Graph analysis of Twitter feed network maps The detection of network patterns within the South African Twitter community

Stefanie Strachan[1] and Aurona Gerber[1,2[0000-0003-1743-8167]]

[1] Department of Informatics, University of Pretoria, Pretoria, South Africa
[2] Centre for AI Research (CAIR), South Africa

Keywords: Social Network Analysis, Graph Analysis, Twitter Networks, Community Clusters, Network Visualisation

Extended Abstract. In the past decade various studies were done on the usage and influence of the internet, technology, and social media in society [1, 2, 3]. These studies indicate that social media has become a popular means of communication and therefore forms a significant part of what determines the views and opinions of society [4]. It is therefore increasingly important to analyse and understand social media networks and information flows. Understanding the ways that online communities form and communicate, as well as the different underlying network structures, could possibly assist with identifying key influencers as well as the formation and dynamics of social networks and communities [5]. The information can be utilised, for example, in the design of necessary interventions, or the identification of targeted campaigns [6].

The aim of this study was to extend previous research [7] and investigate the potential network structures that are formed within the South African Twitter community around the 2019 elections by means of knowledge graphs with an existing toolset namely NodeXL. Previous research and findings suggest that social network maps of Twitter communities can be analysed and visualised to provide insight into the setting of social media. The network maps point out previously unknown information about the persons and topics that drive conversations and group behaviour [8]. The main research question answered by this study can be stated as follows: *Are there any distinct network patterns that can be detected within the South African Twitter community, specifically concerning recent political concerns?*

Twitter datasets on the SA elections during three specific periods were collected, namely, the last week that political campaigning was allowed (28 April 2019 – 4 May 2019), Election day on 8 May 2019, and Inauguration day on 24 May 2019. The datasets were imported into NodeXL, and unique entities, or vertices, were determined. The vertices were grouped into clusters using the Clauset-Newman-Moore algorithm that analyses how the vertices are connected to one another. A range of measures of the graph was calculated for the overall network, as well as for each vertex in the network. Each of the network metrics summarizes a different aspect of the size and shape of the overall network and the location and connection properties of each entity in the network graph. For all three datasets groups with less than ten vertex connections as well as neighbourless edges were removed from the graph in order to focus the network

analysis on the core network. Subgraph images were created for each vertex before the overall network was visualized using the NodeXL visualisation feature. Figure 1 below depicts the graph of Dataset 3 – Inauguration Day (23-24 May 2019).

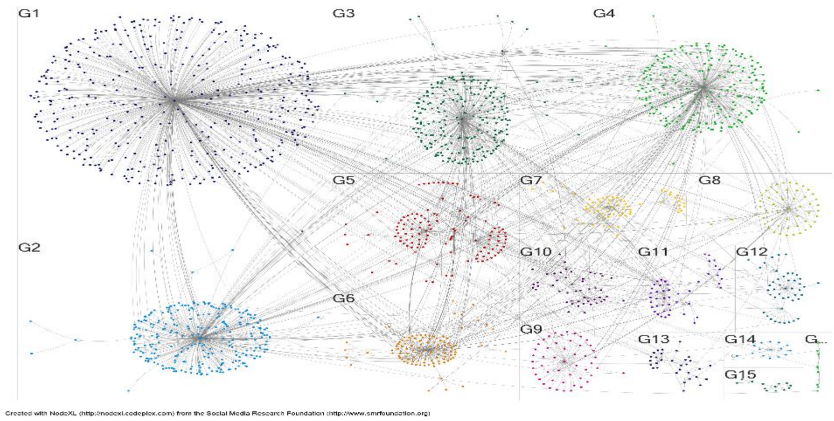

Figure 1: Network visualisation for Dataset 3

All three of the datasets that were analysed depicted community clusters that have a large number of disconnected entities that contribute to the network by mentioning the topic but does not link to other entities within the network. Both the main, larger groups as well as the smaller, sub-groups are highly interconnected and only share a small percentage of connections across groups. Community clusters often reveal diversity of opinion on a specific topic that exists within a society. Since South Africa has multiple political parties, multiple groups are formed, both on the general subject of the elections, as well as centred around specific political groups or influencers. In Datasets 1&2 groups shared a common topic but had a different focus, while the groups in Dataset 3 remain segregated.

In conclusion Twitter has become a popular means of communication and conversations create networks with identifiable groups and connections as users reply to and mention one another. Mapping social networks can assist in understanding the different ways that individuals form communities and organize online. Previous research identified six distinct patterns in the social structures depending on the topic in question [9]. The results of this study on South African Twitter feeds around the 2019 elections identified the community clusters pattern, which means a large number of disconnected entities contribute to the network by mentioning the topic but does not link to other entities within the network. Both the main, larger groups as well as the smaller or sub-groups are highly interconnected and only shared a small percentage of connections across groups. The most central or influential users for each network could also be determined.

References

1. R. Hanna, A. Rohm and V. Crittenden, “We’re all connected: The power of the social media ecosystem,” Business Horizons, vol. 54, pp. 265-273, 2011.
2. A. Perrin, “Social Media Usage: 2005-2015,” Pew Research Center, 2015.
3. A. M. Kaplan and M. Haenlein, “Users of the world, unite! The challenges and opportunities of Social Media,” Business Horizons, vol. 53, pp. 59-68, 2010.
4. B. Schivinski and D. Dabrowski, “The effect of social media communication on consumer perceptions of brands,” Journal of Marketing Communications, vol. 22, no. 2, 2014.
5. J. Kietzmann, K. Hermkens, I. McCarthy and B. Silvestre, “Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media,” Business Horizons, vol. 54, pp. 241-251, 2011.
6. L. Bennet, “The Personalization of Politics Political Identity, Social Media, and Changing Patterns of Participation.,” The ANNALS of the American Academy of Political and Social Science, vol. 644, no. 1, pp. 20-39, 2012.
7. Social Media Research Foundation, “Social Media Research Foundation,” 2016. [Online]. Available: https://www.smrfoundation.org. [Accessed April 2019].
8. M. Smith, L. Rainie, B. Shneiderman and I. Himelboim, “Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters,” Pew Research Internet Project, 2014.
9. M. Smith, C. A., N. Milic-Frayling, B. Shneiderman, E. Mendes Rodrigues, J. Leskovec and C. Dunne, “NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016,” Social Media Research Foundation, 2010.
10. S. Wasserman and K. Faust, “Social Network Analysis in the Social and Behavioral Sciences,” in Social Network Analysis: Methods and Applications, Cambridge University Press, 1994, pp. 1-27.
11. E. M. Rodrigues, N. Milic-Frayling, M. Smith, B. Shneiderman and D. Hansen, “Group-in-a-box Layout for Multi-faceted Analysis of Communities,” IEEE Third International Conference on Social Computing, 2011.
12. D. Ma, “Visualisation of Social Media Data: Mapping Changing Social Networks,” University of Twente, Enchede, 2012.
13. D. Hansen, D. Rotman, E. Bonsignore, N. Milic-Frayling, E. Rodrigues, M. Smith and B. Shneiderman, “Do You Know the Way to SNA?: A Process Model for Analyzing and Visualizing Social Media Data,” University of Maryland Tech Report: HCIL-2009-17, 2009.
14. Twitter, “Twitter,” 2019. [Online]. Available: twitter.com. [Accessed 2019].
15. Pew Research Center, “Pew Research Center,” 2019. [Online]. Available: https://www.pewresearch.org/. [Accessed May 2019].
16. D. Hansen, B. Shneiderman and M. Smith, “Analyzing Social Media Networks: Learning by Doing with NodeXL,” in Analyzing Social Media Networks with NodeXL: Insights from a connected world, San Francisco, CA, USA, Morgan Kaufmann Publishers Inc., 2010.