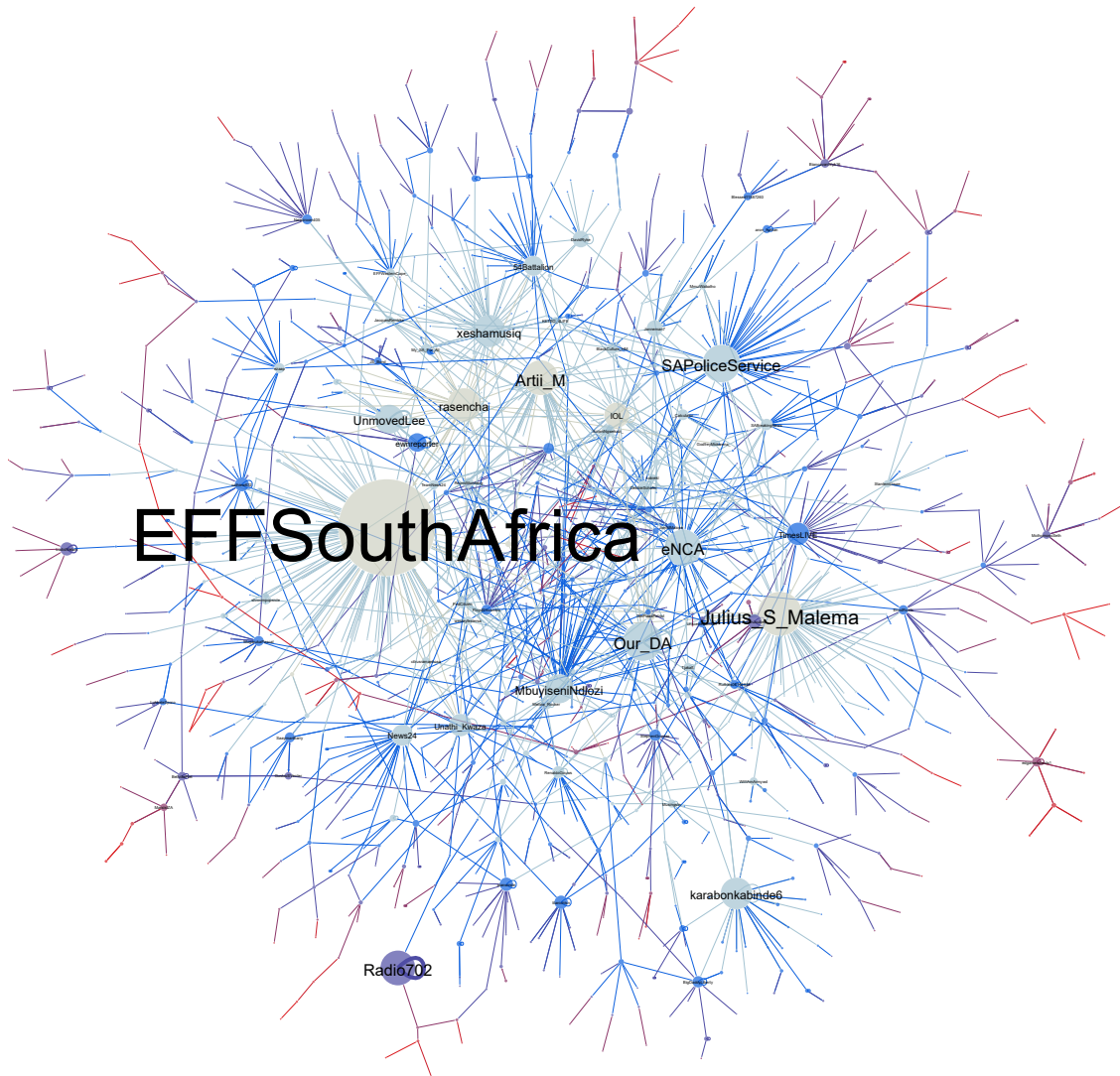
Minimum Eccentricity:

	label
0	EFFSouthAfrica
1	DiePlaatPatriot
2	TheReality1
3	Julius_S_Malema
4	Artii_M
5	Dennistalksfact
6	PearcesVoice
7	IOL
8	Matewis_Dyches


```

9      ZakirahTriveni
10     Rosemary_Hardy
11         rasencha
12         gee_juz
13     mfanafuthishaba

```

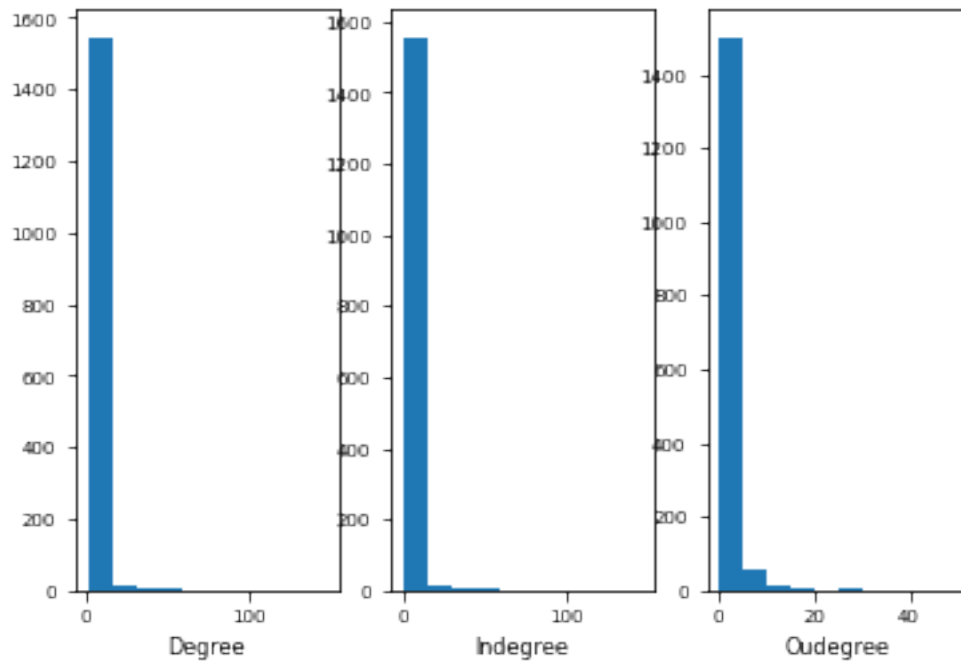


1.2 Network Centrality

To determine relationships between nodes and their position in the network, I measure a node's ties relative to the ties present in the network and the distribution of ties throughout the network.

Network Centrality: Summary Statistics

Descriptive Analysis: Degree Distribution



Average Degree:

3.011406844106464

Density:

0.0008503093824165985

Average Path Length:

2.9253503538226724

2 CONNECTIVITY AND COHESION

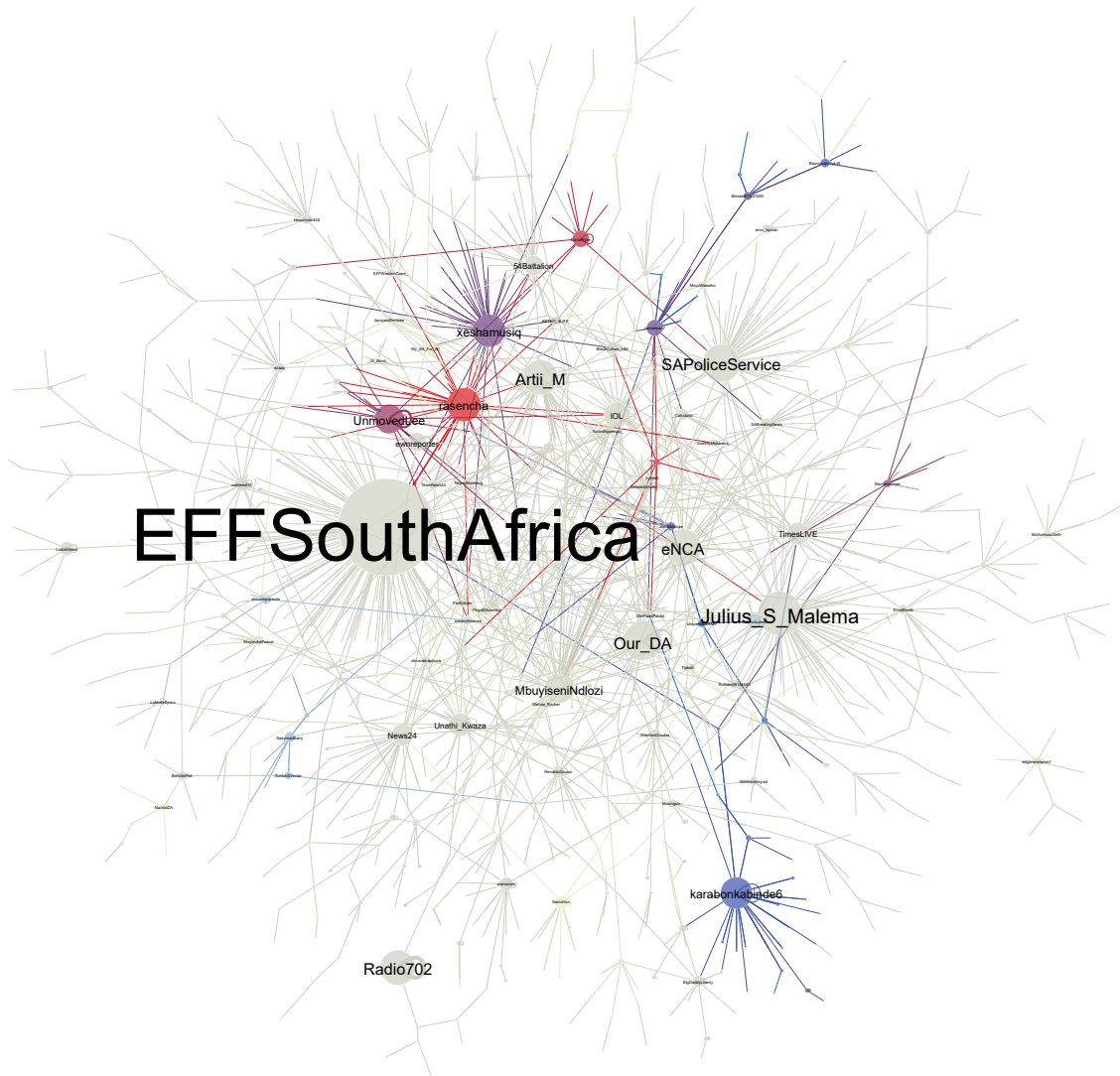
To examine how tightly connected or clustered the network is, I examine the direction, frequency and consistency of relations between nodes and the nodes in their neighbourhood.

Connectivity and Cohesion: Summary Statistics

	reciprocity	transitivity	hierarchy
count	1578.000000	1578.000000	1578.000000
mean	9.060837	0.050697	3.557034
std	41.176039	0.366602	30.735183
min	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000
75%	0.000000	0.000000	1.000000
max	597.000000	10.000000	940.000000

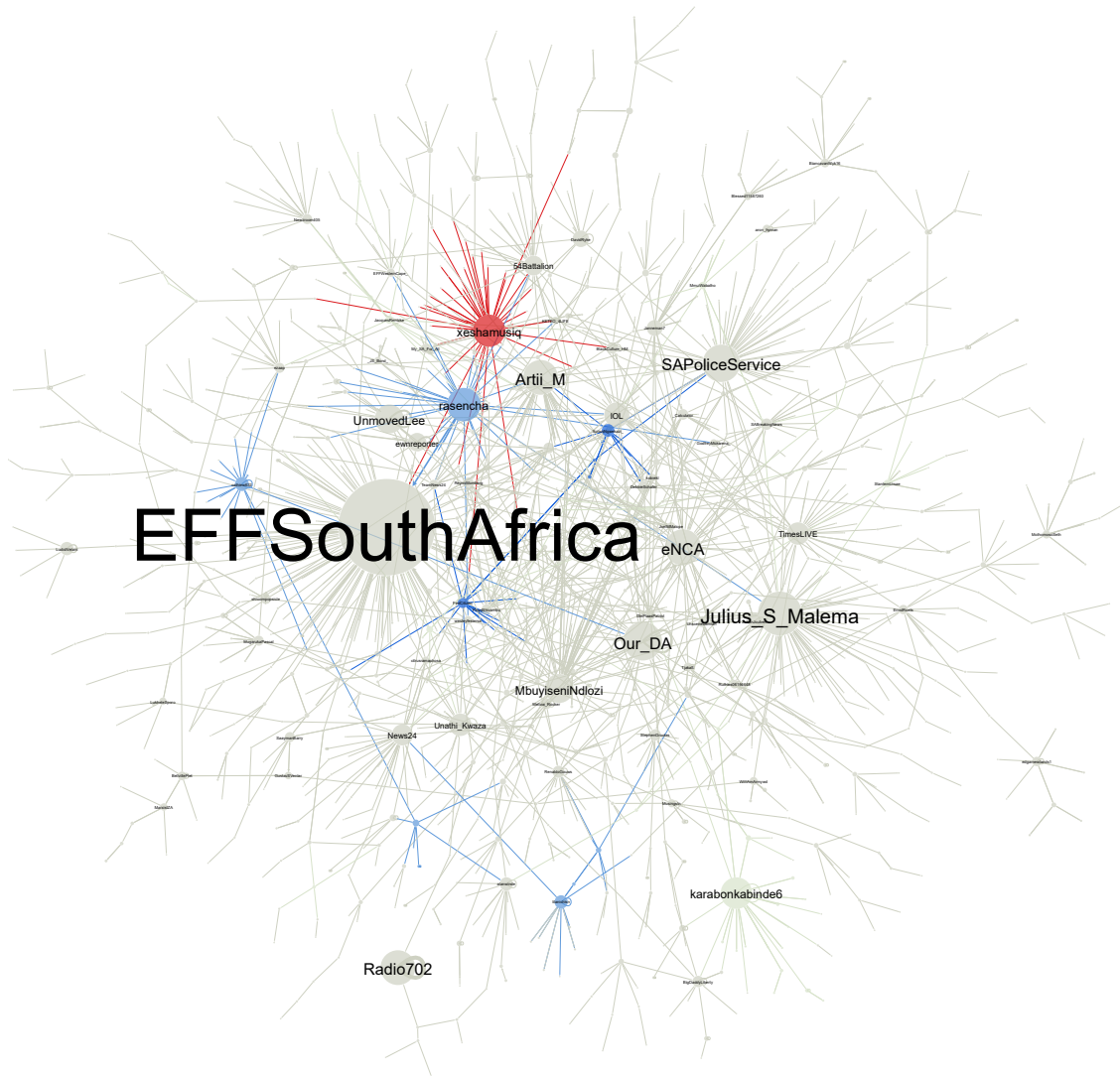
Maximum Reciprocity:

rasencha



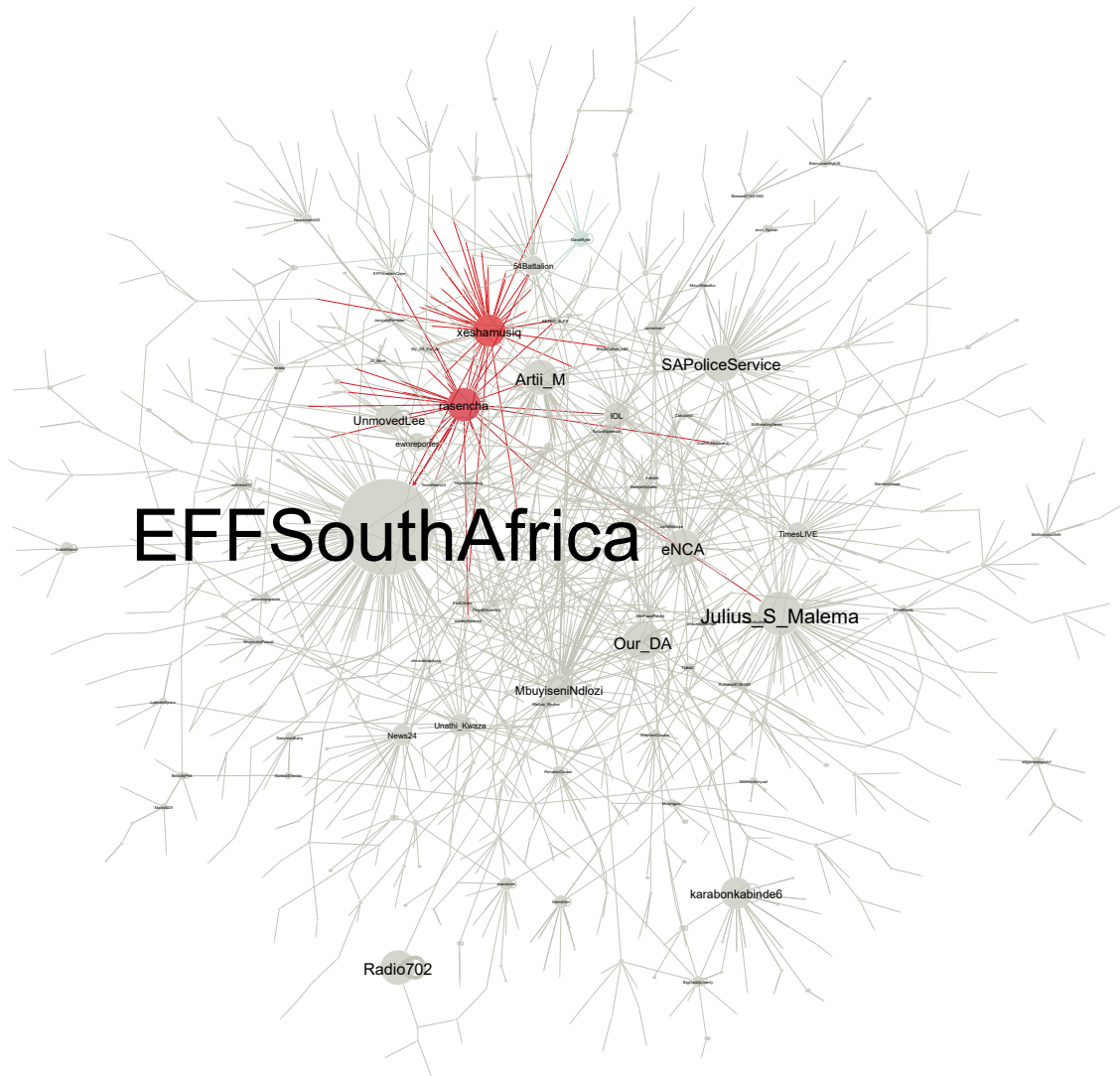
Maximum Transitivity:

xeshamusiq



Maximum Hierarchy:

xeshamusiq



Average Clustering Coefficient:

0.04255027112376041

3 EVALUATION

To determine if the current network exhibits small world properties, I simulate multiple random networks and compare the average distances and clustering features of either network.

Random Network:

Average Path Length:

16.67812299194082

Average Clustering Coefficient:

0.001656284988354186

4 SUMMARY

In this section, I perform descriptive analysis on the #Brackenfell twitter reply network to determine certain node features and network properties that identify significant players in the discourse. I summarise a few findings below.

In addition to skewed or unequal degree distributions, which indicates a tendency of preferential attachment toward popular users, this network exhibits high clustering (26.1x higher) and small average distances (5.8x smaller) than randomly simulated networks of the same size. Consistent with the impression of small worlds.

Twitter accounts found to rank highly on certain network measures are listed below:

- **EFFSouthAfrica** (Indegree, Closeness)
- **Our_DA** (PageRank)
- **JoelMMalope** (Betweenness)
- **rasencha** (Outdegree, Reciprocity)
- **xeshamusiq** (Transitivity, Hierarchy)

Other than the usual suspects, the user accounts **rasencha** and **xeshamusiq** are notable in measures of *reciprocity*, *transitivity* and *hierarchy*. After examining the data further, it appears that in an attempt to promote their music, the user account **xeshamusiq** was targeting and trolling prominent users within the discourse by actively replying to their tweets asking them to retweet and like his pinned tweet.



The twitter account **rasencha**, on the other hand, appears to actively be involved in the discourse and has formed a neighbourhood that includes other highly responsive users. Currently the user behind this account has been suspended off twitter, likely due to the racially discriminatory undertones of their tweets.

← **Tweet**



Rasencha Mohale
@rasencha



Replying to @ReneMalika2

Even today at #EFFinBrackenfell police are taking order from white thugs



In the next section, I explore the network further by using community detection methods to identify groups (communities) of users that are densely connected and have high levels of interaction within the discourse.