HERE BE DRAGONS

Issued by the Centre for Analytics and Behavioural Change

17 August 2020

Interim report on Xenophobia on South African Social Media:

Investigating the content, users, and tracking the intentional promotion of xenophobic conversation.

From 1 April to 31 May 2020

Email:info@cabc.org.za

Background: Promotional accounts on Twitter

Promo Twitter is used to describe accounts or people who partner with brands/celebrities/authors to market or promote their products on the user's timeline. The accounts will then tweet/share pictures and talk about the product until it trends. For a topic to trend, the topic must be tweeted in a short period of time by a sufficiently large number of people: you will need about 10, preferably more, people tweeting and retweeting each other.

For commercial purposes, celebrities are briefed together on what they will be tweeting and retweeting, from hashtags to copy, creative and other accounts they should be engaging with. The inviting tone of the post/copy, imagery and the similar tags and hashtags. These are industry tried and tested methods that have worked to launch music albums, new TV shows or new KFC lunch meals.

Coordinating twitter activity is, however, not limited to benign product promotions or advertising campaigns. Any users can coordinate among themselves to promote a topic beyond its actual reach, be it racism, homophobia, or in this case xenophobia. More problematically, users can avoid using their own accounts, and can instead buy defunct accounts or pay individuals to spread xenophobic content. This makes it easy to hide your identity and your agenda and it allows for a manipulation of the conversation over a longer period. Unlike a one-day campaign, owning accounts means they can be used as frequently as necessary to promote specific content.

During the deep dives into social media data, it soon became clear that there are accounts that have the hallmarks of being bought or paid accounts: almost all their content was tweeting or retweeting xenophobic content.

Isolating the accounts

Initially analysts isolated sample accounts that were presenting as inauthentic and had signs of planned manipulation. Criteria were identified to define suspicious accounts from data that was enriched using a software programme called Brandwatch. A further investigation was done to establish what agenda was being pushed. Some of the criteria included:

- Tweet Volume: People are on Twitter to discuss and follow their interests such as soccer or media, so their page will mirror that. There will be a theme or that "one thing" about the account. There will be tweets and retweets from like-minded accounts. However, if the accounts tweets/retweets only xenophobic or racist tweets and nothing else, there is usually an agenda behind the account. The volume of tweets/retweets of xenophobic content per day was assessed amongst other activity

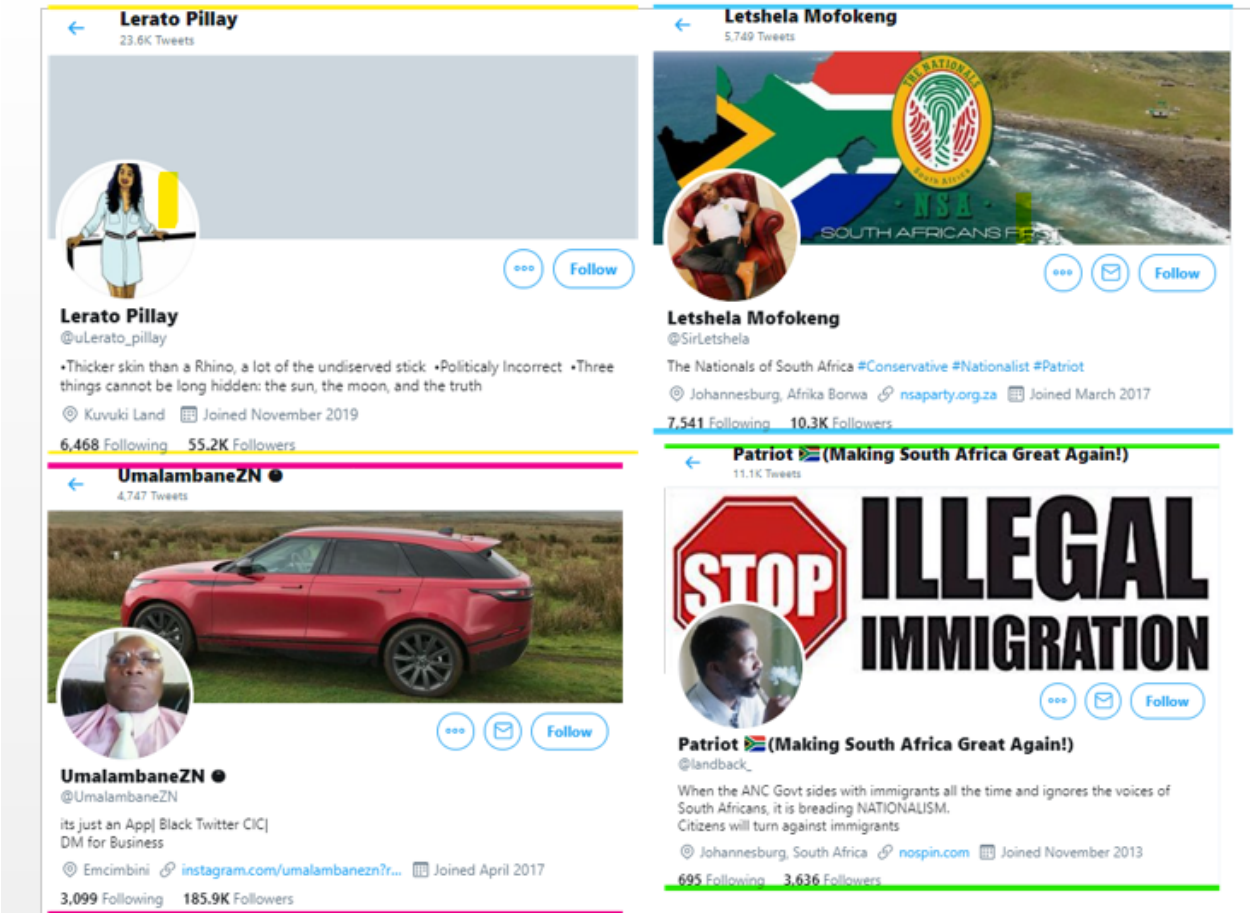
on their timeline. If the account has a higher volume of xenophobic content on their timeline and nothing else, there is something suspicious about the account.

- Political Posts: This political agenda was explored further. This helped to identify if their agenda is right wing or left wing, or if the account is promoting anti-government sentiment. From this an assessment could be made about the account holder's possible political affiliation and whether they were, for instance, anti-ANC, anti-EFF, anti-DA, pro-Trump, etc.
- Connections: These are important criteria because it indicates who is playing from the same team or playbook. The leaders of the movement become evident and the handles of the frequently tweeted (ie: @handle, @handle, @handle)
- Other aspects of the accounts that present as suspicious include:
    - Lack of personal tweets;
    - Retweets only;
    - Incongruent (The dissociation between the profile and tweets/timeline);
    - Xenophobic agenda; and
    - Mostly political content.

The accounts were grouped based on particular characteristics, such as whether the account appears to be professional, or has a particular political agenda. (The list is available on request.)

Who engages who in this conversation:

Popular or influential users tweet their status and are retweeted, mentioned, or replied to by their audience. To break down the criteria further, it was decided to identify the popular accounts in this conversation. These are the accounts that have the most followers and who get the most engagements. It also provides an idea of where the content and traffic is generated from. This directed analysts to the smaller and newer accounts. The screenshots below are of the 3 biggest accounts, by follower size and leadership roles in the xenophobia conversation.



Analysts went through each timeline to see which other accounts on the suspicious list they engage with and the kind of engagement they have.

Things that were considered included the number of followers, number of tweets, reach (how many times a tweet from this accounts get retweeted/liked or replied to) when the account started and if the account has more xenophobic content than any other content. The table below shows the findings:

@ulerato_pillay @umalambanezn	**Can't distinguish a connection**	**@SirLetshela**
@ITKHUMALO @SANEEDSAWALL @Lerato_More @mthosovilakaz @blackieEaster	@FelizBslrallele @sa_illegal @GillianGoldman @zmobilefashion @MakopoK	@landback_ @MarioKhumalo @OwaseMzantsi @Thelandthief @VictoriaAfrica4

@MissB20_10 @Le2X3 (another lerato account, suspended now) @MaleselaMatlou @BlessingsRamoba @KaraboKhama @AlomNyama @LorenzoSander12 @Savali11608784 @metse_keitu @mudzu_thabe	@Motheo2009 @Nthapeleng_ @Laetitiagg2 @AfricaFacts7 @MelusiTweets @Bally12389363 **@ImmigrantsGo**	@victoriaafrica3 @RambauRicky @uncaged_gal @Cindilopa @GImpisi @SAF_RSA @BhekabantuShab2 @GuguSibeko @VictorGreatree @1Grootman (uses a new handle@PutSAfnsFirs @NewErraSA @Trump2020Black @Laetitiagg2 @thestirrerhagar @Commander_C007 @G_Manchidi

- The accounts/handles marked in red are new accounts that were created between January 2020 to May 2020. There was also a noticeable spike in the xenophobia dialogue from March 2020. What is interesting to see with these new accounts, is the obvious similarity in the tone and online behaviour with some of the bigger suspicious accounts.
- The accounts/handles marked in blue were created years ago and were dormant until November 2019 or March 2020. Those two dates are significant because November 2019 was the inception of @uLerato_Pillay account and March 2020 marks the spike in the xenophobia conversation in South African Twitter.

Influential accounts that we should not overlook

Below are a selection of accounts that have supported the hashtag, #PutSouthAfricaFirst by retweeting some of their content. The @uLerato_Pillay network tries to use the reach of these individuals to grow their own network.

- @ZungulaVuyo - President of ATM and Member of Parliament.
- @HermanMashaba - Former Mayor of Johannesburg.
- @VusiThembekwayo - Businessman and Motivational Speaker.
- @BlessingsRamoba - claims to be the President of the Mining Forum of South Africa (no website for the said Forum) and no information about it on the internet to back up its legitimacy. Except on Blessings Ramoba' social media accounts.

Sample of content sharing and recycling

For another layer of evidence that these accounts may be connected, we looked at the content they shared. In most cases there were similarities with the content shared including pictures, videos, copy and captions. Most of these accounts are basically

posting the same content and retweeting each other. For example, the popular account would share a post, a week later one of the other smaller accounts would post the same tweet without even changing a word. In some weeks they will even use the same engagement methods, like creating a poll or a follow for follow trend.

Recycling of content:

Same engagement method:

Familiarity and acknowledgement of each other:

Network Analysis

Data Collection and Methodology

In combination with the investigation into the specific accounts identified as suspicious, analysts also conducted a Network Analysis that attempted to understand the nature of the network of accounts in the overall xenophobia conversation. This analysis drew from the data acquired on Brandwatch that tracked all mentions of the hashtag, #PutSouthAfricaFirst, which was used as a proxy for identifying xenophobic content given how prevalent it was within the xenophobia conversation. According to Brandwatch, the individual accounts that are most commonly tweeting this hashtag account for a significant proportion of all conversation identified as xenophobic on South African Twitter.

The data set was limited to twitter activity surrounding this hashtag from 1 April 2020 to 31 May 2020. All accounts that used this hashtag or retweeted a tweet during this period were assigned as nodes in the network and connections were drawn between them when one account retweeted a post from another. This created a map of the network that could be analysed or subdivided to understand the dynamics of those connections. Once the full network was rendered, the suspicious accounts identified above were isolated and the connections between them visualised. This analysis shows clear evidence of strong connections between these accounts, indicating that they are coordinating their activities.

Full Network Render:

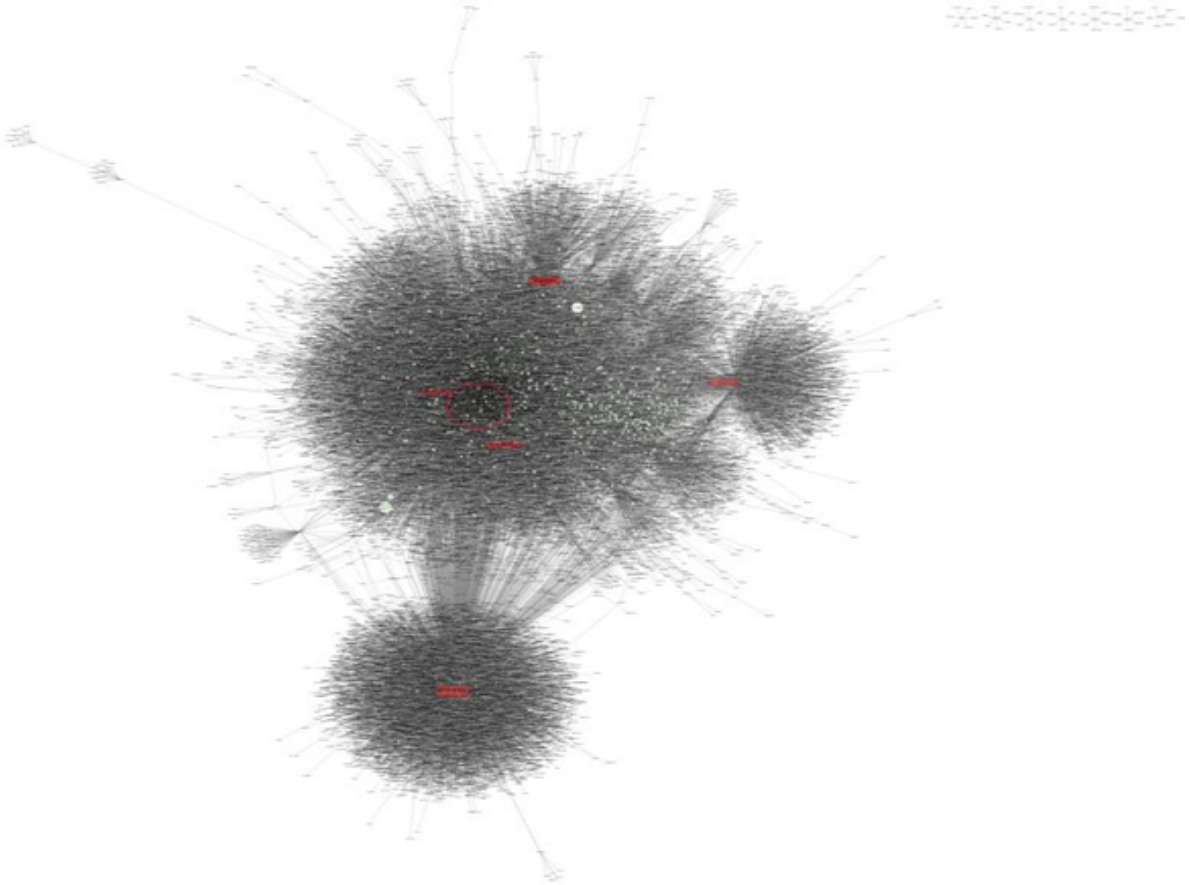


Image 1: #PutSouthAfricaFirst Conversation Map.

A render of the full network is shown in Image 1. This image displays all the users that were either tweeting the hashtag #PutSouthAfricaFirst or were tagged in tweets that used the hashtag. Two distinct clusters have been formed. The reason for this is that connections between nodes are not singular. Individuals who engage in a conversation will engage with others in that conversation, forming multiple connections. This creates a web of connections, that is more dense in the centre where the individuals with the most connections are located.

The existence of a second, smaller cluster is interesting because it has been created by a single account “tweetjacking” (or using a twitter hashtag to promote content unrelated to the conversation). Essentially, this single account is connected to others in the main conversation by retweeting their content, but their audience is not connected to that main conversation. Instead, they are only interacting with the tweetjacking account. As such, this cluster can be ignored, and the main cluster is the focus of the analysis.

Suspicious Accounts Cluster:

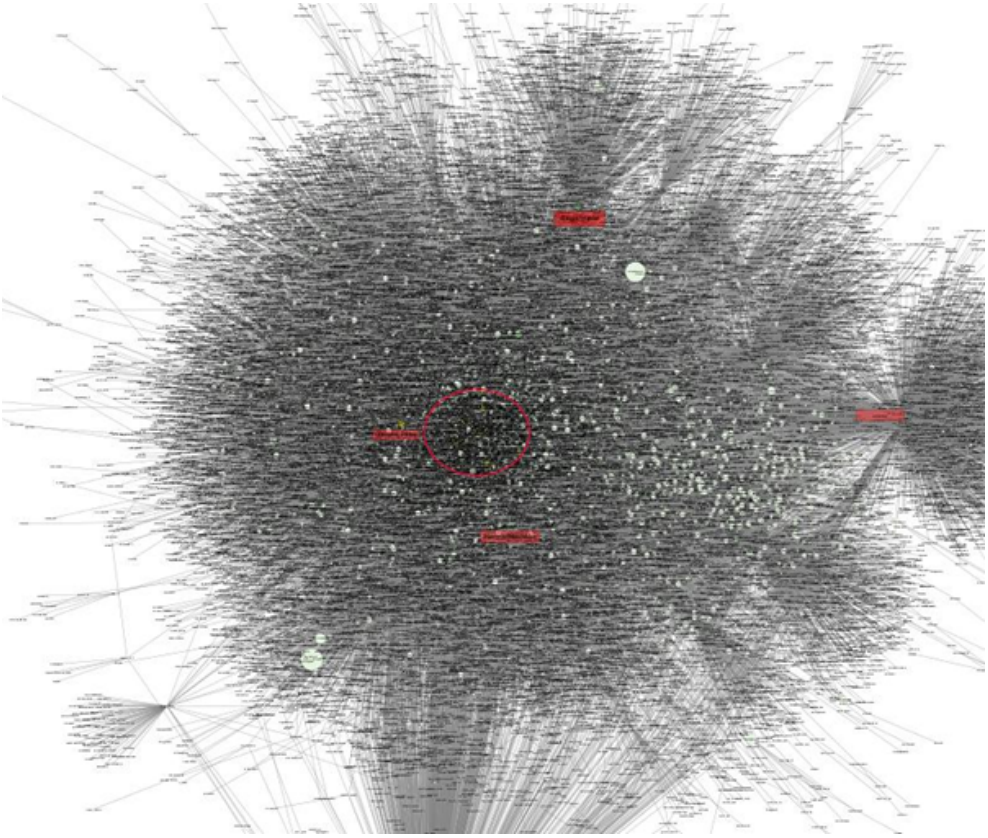


Image 2: #PutSouthAfricaFirst Conversation Map, suspicious account cluster isolated.

Within the network, the suspicious accounts are located within the centre of this larger cluster, and this is shown in Image 2 with the accounts located within the red ring. These accounts are driving a large volume of the xenophobic content within the conversation, and are highly connected to one another, indicated by the frequency of their interactions and retweets. To demonstrate this, the list of suspicious accounts was isolated to examine the specific connections between them.

This is shown in Image 3. The account uLerato_Pillay is deeply connected to the rest of this network, with almost every single account identified being connected to it. This is emphasised in Image 3 below which isolates uLerato_Pillay at the bottom of the map, as well as two other accounts that are particularly well connected: thelandthief at the top and landback_ on the right. This map indicates that these individuals are being retweeted or mentioned frequently by almost all the other accounts within the network. They are driving most of the activity, not only within this network of suspicious accounts, but in the conversation as a whole.

It is important to point out a limitation of this map at this point in the analysis. This data set, by its nature, is agnostic to the content of tweets. If an individual was to tweet at a member of the network “Good morning”, this would be counted as a connection - in the same way that a tweet urging action against Spur for hiring foreign workers would be counted. This is why it was necessary to combine the network data with the analysis of specific content of the accounts conducted by the CABC Dialogue Facilitation team. Because of that analysis, it can be concluded that these individuals are not only important within the network because of the extent of their connections, but they are also responsible for much of the original content that is being shared by accounts within this network.

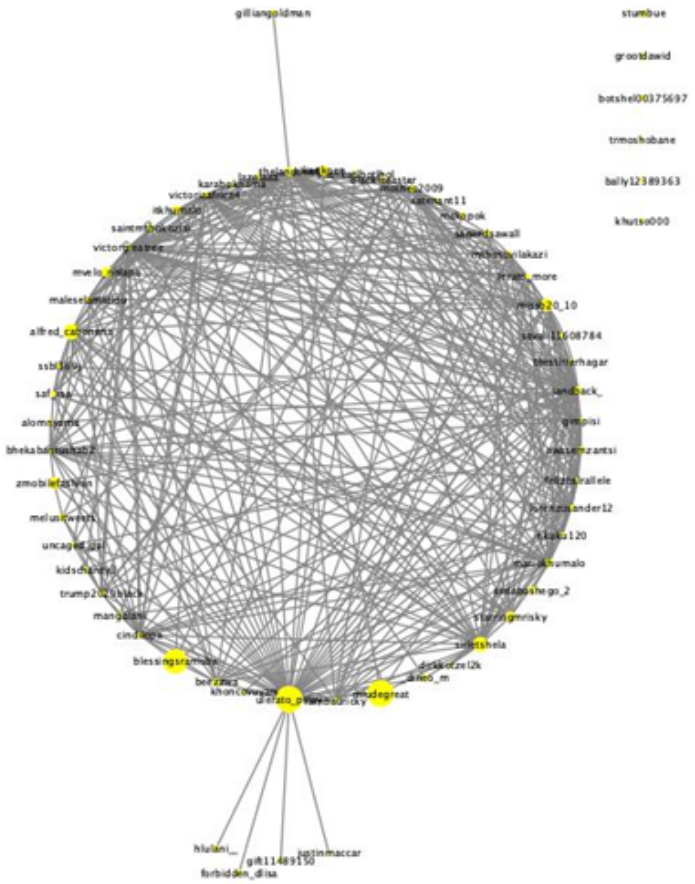


Image 3: Network of Suspicious Accounts.

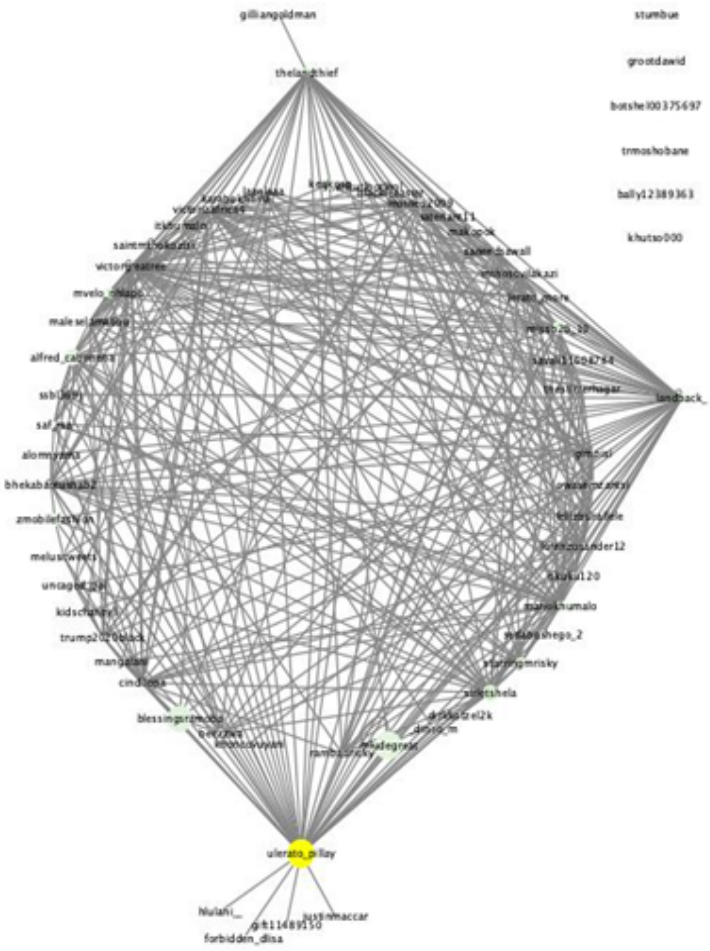


Image 3: Network of Suspicious Accounts, highlighted accounts moved outwards to emphasise connections.

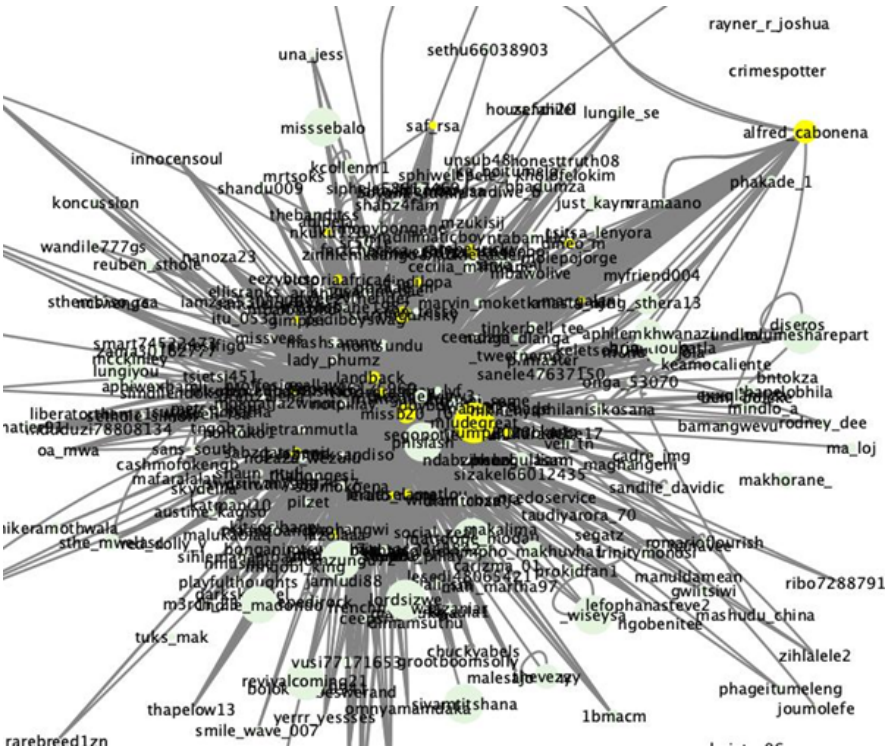


Image 4: Retweet Sources, suspicious accounts highlighted in yellow.

To demonstrate the importance of uLerato_Pillay in the overall network, two maps were generated. Image 4 shows the connections between accounts that retweeted content from uLerato, and how they are connected to one another. Suspicious accounts are highlighted in yellow and are clustered within the middle of this map, indicating that they are coordinating efforts to some degree.

Similarly, Image 5 shows a map of all individuals in the network who retweeted another account more than ten times in the two months (1 April 2020 to 31 May 2020). uLerato_Pillay is the point that connects the circle of accounts to the fan at the bottom. The circular map offers further evidence of a coordinated effort among accounts, as they are frequently interacting with one another.

However, the fan at the bottom indicate that individuals are engaging with uLerato_Pillay's content for the purpose of sharing it to their own networks. This demonstrates that uLerato_Pillay is the source of an enormous volume of content that is being shared within the xenophobia conversation. It also shows that this conversation is not contained within a single network of accounts promoting a similar narrative, but has extended into other networks. This shs why uLerato_Pillay's account is central to the #PutSouthAfricaFirst network map shown in Image 1, as this is the origin point for much of the content that is diffusing outwards throughout South Africa Twitter.

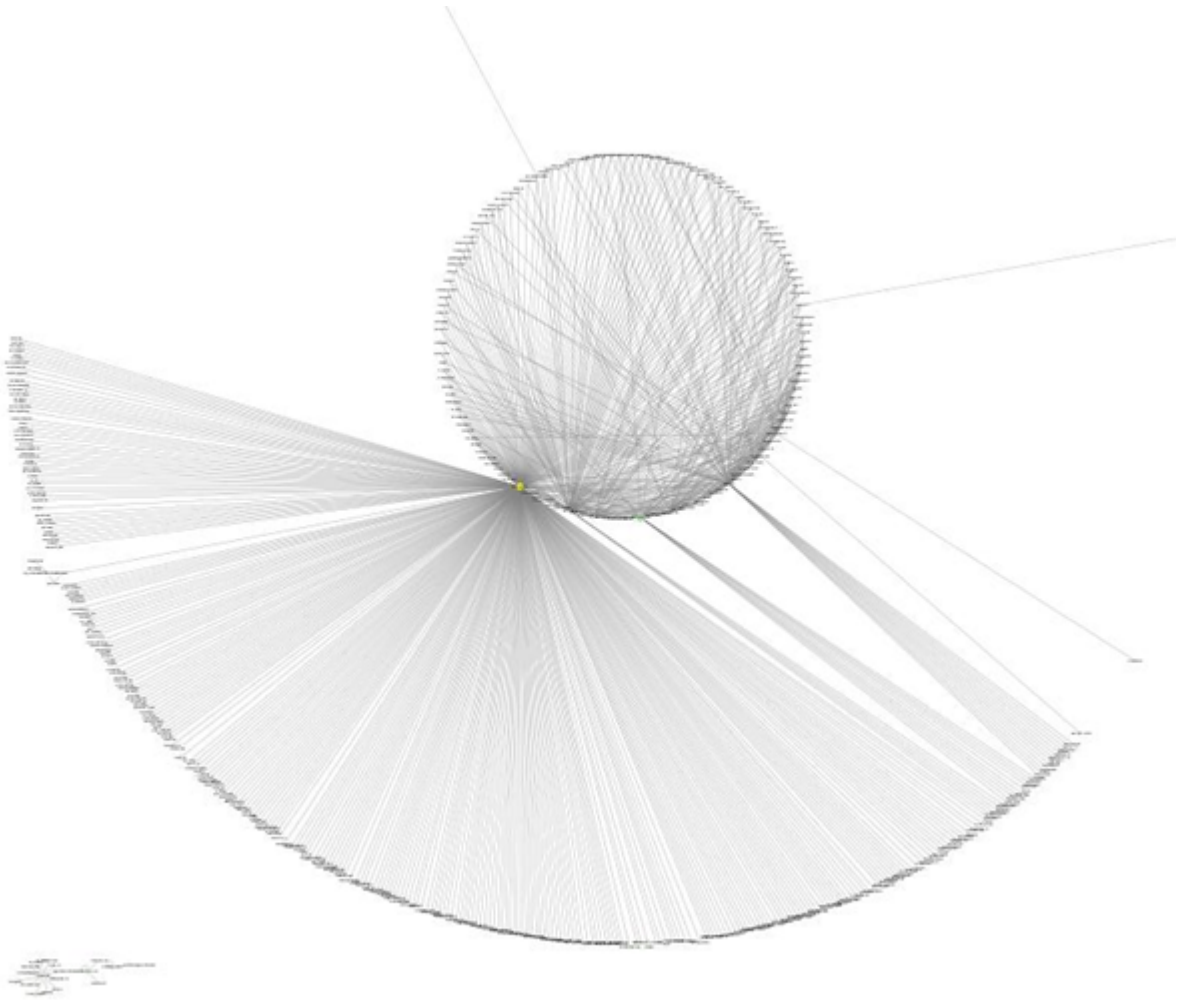


Image 5: Accounts that retweeted another 10 or more times.

Conclusion and Summary:

- The hashtag, #PutSouthAfricaFirst emerged on 27 April 2020 and was used more than 16 000 times in one day. The 80 accounts in the network around uLerato_Pillay were responsible for 50% of the use of the hashtag. This does not indicate organic growth in the conversation, but rather an organised network.
- uLerato_Pillay has tweeted about 5 000 times during the reporting period. It is likely that there are several people managing this account, as it is improbable that one person can manage such a large volume of Twitter activity.
- Several of the 80 accounts were dormant for long periods and were suddenly activated on 27 April 2020. Many of these accounts serve no other purpose but to “signal boost” negative sentiment about foreign Africans. There is no meaningful personal content from any of the accounts.
- There were approximately 12 000 people engaged in organic conversation on 1 April. This grew to about 50 000 in two months and continues to grow daily. This

indicates that about 80 accounts managed to artificially grow a significant organic conversation by targeting foreigners living in South Africa.

- The number of tweets and the signal-boosting through the inner circle of key accounts suggest that it is highly likely that the network has significant financial support.
- The network's playbook is anti-EFF and anti-ANC, who have publicly opposed xenophobia.
- The network is sophisticated and highly organised. They know how to jump hashtags to allow people in the know to continue to follow them, while frustrating attempts to track their activity.
- The network regularly tweets at Herman Mashaba and Julius Malema, who are not an integral part of the network. It appears that they want to capitalise on the extensive online following of these two prominent politicians.

The qualitative and quantitative data indicates that this network collectively propagates the xenophobic narrative in South Africa, and incorporates targeted political content which is predominantly anti-ANC and anti-EFF.

Although the data reveals a significant amount of the xenophobia conversation is organised, coordinated and probably funded, there are still a significant number of individuals within the xenophobia conversation who did not link to this suspicious network.

End

Note to editors:
This is a developing story and the information emerging from current analysis changes rapidly. This report is based on the findings of the first reporting period. The research is ongoing, and further updates will be issued. The information presented here should therefore not be considered as a final report.

Compiled by the Dialogue Facilitation team at the CABC.

About CABC
The Centre for Analytics and Behavioural Change (CABC) is a non-profit organisation based at UCTs Graduate School of Business and incubated by the Allan Gray Centre for Values-Based Leadership. It was established to track and counter mis- and disinformation, fake news and divisive and polarising rhetoric that is promulgated online to undermine social cohesion, democratic integrity, and the stability of nation states.

Cabc.org.za
info@cabc.org.za