#Liberty breach: An exploratory usage case of NodeXL Pro as a social media analytics tool for Twitter

Title: Prof.
Name: Ilse Struweg
Affiliation: University of Johannesburg
Postal address: PO Box 524, Auckland Park, 2006
Telephone number: 011 559 4021
Email: istruweg@uj.ac.za

Abstract
Social media analytics uses data mining tools, platforms, and analytics techniques to collect and analyse infinite amounts of social media data. Social media analytics tools extract patterns and connections from data, for insight into market sentiments and requirements, to enhance business intelligence. 'Network Overview, Discovery and Exploration for Excel Pro' (NodeXL Pro) is a social media analytics tool that simplifies basic network analysis tasks and supports the analysis of social media networks. NodeXL Pro does sophisticated 'crawling' (extracting data) across a range of social media platforms. Through a qualitative case study design, this study explores and describes the use of NodeXL Pro through empirical and multimodal analysis and social network visualisation of social media data of the Liberty Holdings Ltd data breach crisis case in June 2018. The hashtag '#Liberty breach' resulted in 10 000 data sources ('tweets') from the social media platform Twitter. This study is unique on two levels. Firstly, it appears to be the first study in the South African marketing literature to use NodeXL Pro in social media analytics. Secondly, it presents the case study as a usage case to describe, in a step-by-step way, the functionalities of NodeXL Pro through social network analysis. The main finding of the paper focuses on the usability and manifold features (including the integrated visualisation tool) of NodeXL Pro. This social media analytics tool can open doors for marketing scholars and practitioners alike to measure, map, and model collections of connections.

Keywords: Social media analytics; NodeXL Pro; social network analysis; Twitter

1. Introduction

Social media analytics has emerged as an advanced research field after years of rapid and augmented adoption of social networks by consumers and organisations alike. As early as 2010, scholars indicated that there are opportunities for theoretical and practical inquiry to create new knowledge and scientific possibilities by leveraging data, technology, analytics, business, and society (Culnan, McHugh & Zubillaga, 2010). This argument for theoretical and practical inquiry into social media analytics was based on the richness and the dynamic nature of social data.

More recently there has been growing evidence that social media analytics provides a broader view of consumers, groups, and society, and creates business value by identifying new patterns and opportunities (Batrinca & Treleaven, 2015; Moe, Netzer & Schweidel, 2017). These developments in the realm of social media analytics provide opportunities for innovative, non-traditional research. Chen, Lu, Chau and Gupta (2014) support this notion, arguing that the pervasive impact of social media as a source of information has triggered renewed interest in social media analytics research. Wamba, Akter, Kang, Bhattacharya and Upal (2016) agree that the power of social media remains on the increase, but that, similarly, its measurement continues to be a challenge.

To achieve this goal, advances in social media analytics tools could be more effectively applied if scholars constantly explored these tools in different contexts, methodologies, and disciplines. One such social media analytics tool, NodeXL, has recently received considerable scholarly attention (Batrinca & Treleaen, 2015; Salge & Karahanna, 2018; Feng, 2016; Platt & Soens, 2018; Bokunewicz & Shulman, 2017). More specifically, the use of NodeXL in social media analytics has related to Twitter usage during various crisis situations. Examples of these studies include Ferra and Nguyen (2017), who studied #migrantcrisis during the European migration crisis; Ahmed's (2018) study, which used Twitter data to provide qualitative insights into pandemics and epidemics; Brummette and Fussell Sisco (2018), who used Twitter data to frame the Chipotle restaurant chain crisis; and a study by Keib, Himelboim and Han (2018) of the #BlackLivesMatter controversy in the United Kingdom. NodeXL Pro is a licence-based, technologically driven social media analytics tool that uses advanced 'crawling' capabilities over a number of social media platforms. It also supports the capturing, analysis, and social network visualisation of available public information. Against this backdrop, the following research objective guides this paper: *To explore and describe the usage case of NodeXL Pro (a social media analytics tool) to conduct an empirical and multimodal social media network analysis of '#Liberty breach' on Twitter.*

The significance of this paper is, firstly, that, according to available data, this is the first study in the South African marketing literature to explore the use of NodeXL Pro as a social media analytics tool; and secondly, that the methodological design of the paper is unique, in that it is presented as a usage case of the functionalities of NodeXL Pro.

The paper starts with a literature review, consisting of a brief discussion of social media in general, and of Twitter as one such social media platform. Subsequently, 'social media analytics' is defined, which leads to a discussion of social network analysis. In the methods section, the methodological choices of the study are presented, including a description of the Liberty Holdings Ltd data breach crisis case. This is followed by a discussion of the use and features of NodeXL Pro as an innovative social media analytics tool. The results and discussion are followed by the conclusion and recommendations.

2. Review of literature

The meanings of the concept 'social media' are manifold. Defining it is widely argued, and views abound about which tools, platforms, and social phenomena can be regarded as 'social'. Yet its integration into daily life at grassroots level is unquestionable (McCay-Peet & Quan-Haase, 2017). The broad social media definition presented by McCay-Peet and Quan-Haase (2017), which has the potential to include numerous technologies with fundamental social elements, is accepted for the purpose of this paper: "Social media are web-based services that allow individuals, communities, and organisations to collaborate, connect, interact, and build community by enabling them to create, co-create, modify, share, and engage with user-generated content that is easily accessible" (McCay-Peet & Quan-Haase, 2017:17).

Gruzd, Staves and Wilk (2012) present a variety of social media technologies, including Skype, Flickr, Twitter, Facebook, and Academia.edu. Pinterest, YouTube, Yelp, Weibo, Snapshot, and LinkedIn could also be added to the list of Gruzd *et al.* (2012). McCay-Peet and Quan-Haase (2017) argue that scholars have tended to favour Facebook and Twitter in social media research, and state that Twitter, specifically, has transformed the diffusion of information and news around the world. Consequently, Twitter is discussed in the next section.

2.1 Twitter as a social media technology platform

As more people use Twitter to communicate with one another, there is an urgent need to look into methodologies to study this interaction, in order to understand better the patterns, influences, and meanings of communication in that setting. Sanawi, Samani and Taibi (2017) argue that Twitter is one of the most dominant and persuasive social media platforms. Evidence from Statista (2018) supports this notion: it states that Twitter remains one of the most popular social networks worldwide (averaging 336 million monthly active users at the end of the first quarter of 2018), as a result of (a) the ability of users to follow any other user with a public profile, and (b) enabling users to interact with entities who regularly post on the social media site. These could include, for example, Twitter users, news agencies, governments, and organisations, depending on the context.

Social media technology platforms focus on idiosyncratic groupings of content creators and content consumers. Lee, O'Donnell and Hust (2018) regards Twitter as a real-time information network that connects users and followers to the latest stories, ideas, opinions, and news. Twitter is a social networking and micro-blogging platform (McCay-Peet & Quan-Haase, 2017) that empowers registered users to read the views of others and to express their own views in the form of 'tweets' (Pujari, Pujari, Bhat & Dixit, 2018). Micro-blogging, according to Liu, Min, Zhai and Smyth (2016), focuses on information diffusion and interactivity among open-platform users, whereas social network sites (such as Facebook) enable users to advance social relationships with their social networks. Another key feature of Twitter is that it differs from traditional blogging in that its content is typically briefer. The social network's original text limit of 140 characters per message was set at the company's launch in 2006. This was changed late in 2017 to a limit of 280 characters (Larson, 2017). Beyond text characters, Twitter engagement also allows users to upload photos or short videos. Three types of tweets are found on Twitter: original tweets, replies, and retweets (Chae, 2015).

In recent years, these tweets (Twitter data) have become one of the most popular information sources for academic research and practical applications alike. However, the key to using tweets as an information source is that intelligence and knowledge should be extracted from them. Therefore, the next section discusses social media analytics as a whole, and then focuses on two types of social media analytics that are widely used in Twitter analytics.

2.2 Defining 'social media analytics'

The term 'social media analytics' is defined for the purpose of this study predominantly on the basis of the descriptions of He and Xu (2016) and Zeng, Chen, Lusch and Li (2010). It is the development and evaluation of informatics tools and frameworks, in order to gather, scrutinise, condense, and visualise social media data to enable dialogue and connections, and so to derive useful patterns and intelligence.

Chen, Chiang and Storey (2012) adopt a slightly different perspective: they define social media analytics more from a customer perspective, as an approach to revealing what customers think and feel through the analysis of both structured and unstructured online data from online sources. Stieglitz, Dang-Xuan, Bruns and Neuberger (2014) focus on the purpose of social media analytics, concluding that it aims to combine, extend, and adapt methods for the analysis of social media data.

Some confusion appears in the literature in relation to social media analytics, especially when referring to big data analytics. Therefore, for the purpose of this paper, the interpretation of Wamba *et al.* (2016) of the concepts of big data analytics and social media analytics is supported. Their view is that *social media analytics* share similarities with *big data analysis*. Their argument is based on the work of Kiron, Perguson and Prentice (2013), which refers to big data analysis as "accumulated traces of consumers' online activities". Thus both social media analytics and big data analytics involve the analysis, management, and visualisation of similar types of datasets, and both seek to understand the fundamental relational and interactive components of consumers' social media activities.

Akter, Bhattacharya, Fosso Wamba and Aditya (2016) and Wamba *et al.* (2016) describe eight types of social media analytics: topic modelling, opinion mining, sentiment analysis, social network analysis, trend analysis, popularity prediction, customer engagement analysis, and visual analytics. As indicated in the introduction to this article, it will focus on only one social media analytics type related to the functionalities of NodeXL Pro and Twitter analytics – namely, social network analysis.

Social network analysis is used to model relations and associations, developments, and dynamic forces in networks and activities on social media platforms. Social network theory and analysis builds on and uses concepts from the mathematics of graph theory (Hansen, Schneiderman & Smith, 2011; Lee, 2018). It is the process of studying the structures of social networks, and seeks to explain how networks organise and analyse the complex set of relationships inside a social network of individuals or organisations (Scott, 2012; Wasserman & Faust, 1994).

In these social networks, 'nodes' are the individual actors and 'ties' are the relationships between the actors (Lee, 2018). Social network analysis produces a mathematical and visual analysis of the actor relationships within a network. This is done by modelling the social network dynamics and developments – for example, network density, network centrality, and network flows. Nodes and ties – both strong and weak – are important in diffusing information (Brown & Reingen, 1987; Datta, Chowdhury & Chakraborty, 2005).

Social network analysis further uses a range of techniques to give insight into the structures of the network (Scott, 2012). Fan and Gordon (2014) maintain that these techniques range from uncomplicated methods (for example, counting the number of edges a node has, or calculating path lengths) to more sophisticated methods (for example, computing eigenvectors to determine key nodes in a network). Therefore, in social network analysis, complex sets of relationships of connected symbols are visualised as maps (such as graphs or sociograms), but precise measures of the size, shape, and density of the network as a whole and the positions of each node within it are also calculated. It is argued, therefore, that social network analysis enables researchers to see phenomena as a collection of interconnected pieces, using relationships to create emergent patterns of linkages *between* individuals, not *within* individuals.

3. Methodology

This qualitative study uses an exploratory, descriptive, single case study design. The choice of a qualitative case study design is based on Bassey's (1999) argument that the case study approach provides not only a mechanism for theory-seeking and theory testing, but also for storytelling – in this case, the account of the NodeXL Pro usage case, through the empirical and multimodal analysis of 10 000 #Liberty breach tweets. However, Meyer (2001) maintains that the case study is a rather loose design, and therefore cautions that a number of choices need to be addressed in a principled way. For the purpose of this study, these choices are outlined in Table 1.

Table 1: Methodological choices for the purpose of this study

Methodological consideration	Methodological choice
Research paradigm	Qualitative research
Research design	Exploratory, descriptive, single case study
Sampling strategy	Case selection
The case	Liberty Holdings Ltd IT security breach crisis
Data sources	10 000 tweets: #Liberty breach
Data collection	Social media mining approach
	NodeXL Pro API
Data analysis	NodeXL Pro social network analysis
	NodeXL Pro advanced network metrics

The next section provides a brief explanation of the Liberty Holdings Ltd IT security breach case, followed by an overview of NodeXL Pro as a social media analytics tool.

3.1 The case: Liberty Holdings Ltd IT security breach crisis

On Thursday, 14 June 2018, South African financial services provider, Liberty Holdings Ltd, fell victim to an IT systems breach. Specifically the breach occurred in one of Liberty Holdings Ltd's subsidiaries, Liberty Life. In a Fin24 article on 19 June 2018, it was reported that the CEO of Liberty, David Munro, confirmed that the data breach was limited to Liberty's insurance clients, and that none of its other businesses was compromised (Niselow, 2018). The company informed their clients about the incident on Saturday evening, 16 June 2018, via email and text message, that it had suffered "unauthorised access to its IT infrastructure".

According to a report in the *Sunday Times* on 17 June 2018, the hackers demanded millions of Rands from Liberty, threatening to make the data public unless they were paid (Afrika, 2018). In a statement on IT Web on 18 June 2018, a cyber security expert stated that "the news, of course, has sent ripples through the insurance, finance and cyber security industry" (Ukuvuma Security, 2018). Amid these 'ripples', Liberty CEO David Munro affirmed at a media briefing on Sunday 17 June 2018 that "investments remained uncompromised and that an investigation was underway" (Seeth, 2018). Despite the assurances given by Liberty, calls were made for transparency about exactly how Liberty was breached, why no-one detected the breach until the hackers themselves informed Liberty, how the hackers acquired access, why Liberty was slow in communicating to its stakeholders during the crisis, and the like, across a broad range of media platforms (Ukuvuma Security, 2018; McLoughlin, 2018). As a result of the data breach, Liberty Holdings' share price fell almost five per cent – from R124 to R119,16 — on Monday, 18 June 2018 (Seeth, 2018).

3.2 NodeXL Pro as social media analytics tool

NodeXL (Network Overview for Discovery and Exploration in Excel) was developed by the Social Media Research Foundation (https://www.smrfoundation.org/). It consists of two main options: NodeXL Basic (open source), and NodeXL Pro (licence-based). Both are plug-ins for Microsoft Excel, and support social network and content analysis. Smith, Shneiderman, Milic-Frailing, Mendes Rodrigues, Barash, Dunne and Gleave (2009) state that NodeXL was intended to be easy for existing Excel users to adopt. This was done by using the common spreadsheet capabilities in Excel, and expanding the spreadsheet into a network analysis and visualisation tool by incorporating a library of basic network metrics (for example, degree, centrality measures, elementary clustering) and graph visualisation facets.

Node XL uses a highly systematised workbook template that includes multiple worksheets to keep all the information needed to represent a network graph. Network relationships (referred to as 'graph edges' in NodeXL) are represented as an 'edge list'. The 'edge list' contains all pairs of entities that are connected in the network. Corresponding worksheets comprise information about each vertex and cluster. The visualisation features in NodeXL display a range of network graph depictions and chart data attributes to visual properties, including shape, colour, size, transparency, and location (Hansen, Rotman, Bonsignore, Milic-Frayling, Rodriques, Smith & Shneiderman, 2012).

The Social Media Research Foundation (2018a) describes NodeXL Basic as being positioned as a free browser for files created with NodeXL Pro, which offers more advanced features. These advanced features build on the features in NodeXL Basic to include the following, among others (Social Media Research Foundation, 2018b):

- *Advanced network metrics*, which include, among other functions, determining betweenness centrality, closeness centrality, and Eigenvector centrality.
- *Content analysis*, which includes text analysis, sentiment analysis, time series analysis, and top items (words, word pairs, URLs, and hashtags).
- *Access to the social network application programming interfaces (APIs)* of Flickr, Facebook, YouTube, and Twitter, as well as third-party graph data importers.

The features in NodeXL Pro used for the purpose of this study included (a) network visualisation, (b) social network analysis, (c) social network APIs, (d) data import, (e) data export, and (f) task automation.

4. Results and discussion

To address Meyer's (2001) concern about the 'looseness' that is sometimes evident in case study designs, this section is presented according to the eight-step system description and workflow for NodeXL and NodeXL Pro, as suggested by Smith *et al.* (2009). This workflow moves from data import through processing, calculation, and refinement, before creating a network graph that "tells a useful story" (Smith *et al.,* 2009:4).

4.1 Step 1: Import data

As this paper focuses on exploring NodeXL Pro as a social media analytics tool to examine the information of the Liberty Holdings Ltd IT security breach crisis, only the most popular hashtag identified by Twitter was used – namely, '#Liberty breach'. The NodeXL Pro Twitter data import feature was used to extract networks for '#Liberty breach'. The collection of the '#Liberty breach' data was performed on Tuesday, 19 June 2018. NodeXL Pro allows for data from the past seven days and/or 18 000 tweets. For the purpose of this study, 10 000 tweets were extracted.

The data was then automatically entered into the NodeXL Pro template in the 'edges' and 'vertices' worksheets. These two concepts are central to network analytics theory (Chae, 2012). Firstly, Hansen *et al.* (2009) argue that vertices (also referred to as 'nodes', 'agents', 'entities', or 'items') can include (a) individuals, (b) social structures (such as workgroups, teams, organisations, institutions, states, or countries), (c) content (such as web pages, keyword tags, or videos), or (d) even locations (physical and virtual) and (e) events. Attribute data of vertices are also available in NodeXL Pro, which could include demographic data, data that describe the vertices' use of a system (for example, number of logins, messages posted, and edits made) or location. However, for the purpose of this paper, attribute data were not considered.

Secondly, 'edges' (also known as 'links', 'ties', 'connections', or 'relationships'), on the other hand, can be said to occur if they have a particular official status, if they are recognised by the participants, or if they are observed by exchange or collaboration between them (Hansen *et al.,* 2009). Alhajj and Rokne (2014) regard these edges as social interactions, organisational structures, physical proximities or abstract interactions such as hyperlinks. In essence, therefore, an edge connects two vertices together.

According to Hansen *et al.* (2009), edges can be divided into two type of connection: directed or undirected. Directed edges (or asymmetric edges) have a distinct source and end-point. These edges are represented on a graph as a line with an arrow pointing from the source vertex to the recipient vertex (the end-point). However, undirected edges (or symmetric edges) only occur between two people or things, with no clear source or end-point in these mutual relationships. These edges are represented on a graph as a line connecting two vertices, with no arrows.

In the '#Liberty breach' case, a total of 1 015 edges (including 847 unique edges and 168 edges with duplicates) and 767 vertices were identified through NodeXL Pro. The edges in this study were all presented as directed edges.

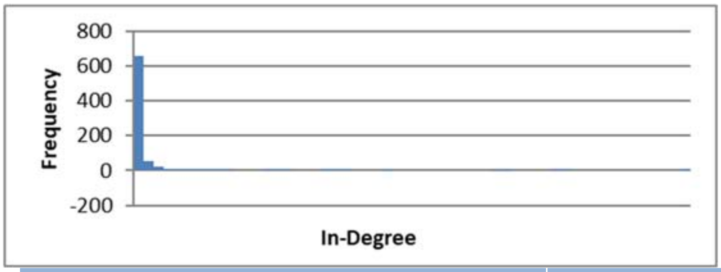
4.2 Step 2: Clean the data (if required)

This step entails the elimination of duplicate edges. Smith *et al.* (2009) argue that, in some instances, network measures cannot be accurately determined if multiple edges exist between the identical pair of entities in a single data set. Then the redundant edges may be aggregated into a single edge with a weighting, reflecting the number of original instances. For the purpose of this paper, data cleaning was done, resulting in 168 edges after duplicates had been merged.

4.3 Step 3: Calculating the graph metrics

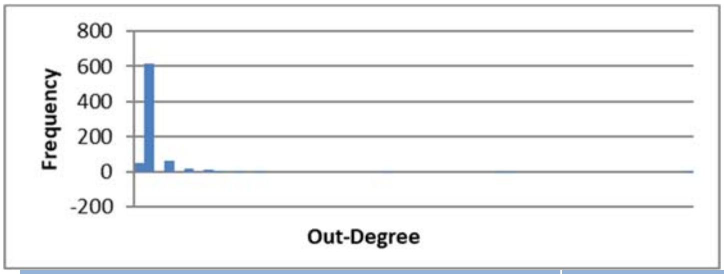
The third step in the eight-step system description and workflow for NodeXL and NodeXL Pro calculates the graph metrics. NodeXL Pro calculates several network graph metrics to capture the size and internal connectivity of a network, and also the attributes of each vertex, based on in- and out-degree, betweenness, closeness, and eigenvector centrality (Hansen *et al.,* 2009). The data from the '#Liberty breach' case in relation to each of the graph metrics are discussed below.

The degree of a vertex (sometimes referred to as 'degree centrality') is a count of the number of unique edges that are connected to it (Hansen *et al.,* 2009). Kim and Hastak (2017) explain it similarly, in that 'degree centrality' refers to the number of edges a vertex has that connect to other vertices. In the '#Liberty case', which analysed a directed graph, this single degree metric is split into two metrics, in- and out-degree centrality. Figures 1 and 2 illustrate the in- and out-degree centrality metrics for the '#Liberty breach' case.



Minimum in-degree	0
Maximum in-degree	76
Average in-degree	1,181

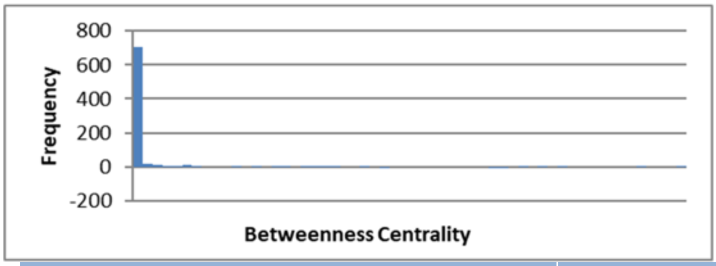
Graph 1: '#Liberty breach': In-degree centrality



Minimum out-degree	0
Maximum out-degree	31
Average out-degree	1,181
Median out-degree	1,000

Figure 2: '#Liberty breach': In-degree centrality

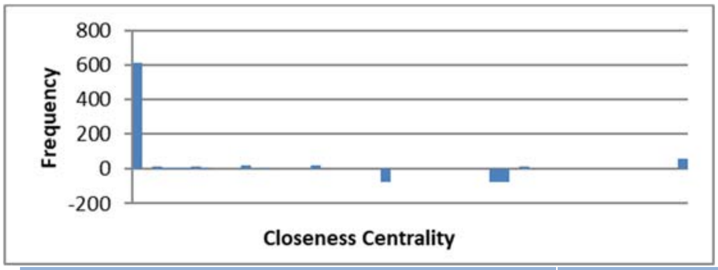
From Graph 1, therefore, it can be deduced that the average number of edges that point toward the vertex of interest is 1 181, and from Graph 2 that the out-degree centrality for '#Liberty breach' has an average of 1 181 and a median of 1 000 edges towards which the vertex of interest points.



Minimum betweenness centrality	0,000
Maximum betweenness centrality	90552,967
Average betweenness centrality	1174,866
Median betweenness centrality	0,000

Graph 3: '#Liberty breach': Betweenness centrality

From the social network perspective, Wasserman and Faust (1994) described the significance of high betweenness: "interactions between two nonadjacent actors might depend on other actors in the set of actors, especially the actors who lie on the paths between the two". Wasserman and Faust (1994) refer to 'actors', which should be read as 'vertices' in the NodeXL Pro context. From Graph 3 it can be deduced that, in the case of '#Liberty breach', the betweenness centrality was very high, which means that these high betweenness vertices played the role of what Kim and Hastak (2018) refer to as 'gatekeepers' in handling the information flow between Liberty Holdings Ltd and other communities.



Minimum closeness centrality	0,000
Maximum closeness centrality	1,000
Average closeness centrality	0,110
Median closeness centrality	0,000

Graph 4: '#Liberty breach': Closeness centrality

For Hansen *et al.* (2009), closeness centrality is viewed differently from the other network metrics. Closeness centrality captures the average distance between a vertex and every other vertex in the network. Assuming that vertices can only pass messages to or influence their existing connections (vertices), a low closeness centrality means that a person (vertex) is directly connected to, or "just a hop away" (Hansen *et al.*, 2009) from, most other vertices in the network. By contrast, vertices in very peripheral locations may have high closeness centrality scores, indicating the number of vertex 'hops' they need to take to connect to distant others in the network. In the social media '#Liberty breach', it is significant that the average closeness is very localised, as seen in Graph 4 above. This might be because the vast majority of tweets in this case were sent by media houses, which use content across news channels and other media.

Much like degree centrality (depicted in Graphs 1 and 2), eigenvector centrality favours vertices that have high correlations with many other vertices. However, in contrast to degree centrality, eigenvector centrality specifically favours vertices that are connected to vertices that are themselves central within the network (Lohmann, Margulies, Horstmann, Pleger, Lepsien, Goldhahn, Shloegl, Stumvoll, Villringer & Turner, 2010). Thus it takes the entire pattern of the network into account. Hansen *et al.* (2009) agree, and state that the eigenvector centrality network metric takes into account not only how many connections a vertex has (i.e., its degree), but also the degree of the vertices to which it is connected. The eigenvector centrality for the '#Liberty breach' was especially low. This indicates that there were very limited influences over other vertices in the network. Also, the vertices in the '#Liberty breach' analysis were not well connected, and had limited links to other vertices through the network.

Steps 4 to 8 in the eight-step system description and workflow for NodeXL and NodeXL Pro application to the '#Liberty Breach' case are combined in the discussion under Step 8.

4.4 Step 4: Creating clusters

In network analysis, vertices can share attributes. Therefore, NodeXL Pro allows for the creation of clusters and cluster vertices worksheets in the fourth step. This clustering algorithm of NodeXL Pro is useful to group and analyse these vertices together. Hansen *et al.* (2009) encourage the use of these algorithms, as each cluster can have its own display features with a distinctive shape, colour, size, transparency, or image.

4.5 Step 5: Creating sub-graph images

Whole graph images can, in some cases, be too dense to reveal details about individual vertices or clusters. In this step of the eight-step system description and workflow for NodeXL, therefore, sub-graph images are created. Sub-graph images are useful depictions of the range of variation in the group vertices of the population in the network. This, according to Hansen *et al.* (2009), produces confined networks that focus on each vertex at a time, containing only the vertices with which that vertex is directly associated.

4.6 Step 6: Preparing edge lists

Vertices and edges hold attributes that can be used to order the data (for example, ordering vertices by the number of connections to other vertices). In Step 6, 'lay-out order', which directs the presentation of vertices in the graph display, is carried out in the whole graph visualisation.

4.7 Step 7: Expanding the worksheet with graphing features

In this step, NodeXL Pro allows for columns to be auto-filled to plot data in order to present different features. Hansen *et al.* (2009) provide the following suggestions:

- Graphical features of vertices and edges (their shape, colour, opacity, size, label, and tooltip) can be improved to suggest supplementary information in the network visualisation.
- Images in the 'images' worksheet can replace the shapes used to represent the vertices.
- Additional numerical attributes of each vertex could be added in adjacent columns and then, by design, scaled to present characteristics.

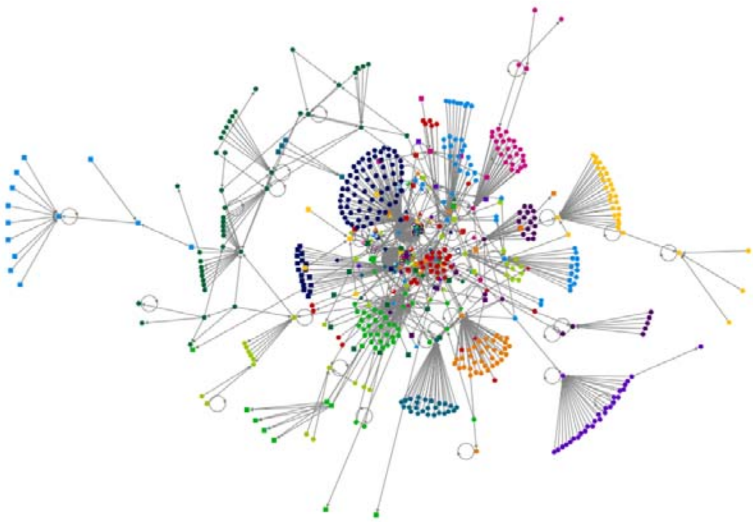
During this step, the layout features simplify the creation of multiple networks for a consistent presentation.

4.8 Step 8: Showing the graph

In the final step of the eight-step system description and workflow for NodeXL and NodeXL Pro, the social media analytics tool presents a visualisation of the network. Graph 5 below presents the 'whole graph', illustrating the network according to the Fruchterm-Reingold layout (the default layout of NodeXL Pro).

Graph 5: The social networks of '#Liberty Breach'

Graph 6 below also presents the 'whole graph', illustrating the network according to the Harel-Koren multiscale layout algorithm. In this layout, the clusters are more visible, clearly indicating the number of influencers in the social network.



Graph 6: The social networks of '#Liberty Breach' by clusters

In concluding this section of the NodeXL Pro '#Liberty breach' usage case, it should be noted that the workflow is not rigidly prescribed. The NodeXL Pro analysis and visualisation of the network can constantly be refined.

5. Conclusion

The purpose of this paper was to explore and describe the usage case of NodeXL Pro in conducting a social media network analysis of '#Liberty breach' on Twitter. The data set included 10 000 tweets, based on the most popular hashtag, '#Liberty breach', in the Liberty Holdings Ltd IT security breach in June 2018. From the results of the paper, it can be concluded that NodeXL Pro enables scholars and practitioners to interpret and create meaningful representations of complex social media networks, using big data, within a fairly short timeframe. The significance in the exploration and description of NodeXL Pro as a social media analytics tool in the usage case of '#Liberty breach' is based on the tool's close combination of spreadsheets, worksheets, and graph visualisations, as well as the variety of ways in which it allows for metrics and attributes to be mapped on to graphs. The result is a deeper understanding of the connected structure of the world.

6. Recommendations

Future studies are needed to delve deeper into all the features of NodeXL Pro. These range from the ability to use data from a variety of social media platforms in a single study, to the more sophisticated features of NodeXL Pro, to advance marketing intelligence for scholars and practitioners alike.

References:

Afrika, W.A. 2018. *Liberty life system falls victim to cyber-attack.* Available online: https://www.timeslive.co.za/sunday-times/news/2018-06-16-liberty-life--system--falls-victim-to--cyber-attack/ Accessed 18 June 2018.

Ahmed, W. 2018. *Using Twitter data to provide qualitative insights into pandemics and epidemics.* Doctoral dissertation, University of Sheffield. Available online: http://etheses.whiterose.ac.uk/20367/1/Final%20PhD%20Thesis%2011%20MAY.pdf Accessed 27 June 2018.

Akter, S., Bhattacharya, M., Fosso Wamba, S. & Aditya, S. 2016. How does social media analytics create value? *Journal of Organisational and End User Computing,* 28: 1-9. 10.4018/JOEUC.2016070101.

Alhajj, R. & Rokne, J. (eds). 2014. *Encyclopedia of social network analysis and mining*. New York: Springer Science and Business Media.

Bassey, M., 1999. *Case study research in educational settings*. London: McGraw-Hill.

Batrinca, B. & Treleaen, P.C. 2015. Social media analytics: A survey of techniques, tools and platforms. *AI & Society*, 30(1): 89-116.

Bokunewicz, J. F. & Shulman, J. 2017. Influencer identification in Twitter networks of destination marketing organisations. *Journal of Hospitality and Tourism Technology*, 8(2): 205-219.

Brown, J.J. & Reingen, P.H. 1987. Social ties and word-of-mouth referral behaviour. *Journal of Consumer research*, *14*(3): 350-362.

Brummette, J., & Fussell Sisco, H. 2018. Holy guacamole! Framing and the Chipotle contamination issue. *Journal of Communication Management*, 22(3): 280-295.

Chae, B. 2015. Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics,* 165: 247-259.

Chen, A., Lu, Y., Chau, P.Y.K. & Gupta, S. 2014. Classifying, measuring, and predicting users' overall active behaviour on social networking sites. *Journal of Management Information Systems*, 31(3): 213-253.

Chen, H., Chiang, R.H. & Storey, V.C. 2012. Business intelligence and analytics: From big data to big impact. *Management Information Systems Quarterly*, 36: 1165-1188.

Culnan, M.J., McHugh, P.J. & Zubillaga, J.I. 2010. How large U.S. companies can use Twitter and other social media to gain business value. *MIS Quarterly Executive*, 9(4): 243-259.

Datta, P.R., Chowdhury, D.N. & Chakraborty, B.R. 2005. Viral marketing: New form of word-of-mouth through Internet. *The business review*, *3*(2): 69-75.

Fan, W. & Gordon, M. 2014. The power of social media analytics. *Communications of the ACM,* 57(6): 74-81.

Feng, Y. 2016. Are you connected? Evaluating information cascades in online discussion about the #RaceTogether campaign. *Computers in Human Behavior*, 54: 43-53.

Ferra, I. & Nguyen, D. 2017. #Migrantcrisis: "Tagging" the European migration crisis on Twitter. *Journal of Communication Management*, 21(4): 411-426.

Gruzd, A., Staves, K. & Wilk, A. 2012. Connected scholars: Examining the role of social media in research practices of faculty using the UTAUT model. *Computers in Human Behavior*, 28(6): 2340-2350.

Hansen, D.L., Rotman, D., Bonsignore, E., Milic-Frayling, N., Rodrigues, E.M., Smith, M. & Shneiderman, B. 2012. Do you know the way to SNA? A process model for analyzing and visualizing social media network data. In *Social Informatics, 2012 International Conference on social informatics*, 304-313. IEEE.

Hansen, D., Schneiderman, B. & Smith, M.A. 2011. *Analyzing social media networks with NodeXL: Insights from a connected world*. Amsterdam / Boston: M. Kaufmann.

He, W. & Xu, G. 2016. Social media analytics: Unveiling the value, impact and implications of social media analytics for the management and use of online information. *Online Information Review,* 40(1): 2-5.

Keib, K., Himelboim, I. & Han, J.Y. 2018. Important tweets matter: Predicting retweets in the #BlackLivesMatter talk on Twitter. *Computers in Human Behavior,* 85: 106-115.

Khan, G.F. 2017. Social media analytics. In *Social Media for Government,* 93-118. Singapore: Springer.

Kim, J. & Hastak, M. 2018. Social network analysis. *International Journal of Information Management: The Journal for Information Professionals*, 38(1): 86-96.

Kiron, D., Perguson, R.B. & Prentice, P.K. 2013. *MIT Sloan Management Review,* 54(3): 1.

Larson, S. 2017. *Welcome to a world of 280-character tweets*. Available online: http://money.cnn.com/2017/11/07/technology/twitter-280-character-limit/index.html Accessed 25 June 2018.

Lee, Y.J., O'Donnell, N.H. & Hust, S.J., 2018. Interaction effects of system generated information and consumer scepticism: An evaluation of issue support behaviour in CSR twitter campaigns. *Journal of Interactive Advertising*, 19(2): 1-37.

Liu, Z., Min, Q., Zhai, Q. & Smyth, R. 2016. Self-disclosure in Chinese micro-blogging: A social exchange theory perspective. *Information & Management,* 53(1): 53-63.

Lohmann, G., Margulies, D.S., Horstmann, A., Pleger, B., Lepsien, J., Goldhahn, D., Shloegl, H., Stumvoll, M., Villringer, A. & Turner, R. 2010. Eigenvector centrality mapping for analyzing connectivity patterns in fMRI data of the human brain. *PloS one*, 5(4): e10232.

McCay-Peet, L. & Quan-Haase, A. 2017. What is social media and what questions can social media research help us answer? *The SAGE Handbook of Social Media Research Methods*. London: Sage.

McLoughlin, J. 2018. Liberty crack: What experts make of it all. *The Citizen* (Gauteng), 26 June, 18.

Meyer, C.B. 2001. A case in case study methodology. *Field methods*, *13*(4): 329-352.

Moe, W.W., Netzer, O. & Schweidel, D.A. 2017. Social media analytics. In *Handbook of Marketing Decision Models,* 483-504. Cham: Springer.

Niselow, T. 2018. Five massive data breaches affecting South Africa. Available online: https://www.fin24.com/Companies/ICT/five-massive-data-breaches-affecting-south-africans-20180619-2 Accessed 20 June 2018.

Platt, A. & Soens, S. 2018. Toward an automatic warning system for reputation damaging social media posts. *Journal of Computing Sciences in Colleges*, 33(5): 33-38.

Pujari, R., Pujari, J., Bhat, V.S. & Dixit, A. 2018. Timeline analysis of Twitter users. *Procedia Computer Science,* 132: 157-166.

Salge, C.A.D.L. & Karahanna, E. 2018. Protesting corruption on Twitter: Is it a bot or is it a person? *Academy of Management Discoveries*, 4(1): 32-49.

Sanawi, J.B., Samani, M.C. & Taibi, M. 2017. # Vaccination: Identifying influencers in the vaccination discussion on Twitter through social network visualisation. *International Journal of Business and Society*, 18(S4): 718-726.

Scott, J. 2012. *Social network analysis.* Thousand Oaks, CA: Sage.

Seeth, A. 2018. *Liberty's share price falls following data breach.* Available online: https://city-press.news24.com/Business/liberty-share-price-falls-following-data-breach-20180618 Accessed 20 June 2018.

Smith, M.A., Shneiderman, B., Milic-Frayling, N., Mendes Rodrigues, E., Barash, V., Dunne, C. & Gleave, E. 2009. Analyzing (social media) networks with NodeXL. In *Proceedings of the Fourth International Conference on Communities and Technologies,* 255-264. ACM.

Social Media Research Foundation. 2018a. *What is NodeXL?* Available online: https://www.smrfoundation.org/nodexl/ Accessed 27 June 2018.

Social Media Research Foundation. 2018b. *NodeXL feature overview.* Available online: https://www.smrfoundation.org/nodexl/features/ Accessed 27 June 2018.

Statista. 2018. *Number of monthly active users worldwide from first quarter 2010 to first quarter 2018*. Available online: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ Accessed 25 June 2018.

Stieglitz, S., Dang-Xuan, L., Bruns, A. & Neuberger, C., 2014. Social media analytics. *Wirtschaftsinformatik*, 56(2): 101-109.

Ukuvuma Security. 2018. *Latest big corporate data breach: Why Liberty's clients should be bringing it to book.* Available online: https://www.itweb.co.za/content/dgp45qaGaYO7X9l8 Accessed 19 June 2018.

Wamba, S.F., Akter, S., Kang, H., Bhattacharya, M. & Upal, M. 2016. The primer of social media analytics. *Journal of Organisational and End User Computing (JOEUC)*, 28(2): 1-12.

Wasserman, S. & Faust, K. 1994. *Social network analysis: Methods and applications.* Cambridge and New York: Cambridge University Press.

Zeng, D., Chen, H., Lusch, R. & Li, S.-H. 2010. Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6): 13-16.