Paper:
Grover, P., Kar, A., Dwivedi, Y. & Janssen, M. Polarization and acculturation in US Election 2016 outcomes – Can twitter analytics predict changes in voting preferences. *Technological Forecasting and Social Change*
http://dx.doi.org/10.1016/j.techfore.2018.09.009

Polarization and acculturation in the 2016 US presidential election: Can Twitter analytics predict changes in voting preferences?

Purva Grover
Information Systems area, DMS, Indian Institute of Technology Delhi
Email: groverdpurva@gmail.com

Arpan Kumar Kar
Information Systems area, DMS, Indian Institute of Technology Delhi, India.
Email: arpan_kar@yahoo.co.in

Yogesh K. Dwivedi
Professor Digital Marketing and Innovation and Director of Emerging Markets Research Centre, School of Management, Swansea University Bay Campus, Swansea, SA1 8EN, UK.
Email: y.k.dwivedi@swansea.ac.uk

Marjin Janssen
Professor of Technology, Policy and Management of Delft University of Technology, Netherlands.
Email: M.F.W.H.A.Janssen@tudelft.nl

Abstract

Elections are among the most critical events in a national calendar. During elections, candidates increasingly use social media platforms to engage voters. Using the 2016 US presidential election as a case study, we looked at the use of Twitter by political campaigns and examined how the drivers of voter behaviour were reflected in Twitter. Social media analytics have been used to derive insights related to theoretical frameworks within political science. Using social media analytics, we investigated whether the nature of social media discussions have an impact on voting behaviour during an election, through acculturation of ideologies and polarization of voter preferences. Our findings indicate that discussions on Twitter could have polarized users significantly. Reasons behind such polarization were explored using Newman and Sheth's model of voter's choice behaviour. Geographical analysis of tweets, users, and campaigns suggests acculturation of ideologies among voting groups. Finally, network analysis among voters indicates that polarization may have occurred due to differences between the respective online campaigns. This study thus provides important and highly relevant insights into voter behaviour for the future management and governance of successful political campaigns.

Keywords

Social Media; Twitter Analytics; Polarization in Elections; eParticipation; Public policy; Hashtag community; Acculturation in Social Media;

1. **Introduction**

Social media plays a pivotal role in impacting the outcome of national elections (Bruns and Stieglitz, 2013). The United States presidential election of 2016, held on 8 November, resulted in a victory for the Republican party; the Republican ticket of Donald Trump and Mike Pence defeating the Democratic ticket of Hillary Clinton and Tim Kaine. Using data from 784,153 tweets collected over the 120 days from 13 August to 10 December 2016 – and employing Twitter search terms such as 'Hillary Clinton', 'Donald Trump' and 'USA Election' – this paper offers insights into how Twitter was used by the 2016 presidential candidates and the way in which this reflects the political engagement of US citizens over the election period. The study also describes the Twitter campaigns run by the presidential candidates for the 58th quadrennial American presidential election, the drivers of their engagement and their potential impact.

The presidential election of the United States of America (USA) is a highly significant event for both the country and the rest of the world. Existing literature shows that increased use of digital media leads to increased political participation; raising the political knowledge of citizens and engaging them in the election campaigns (Dimitrova et al., 2014; Hossain et al., 2018; Ogola, 2015). Social media platforms support two-way communication (Kapoor and Dwivedi, 2015; Vaccari and Valeriani, 2015). According to the Pew Research Centre and the American Life Project, 69% of online adults use social networking sites (Social Media Fact Sheet, 2016). Online campaigning was one of the biggest drivers behind the Democrat victory of 2008 and Barack Obama presidential campaign (Stirland, 2008).

Social media allows people to – without meeting physically – create, share and exchange their thoughts, ideas, opinions, information, videos, images and other digital content in virtual communities such as Facebook, Twitter, LinkedIn, Google+, Slideshare, Flickr, Instagram and many more. These platforms allow users to form online communities in which they can share personal information and perspectives through user-generated content. Authors have described social media platform as a means for large-scale communication (Boynton and Richardson, 2016) and sharing purposes (AlAlwan et al. 2017; Barnett et al., 2017; Dwivedi et al. 2015; Kapoor et al., 2018; Hollander, 2008). Social media is able to empower voters by enhancing deliberative democracy among voters (Lawrence et al., 2010; Yardi and Boyd, 2010). Deliberation may help voters in: (a) refining their own opinions; (b) listening to different opinions; and (c) identifying common ends and means (Lawrence et al., 2010). However, research also indicates that online discussions may amplify division among social groups with differing views, rather than building consensus among them (Lee, 2007; Yardi and Boyd, 2010).

According to Pew Research, around 225.78 million American citizens were of legal voting age in 2016. The Statista portal estimates that in the USA there are around 67 million monthly active users on Twitter. Twitter data can thus become a significant source of information, with the potential to impact election outcomes owing to four overarching factors. First, the numbers presented above highlights that almost a quarter of the voting population of the USA is present on Twitter. Second, Twitter has been used by the presidential candidates to interact with the public and the media for reasons of public conversation (Shapiro and Hemphill, 2016; Vaccari and Valeriani, 2015; Waisbord and Amado, 2017). Third, Twitter is highly associated with non-personal engagement (Mosca and Quaranta, 2016). Finally, Twitter data has been used for electoral forecasting (Burnap et al., 2016), for indicating social tension (Burnap et al., 2015) and to estimate public engagement over the election period in various countries (Adams and McCorkindale, 2013;

Ahmed et al., 2016; Bode, 2016; Burnap et al., 2016; Ceron et al., 2014; Domingo and Martos, 2015; Ernst et al., 2017).

To the best of our knowledge, this study is the first within the political domain in which the social activity created by a presidential candidate's tweets were mapped to citizens' responses. The study aims to explore the following areas: (a) relationship between activity and engagement on social media platforms; (b) consecutive campaigns effects on popularity and engagement; (c) tweets sentiments effects on popularity and engagement; (d) relationship between drivers of voter's choice behavior and engagement on social media platforms; (e) acculturation of ideologies through hashtags; (f) opinion polarization of users within political deliberation and the subsequently formation of communities.

The contents of this study may position it within the sphere of computer-mediated communication and digital politics. The study contributes to the field through analyzing the social engagement from both the presidential candidate's and the voter's perspective. It presents the Twitter discussions concerning party policies and campaigning that, theoretically, may have led to the acculturation of political ideologies among voters, and subsequently to polarizations in voter opinion – thus potentially impacting the outcome of the 2016 election. In short, the *buzz* created by presidential candidates Twitter presence has been mapped according to the concept of acculturation of ideologies (i.e. hashtags) and opinion polarization within virtual communities.

The remaining sections are organized as follows. Section 2 summarizes a literature review regarding political communications, social media, polarization in elections, acculturation in social media and the usage of social media platforms for political communication, along with the knowledge gap identified, research questions and potential contribution of the study. Section 3 focuses on hypothesis development and contains the key sources identified by the literature review instrumental in hypothesis development. Section 4 illustrates the methodology for collecting and analyzing the tweets. Section 5 presents the results of the analysis of the tweets. Further discussions are presented concerning the contribution of the study, the implications to practice and policy, limitations and future research directions.

2. Literature Review

The literature review is divided into the five sections, namely political communication, social media, polarization, acculturation in social media, and how political actors are using social media for public communication. The last section of the literature review presents the knowledge gaps identified, research questions and the potential contribution of the study.

2.1 Political communication

Traditional media follows a model of unidirectional communication and offers asynchronous communications. In contrast, social media communication is multi-directional and offers interactive communication (Kruikemeier et al., 2016; Ross and Bürger, 2014). This facility of social media enables political discourse to shift from the traditional mass media to social media platforms like Facebook and Twitter (Heo et al., 2016). The use of the social media platforms in western democracies is very high for purposes of political communication (Mosca and Quaranta, 2016) and varies between countries due to factors such as broadband facilities, internet penetration, and media literacy (Klinger, 2013).

Politicians and journalists – through such online interaction – are emerging as both actors and sources of information (Ekman and Widholm, 2015). In this light, many have highlighted the significant role that social media plays in the modern media environment (e.g. Bode, 2016). Politicians have used social media for distributing information (Klinger, 2013; Ross and Bürger, 2014) and campaigning purposes (Jungherr, 2014); seeking to mobilize voters through drawing their attention to a party's agenda (Skogerbø and Krumsvik, 2015). Social media sites are emerging as journalistic sources (Ogola, 2015; Skogerbø and Krumsvik, 2015) and as a way to connect politically involved citizens to non-involved citizens in political discourse (Mosca and Quaranta, 2016).

Communication between like-minded users can strengthen a group identity, whereas communication between different-minded users leads to in-group and out-group affiliations (Yardi and Boyd, 2010). *In-group* refers to connections within the group to which a user already belongs, whereas *out-group* refers connections to a group which a user does not belong to (Iyengar and Westwood, 2015). In the deliberation of duos, one user rates their self-opinion more positively when other users are in support of opinion (Lee, 2007). Users with similar political views flock together (Gruzd and Roy, 2014; Kim, 2015; Lawrence et al., 2010; Lee et al., 2014; Yardi and Boyd, 2010). However, voters with little interest in politics have been shown to be ideologically moderate and can be polarized easily (Lawrence et al., 2010).

Research has further shown that the reach of protest messages increases through the use of social media platforms (Barberá et al., 2015) which can enable crowd mobilization (Ems, 2014; Theocharis et al., 2015). Communication on social media gets accelerated (Ernst et al., 2017; Poell, 2014) and user-generated content within small span of time reaches to thousands of people present on social media platform (Heo et al., 2016).

2.2 Social media

Social media data (i.e. user-generated content) has been extensively used in the analysis of issues such as electoral forecasting (Burnap et al., 2016), engaging with voters (Adams and McCorkindale, 2013), identifying social tensions (Burnap et al., 2015), evaluating voting intentions (Ceron et al., 2014) and measuring behavior transition in national events (Lakhiwal and Kar, 2016). Domain-specific understanding may be developed by analyzing user-generated content through the use of social media analytics (Aswani et al., 2017;2018; Grover et al., 2017; Rathore et al., 2017; Joseph et al., 2017) using big data analytics (Grover and Kar, 2017; Gupta et al., 2018).

Twitter has been used for announcing and promoting awareness of various public policies, such as campaigns regarding electronic cigarettes (Harris et al., 2014), early warning announcements concerning natural hazards (Chatfield et al., 2013), understanding social sensitivity towards the environment (Cody et al., 2015) and emergency management (Panagiotopoulos et al., 2016; Singh et al. 2017). Voters have also used Twitter for seeking and sharing information related to social support (Yardi and Boyd, 2010). The potential for using Twitter to uncover unbiased information from user-generated content was one of the drivers behind using Twitter data in our study.

The hybrid of television and social media can lead to positive outcomes regarding democratic engagement in elections (Chadwick et al., 2017). Literature indicates online engagement on social media impacts user's sentiments (Ibrahim et al., 2017). Highly engaged users are often highly educated followers (Scott et al., 2017) belonging to higher socio-economic equity. Post tagged with the hashtags influence users more as compared to untagged posts (Chadwick et al., 2017)

2.3 Polarization in elections

Polarization can be defined as a state as well as a process (DiMaggio et al., 1996). Polarization is a state in which an opinion on an issue has generated an opposing opinion to a theoretical maximum value. Polarization is a process whereby this opposition increases over the time. In this study, polarization had been treated as a state. The study considers two states (positive and negative) of polarization. A voter is in the positive state when the voter holds a positive opinion of the presidential candidate. Similarly, a voter is in the negative state when the voter holds a negative opinion of the presidential candidate. Opinion polarization is relevant in fields of political conflict and social volatility (DiMaggio et al., 1996). Existing literature indicates that polarization within American society has increased over the past four decades (Iyengar and Westwood, 2015).

DiMaggio et al. (1996) highlight four dimensions of the polarization: dispersion, bimodality, constraint, and consolidation. *Dispersion* takes into the account the diversity of the opinions among the public. As dispersion of opinions increases among voters, difficulty in establishing and maintaining a consensus within the political system also increases. *Bimodality* refers to polarization occurring between opinions; the authors suggesting that people with different positions cluster into separate camps regarding an issue. *Constraints* consider whether the extent of opinion is associated with any other opinions within an opinion domain. *Consolidation* refers to differences in the responses to an issue on the basis of demographics such as gender, race, occupation, age, graduation, and income. DiMaggio et al. (1996) surmise that opinion polarization increases when opinion distribution becomes dispersed, bimodal, closely associated and closely linked to social identities.

Political leaders act as the polarizing cues for voters (Nicholson, 2012). Iyengar and Westwood (2015) suggest that followers of a presidential candidate – those present on social media – can play a significant role in polarizing the political choices of voters. Political polarization towards party is strong as race polarization (Iyengar and Westwood, 2015). Polarization stimulates voters towards political participation (Abramowitz and Saunders, 2008). Polarization among in-group leaders tends to decrease voters' trust in the party (Layman et al., 2006).

In attempting to explain political polarization, authors have described what is termed the *echo chamber* effect of social media platforms (Gruzd and Roy, 2014; Iyengar and Westwood, 2015; Lawrence et al., 2010). This refers to the environment in which voters are exposed only to information and communities that support and reinforce their views and opinions. Some authors, however, have sought to downplay this effect, offering the opinion that suggests that the use of social media for political news distribution and policy-based deliberation by the voters can lessen any echo chamber effect since discussions take place in open platforms and are accessible to all (Lee et al., 2014).

Public self-awareness increases group polarization within communities (Lee, 2007). Group polarization can be enhanced within the user with group discussions (Chadwick et al., 2017; Isenberg, 1986). Disagreement of the user was negatively associated with group polarization (Kim, 2015). The group has the potential of creating or distorting a user's opinion (Zhu, 2013; Moscovici and Zavalloni, 1969). Literature indicates group opinions had been often adopted by individuals as their personal opinion (Lee, 2007; Moscovici and Zavalloni, 1969). Demographic homogeneity and minority expertise reduce group polarization (Zhu, 2013).

On Twitter, various social groups participate in discussions - leading to diversity in opinions (Yardi and Boyd, 2010). Divergence in opinion may increase the representativeness or breadth of governmental policies, leading to a healthy democracy (Hollander, 2008; Layman et al., 2006). Isenberg (1986) found that argumentation effects tend to be larger than social comparison in seeding polarization among social groups. From above literature evidences it can be concluded that social media has the potential of exposing voters to both sides of an argument (i.e. positive and negative), which can lead to opinion polarization among voters, resulting in the amplification of division between social groups holding different views (Lee, 2007).

2.4 Acculturation in social media

Acculturation has been defined as the occurrence of a change in preferences within an individual when exposed to individuals or groups from a different cultural background (Redfield et al., 1936). Various interpretations and caveats to this definition exist. Ferguson (2017), for examples, extends the definition to include what he calls *remote acculturation*: changes experienced by individuals having only intermittent contact with a geographically separate culture. The overarching view across definitions, however, sees acculturation as a process of altering individual identity by exposing them to new ideas through geographically dispersed individuals or groups. This is the definition of acculturation adopted by this study.

Ogden et al. (2004), describes acculturation both at an individual and group level. The writers further identified a series of characteristics of acculturation on both an individual and group level. Changes in perception, attitudes, values, and personality are described as important on an individual basis, whereas group level acculturation characteristics included relationship to socialization, social interaction, and mobility. Ogden et al. (2004) further describe three phases of acculturation: contact, conflict and adaptation. In phase 1 (*contact*), an individual comes into contact with an individual or group of differing ideology, resulting in *conflict* (phase 2) of opinion, and subsequently adaptation (phase 3) of the majority opinion. Acculturation also leads to psychological changes within an individual (Berry, 2008) and influences their behaviour, values and identity (Ferguson et al., 2017).

Berry (1997) suggests four strategies for the process of acculturation: assimilation, separation, integration and marginalization. *Assimilation* is a strategy where an individual belonging to a non-dominant group – who does not wish to maintain their cultural identity – interacts frequently with the dominant group. In contrast, *separation* describes a situation where an individual seeks to retain their values and tries not to interact with other cultures. When both the groups seek to maintain their cultural values but also wish to interact with other groups, a strategy of *integration* is followed. For groups less interested in maintaining their cultural preferences and less interested in maintaining relationships with another group, a *marginalization* strategy is followed. Changes primarily impact the minority group, which is then expected to become more like the majority group (Berry, 2008).

Acculturation theories have been applied to the political domain by Hindriks et al., (2016), in a study of native majority and immigrant minority populations. Their results indicate that (a) using a political assimilation strategy, the interests of only the major groups advance; whereas (b) with a strategy of political integration, the interests of a majority group advances along with those of a minority group; and (c) using a political separation strategy, the interests of the minority group only advance.

Authors have also described how the media can be an important mechanism for remote acculturation (e.g. Ferguson et al., 2017). The branch of the media used by this study for mapping acculturation is the social media platform Twitter. In this study, individual level acculturation had been measured through examining the perceptions of, and attitudes towards, a presidential candidate. Communications taking place on social media have the potential to strengthen or weaken the perceptions and attitudes of users (Croucher, 2011; Li and Tsai, 2015; Mao and Yuxia, 2015).

There are numerous studies that have examined the process of acculturation due to the influence of social media platforms, and various user groups have been studied: Chinese professionals overseas (Mao and Yuxia, 2015), Hispanics in the US (Li and Tsai, 2015), international students (Cao and Zhang, 2012; Forbush et al., 2016), and Lebanese nationals residing in French-speaking urban areas (Cleveland et al., 2009). It seems from the literature that geographical divergence among communities can lead to the acculturation of ideas.

2.5 Political Communication and Social Media

Politicians use social media platforms like Facebook and Twitter for professional communication (Kelm et al., 2017). Political campaigning through social media campaigning can be of two broad styles: party-centric or individually targeted (Karlsen and Enjolras, 2016). Political information shared and discussed on social media engages young people (Vromen et al., 2015). Evidence further suggests that the degree of social media buzz created by political parties has impacted the outcome of general elections in emerging economies such as India (Safiullah et al., 2017).

Microblogging services provide opportunities to politicians with respect to disseminating information, engaging with voters, monitoring public opinion, and making public relations (Frame and Brachotte, 2015; LaMarre and Suzuki-Lambrecht, 2013). If voters acquire political information via social media channels and respond to that information, this increases the likelihood that they will go on to contact politicians and attend offline events (Vaccari et al., 2015). Officials active on social media have more contacts as compared to less active officials (Djerf-Pierre and Pierre, 2016). Therefore, politicians use social media platforms for communication, engagement with voters and marketing purposes. For marketing purposes, Facebook is often the preferred tool, whereas for continuous dialogue Twitter is often preferred (Enli and Skogerbø, 2013). National Assembly members in Korea used Twitter to communicate with fellow politicians rather than with their constituents (Hsu and Park, 2012). Twitter can also be used as a tool for political opposition by politicians (Van Kessel and Castelein, 2016).

Political actors in Western democracies are increasingly using Twitter and Facebook for populist communication (Ernst et al., 2017) and are able to freely circulate their messages and ideology through the use of social media platforms (Engesser et al., 2017). A political leader using Twitter and Facebook receives considerable attention on these platforms (Larsson, 2017).

Twitter has also been used by politicians for broadcasting purposes (Hutchins, 2016; Theocharis et al., 2016), advertising (Domingo and Martos. 2015; Hutchins, 2016) and for engaging with citizens (Ahmed et al., 2016). LaMarre and Suzuki-Lambrecht (2013) have, furthermore, been able to show that Twitter usage by politicians increases their chances of winning an election. The adoption of

Twitter by presidential candidate is conditioned at a personal level (Scherpereel et al., 2017) and driven by candidate's age (Rauchfleisch and Metag, 2016).

Twitter is used by established political parties as well as new and upcoming parties for political communication. Established parties use Twitter to supplement offline strategies, whereas newer political parties use it more for self-promotion and media validation (Ahmed et al., 2016). Politicians who maintain the synergy between social media platforms and traditional media channels can act as influencers on social media platforms (Conway et al., 2015; Karlsen and Enjolras, 2016). The more the politician is active on social media, the more journalism and press the politician receives (Rauchfleisch and Metag, 2016).

2.6 Knowledge Gap

To the best of our knowledge, no study in the existing literature has mapped a presidential candidate's Twitter impact among voters. Further the role of social media in affecting the voting communities has never been explored. Following extensive literature review, four specific knowledge gaps have been identified. These knowledge gaps are listed below: (a) to measure the impact of presidential candidate's tweets on popularity and engagement among followers on Twitter; (b) how political ideologies become acculturated using hashtags on Twitter; (c) how opinion polarization occurs among voters on Twitter; (d) how opinion of a voter plays a role in formation of the communities on Twitter.

The knowledge gaps identified have been visually represented in Figure 1 with the help of four scenarios. Therefore, the first knowledge gap - the specifics of a candidate's tweets - leads us to Scenario 1, which attempts to measure and characterize a presidential candidate's tweets with respect to activity, consecutive campaigning, sentiments expressed, issues and policies discussed on Twitter. The second knowledge gap, concerning how political ideologies become acculturated, leads us to Scenario 2: mapping political deliberation among geographically dispersed voters using hashtags reflecting the activities of the presidential candidate on Twitter. The third knowledge gap, how opinion polarization occurs among voter (Scenario 3), requires us to attempt to map voter polarization. We hypothesize voter polarization - potentially caused by voter acculturation of ideologies - may have subsequently lead to the formation of communities among voters (Scenario 4).

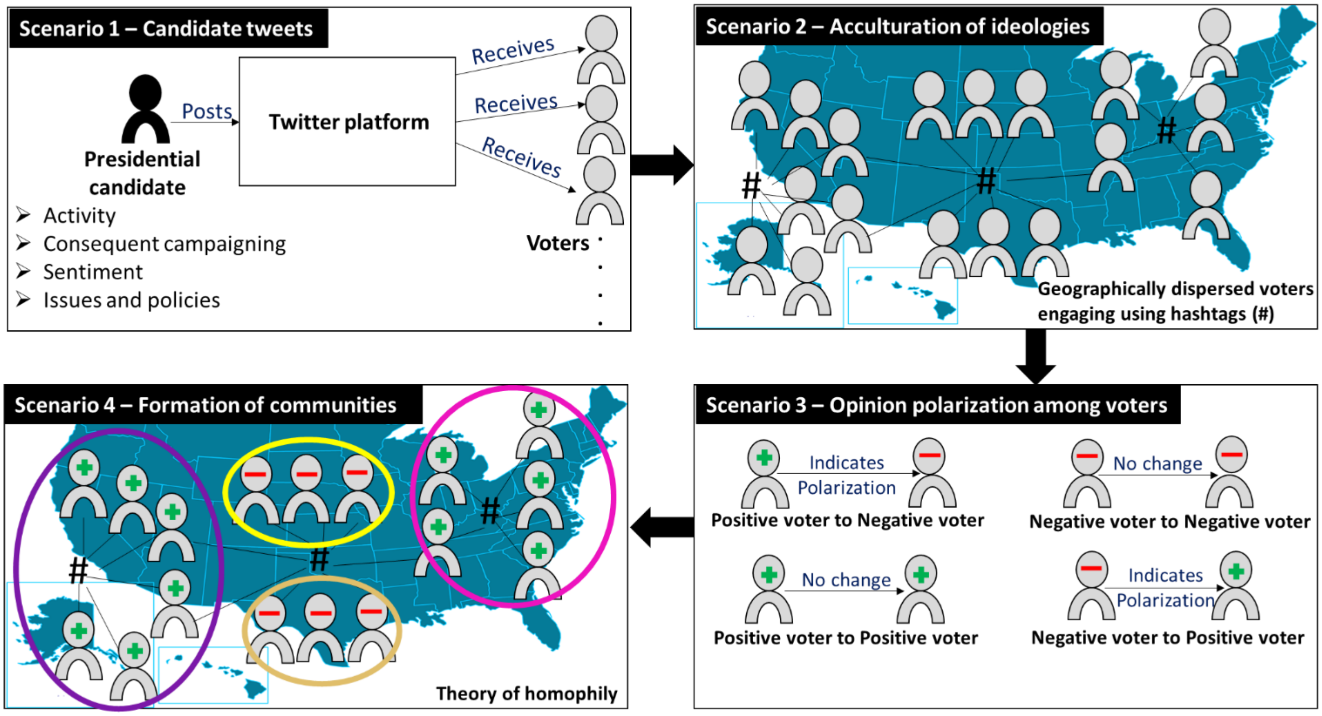

Figure 1. Pictorial representation of knowledge gaps identified for study.

We elaborate on these knowledge gaps in the subsequent subsections, and use them to develop research questions and hypotheses, we attempt to validate through our study.

2.7 Research questions and major contributions

The primary focus of the study is to explore deliberation surrounding the 2016 US election that took place via a social media platform (Twitter), and how these deliberations could have resulted in the acculturation of ideologies and subsequent voter polarization, as illustrated in Figure 1. This study is constructed around three research questions (RQ1, RQ2 and RQ3), listed below:

RQ1: Is the frequency of social media use related to popularity and engagement? Are the topics discussed by Trump more popular than the topics discussed by Clinton on Twitter?

RQ2: How are the drivers of voter's choice behaviour being discussed in the Twitter ecosystem? How do these drivers affect the outcome of the election?

RQ3: Does acculturation have an impact on polarization? What is the nature of this polarization? Do voters undergo transition and polarization of their preferences through Twitter over the course of an election?

In order to answer these questions, the study will analyze tweets using social media analytics such as descriptive analysis, content analysis and network analysis (Chae, 2015) along with data mining approaches such as regression analysis and community detection (Fortunato, 2010). Details of this are provided in subsequent sections. The study showcases how voter engagement occurs on the social media platform during the election period among the different stakeholders in virtual communities. The study also highlights the role of Twitter features such as hashtags, @mention, retweets, and likes, and how these features are being used in political communications. Future political actors can then use the results of the study for planning digital campaigns over the Twitter platform.

3. **Hypotheses Development**

On Twitter, voters are exposed to a diversity of opinions surrounding events and issues (Lee et al., 2014; Yardi and Boyd, 2010). Research indicates that diversity and deliberation are critical components of the online society; therefore, potential voters witnessing deliberations on social media platforms try to participate in it (Yardi and Boyd, 2010). This leads to voters forming connections to other voters with similar ideologies (Gruzd and Roy, 2014): leading to the formation of communities.

Higher activity on Twitter leads to higher visibility, leading to an increased number of online discussions among voters. These discussions can polarize voters towards a candidate and ultimately result in a candidate winning the election (Kruikemeier et al., 2016; Larsson and Moe, 2012). Research shows that the frequency of posts on Twitter is related to voter engagement (Scherpereel et al., 2017). Tweet influence can be measured in terms of the number of followers the author has within their network (Moya-Sánchez and Herrera-Damas, 2016). The *reach metric* (shown in Table 1) attempts to quantify the reach of a political message (Ganis and Kohirkar, 2015).

A candidate who engages heavily with voters on social media platforms is likely to be exposed to more to criticism and harassment (Theocharis et al., 2016). Higher activity on social media can be related to both increased popularity and engagement, but the opposite can also be true, and higher activity on social media can also be negatively related to popularity and engagement among followers (Rauchfleisch and Metag, 2016). Therefore, to examine how social media activity is related to popularity and engagement among followers in the 2016 US election, the first hypothesis looks to test if:

H1: Higher activity on social media is positively related to higher popularity and engagement among followers.

Literature indicates society can radicalize ideas within individuals through communication (Moscovici and Zavalloni, 1969). Campaigns encourage communications on Twitter through responding, retweeting and engaging (Jensen, 2017). Citizens can relate to consecutive campaigns with ease (Iyengar and Westwood, 2015). Campaigns organized at a national level receive more attention than local campaigns (DiMaggio et al., 1996). On Twitter campaigns had been associated with hashtags. Political engagement through hashtags had been considered as most consistent (Chadwick et al., 2017; Vaccari et al., 2015).

Communicative exchanges can be easily tracked using hashtags. Research indicates that the use of free-text on Twitter has a stronger correlation to voting outcomes compared to @mention use (McKelvey et al., 2014). Regular tweeting helps to sustain voter interest in social media campaigns (Mills, 2012), although this has not been established empirically. Therefore, the second hypothesis (H2) attempts to explore whether the frequency of tweets during the election period is of importance, and assists in information propagation.

H2: Less time between consecutive campaigns is positively related to higher popularity and engagement.

Deliberation and argumentation in the online environment mostly surround political news, emotionally charged tweets or controversial issues (Yardi and Boyd, 2010). Some accounts (influencers) play a more significant role in disseminating this information in the social network. Furthermore, tweets with more emotionally charged content may be retweeted more than neutral tweets (Stieglitz and Dang-Xuan, 2013). High Twitter usage by the elected candidates during an election period is likely to increase voter loyalties towards the party (Gruzd and Roy, 2014). Therefore, this hypothesis (H3) attempts to explore whether greater levels of polarity and emotions expressed in tweets have a positive or negative impact on buzz in social media platforms (Twitter).

H3: Higher thresholds of sentiments (polarity) within tweets is positively related to higher popularity and engagement among followers.

Newman and Sheth's model of voter's choice describes seven factors which drive the voter's behaviour in the physical world. The drivers of voter's choice behaviour described by the authors are issues and policies, social imagery, emotional feelings, candidate image, current events, personal events, and epistemic issues (Newman and Sheth, 1985). This model has been widely applied in examining voter's choice behaviour in empirical surveys. However, the utility of this model in analyzing user-generated digital content has not been explored. Therefore, in this study we try to translate model components into the virtual environment using twitter analytics, to determine whether the discussions surrounding these factors are initiating polarization and acculturation processes among voters.

Twitter has been used by candidates to interact with voters (Graham et al., 2013), and voters actively participate in election-orientated discussions on Twitter (Raynauld and Greenberg, 2014). The discussions surrounding these seven domains of voter's choice behaviour can highlight how the Twitter users gets impacted in the virtual world. The drivers of voter's choice behaviour are explained through Twitter analytics in this study.

H4: Greater levels of social discussion – concerning the components of Newman and Sheth's model of voter's choice behaviour – increase engagement among voters, actively or passively.

Mao and Yuxia (2015), in their study of Chinese professionals overseas, show how groups have been able to use Facebook as an acculturation tool for acquiring information regarding contemporary topics in their host countries. Specific to voting populations, Twitter hashtags and internet campaigns have further been shown to influence users political views (Bode et al., 2015; Kruikemeier et al., 2016; Larsson and Moe, 2012; Wu, 2014). Twitter has been used by candidates for purposes of mobilizing their campaigns and for directly interacting with voters (Borondo et al., 2014; Bode et al., 2015; Chae et al., 2015; Graham et al., 2013; Gruzd and Roy, 2014). Prior research has shown that social media platforms are useful in the acculturation process (Li and Tsai, 2015).

Our next hypothesis (H5) is designed to explore how hashtags or campaigns contribute towards the acculturation process among Twitter users located in different geographical locations.

H5: Popular hashtags or campaigns initiate a process of acculturation of ideologies among Twitter users located in different geographical locations.

Voters on Twitter are exposed to a diversity of opinions which, in turn, allows voters to explore and refine their own opinions (Lee, 2007). Political deliberation moderates the relationship between network heterogeneity and ideological polarizations (Lee et al., 2014). Furthermore, In-group leaders can be highly persuasive in these groups (Nicholson, 2012). Kim (2015) suggests that the frequency of voter's participation in deliberation on social media platforms is negatively related to polarization. The social media buzz created by political parties had been shown to result in their favour in terms of votes in an election (Safiullah et al., 2017). Indeed, some electoral campaigns have resulted in only minimal public attention (Hong and Nadler, 2012). Furthermore, polarization may seem to increase even when, in reality, it does not (DiMaggio et al., 1996).

Given the conflicting evidence, it appears debatable as to whether voters can become polarised in the virtual environment, and concrete evidence of polarization is missing from the existing literature. Therefore, this hypothesis (H6) attempts to explore the impact of political deliberation on opinion polarization:

H6: Political deliberation on a social media platforms (Twitter) leads to opinion polarization among users.

Users may potentially be polarized through campaigns, tweets or discussions surrounding the candidate. Polarization is the process by which users undergo a transition of opinion. In this study opinion polarization of Twitter users were tracked from phase 1 to phase 2. This study treats polarization as a state. Two states consider in the study are positive and negative. A voter holds the positive state when he/she has a positive opinion towards presidential candidate. A voter holds the negative state when he/she has a negative opinion towards presidential candidate. In this case, opinion polarization of Twitter users were tracked from phase 1 to phase 2 (positive to positive, positive to negative, negative to positive, negative to negative).

Internet communication has the potential to fragment populations by engaging users (Lawrence et al., 2010). Voters may form their opinions both according to personal, closely held beliefs and in opposition to beliefs that threaten their core values (Hollander, 2008; Kim, 2015). Demographically, men tend to be more politically neutral on social media whereas women tend to be more opinionated on social media platforms, with young people expressing a higher proportion of negative opinions and emotions than older users (Volkova and Bachrach, 2015). Through hypothesis (H7), we attempt to explore how polarization effects formation of communities among voters.

H7: Communities are formed among groups of users polarized during social media discussions, **around political events such as elections.**

Social media users have been shown to cluster into politically homogeneous networks (Borondo et al., 2014). *Homophily* is a central idea in the study of social networks. Himelboim et al. (2016) describe this phenomenon in relation to online political discourse, whereby individuals try to associate themselves with similar users on the social network. This leads to the formation of clusters within the virtual communities (Yardi and Boyd, 2010). Users within these communities are unlikely to be exposed to ideologies from different groups (Himelboim et al., 2013). However, social media is able to – more generally – open up the potential for cross-cultural interaction (Gruzd and Roy, 2014; Li and Tsai, 2015).

4. Research Methodology

A social media analytics framework, for use in the political domain, was adopted from the work of Stieglitz and Dang-Xuan (2013). This framework consists of two parts: data tracking and monitoring, followed by data analysis. The tweets constituting the raw data were extracted through Twitter's APIs (application programming interfaces) over a timeframe of four months. Tweets can be tracked via user timeline, keywords, topics, hashtags, and URL. The data can be extracted from social media using API functions such as ''search API'' and ''streaming API.'' The framework used illustrates that social media data can be analyzed using content analysis, opinion mining, social network analysis and sentiment analysis (Stieglitz and Dang-Xuan, 2013). Twitter allows users to download data posted or discussed around a search term within a particular period. This data can then be analyzed for deriving metrics and developing more in-depth insights.

Techniques for quantitatively comparing communicative patterns on Twitter have been previously described (e.g. Bruns and Stieglitz, 2013; Chae, 2015). A full list of methods used by this study for purposes of Twitter analytics is given in Table 1. This comprehensive overview of Twitter analytics is among the contributions of this study, as, to the best of our knowledge, this has not been attempted before in the scientific literature.

The Twitter analytics have been divided into four broad categories: descriptive analytics, content analysis, network analysis, and geospatial analysis. The descriptive analysis incorporates basic descriptive statistics, such as the number of and types of tweets, number of individual users, hashtags, frequency of @mention and hyperlink modifiers added to tweets, word cloud, and reach metrics. Word clouds help us to visualize the popular words/topics in tweets (Nooralahzadeh et al., 2013). The *reach metric* can be used as a way to measure the reach of the messages (Ganis and Kohirkar, 2015). Similarly, the *reply* and *retweet* features of Twitter allow for measurement of two-way interaction and engagement (Purohit et al., 2013). Hashtags are used in tweets so that the tweet can be shared across a broader community of similar interest (Chae et al., 2015). Similarly, the @mentions analysis helps in identifying the influencers who had influenced users to the extent that they wish to engage in discussion with the influencer on the tweet topic (Shuai et al., 2012).

Content analysis is used to extract the semantic content from text data. It uses principles from natural language processing (NLP) and text mining (Kayser and Blind, 2017) in order to retrieve information from a large amount of text data (Kassarjian, 1977). For example, sentiment analysis is the process of computationally identifying and categorizing opinions present in the text (Zhang et al., 2016). It consists of two analytical components: polarity analysis and emotion analysis. For this study, sentiment analysis of the tweets was performed with R (programming language), using syuzhet, lubridate, and dplyr libraries. Polarity analysis is one of the most commonly used techniques for analyzing Twitter data; classifying the opinions of the users in terms of positive, negative, and neutral. Emotion analysis is a technique in which user-generated content is classified into eight emotions, namely anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Volkova and Bachrach (2015). Topic modelling identifies the key themes within the tweets (Llewellyn et al., 2015). Topic modelling was performed using the *tm* and *topicmodels* libraries of R.

Connections among Twitter users can be visually depicted through the identification of *networks* (HerdaĞdelen et al., 2013; Stieglitz and Dang-Xuan, 2013). Networks analysis further allows us to identify communities and clusterings of users on the basis of their opinions and thoughts expressed on social networks (Abascal-Mena et al., 2015). Information flow on social media can, therefore, represent the information flow within and among these networks (Park et al., 2015).

Geospatial analysis was divided into two broad categories: location-specific analysis, and time-trend specific analysis. The time-trend analysis allows us to examine the evolution of topics and trends over the period of time (Saboo et al., 2016). Geospatial analysis helps us in mining location specific opinions (Stephens and Poorthuis, 2015; Attu et al., 2017).

Table 1: Overview of Twitter analytics method.

Twitter analytics methods	
Descriptive Analytics	
• Retweet (Bode et al., 2015; Yardi and Boyd, 2010;)	Allows one follower to share someone else's tweet.
• URL analysis (Stieglitz and Dang-Xuan, 2013)	Allows users to disseminate information by including the URL within the 140 character tweet.
• Hashtags analysis (Borondo et al., 2014; Bode et al., 2015; Chae et al., 2015)	Hashtags are user-generated keywords preceded by the # symbol, allowing users to cluster opinions.
• @mentions analysis (Borondo et al., 2014; Larsson and Ihlen, 2015; Shuai et al., 2012)	@mentions allow users to draw an individual's attention to a discussion topic (and helps in promoting one to one discussions on Twitter).
• Word cloud (Nooralahzadeh et al., 2013)	Pictorially represents the most frequent words used in Twitter discussions.
• Reach metric (Ganis and Kohirkar, 2015)	Measures the reach of the tweets.
Content Analysis	
• Sentiment analysis (Burnap et al., 2015)	Identifies and categorizes opinions present the text.
I. Polarity analysis	Categorizes user opinions in the text into positive, negative, and neutral.
II. Emotion analysis	Categorize the tweets on the basis of the emotions expressed.
• Topic modelling (Llewellyn et al., 2015)	Identifies the key themes within the text.
Network Analysis	
• Network analysis (HerdaĞdelen et al., 2013; Stieglitz and Dang-Xuan, 2013)	Connection among the users
• Cluster/ Community detection (Abascal-Mena et al., 2015)	Identifies different communities among users.
• Information flow networks (Park et al., 2015)	Depicts the flow of the information across a network.
GeoSpatial Analysis	
• Time-trend analysis (Saboo et al., 2016)	Temporal analysis of trends or topics.
• Geospatial analysis (Attu et al., 2017; Stephens and Poorthuis, 2015)	Location specific analysis

To test our hypotheses, we retrieved data from Twitter – over a period of 120 days – in two main ways. First, daily Twitter searches were performed using the search terms 'USA election', 'Hillary Clinton' and 'Donald Trump', concatenated by 'OR'. Only tweets that were generated within the USA have been included in the analyses. Second, we extracted Twitter timeline data of 'Hillary Clinton' and 'Donald Trump'.

This study uses social media analytics applied to 784,153 tweets, derived from 287,838 users, to attempt to gain insights into changes in voter opinion over the election period, and the specific topics shared and discussed via Twitter. For each tweet, 46 parameters – focusing on the user demographics and tweet characteristics – were analyzed. User demographics captured included name, location and description. Tweet characteristics captured included tweet content, language, retweet count, like count, and status updates. The results from the analysis of tweets were also used to explore and assess the drivers of the outcome of the election.

For the first part of the data extraction, the methodology sub-divides into five-phases (Figure 2). Phase 1 identifies the search terms with which to extract data from Twitter. For this study, the election-related search terms 'USA election', 'Hillary Clinton' and 'Donald Trump' were identified based on Twitter trends. Phase 2 of the study focuses on extracting the data from Twitter. The unstructured data were collected through the Twitter API using Python scripts in JSON format. Phase 3 of the study converts the unstructured data to structured data, i.e. JSON to the structured Excel format. The steps of Phases 2 and 3 were repeated daily over the 18 weeks to extract the data from Twitter; Gonzalez-Bailon et al. (2014) having previously shown that small, online samples do not give an accurate representation of activities on Twitter. Phase 4 is concerned with deriving meaningful insights from the data, through the analytical methodologies described in Table 1. Phase 5 explains the impact of the findings in the framework of Newman and Sheth's model of voter's behaviour, using the seven concepts of issues and policies, social imagery, emotional feelings, candidate image, current events, personal events, and epistemic issues.

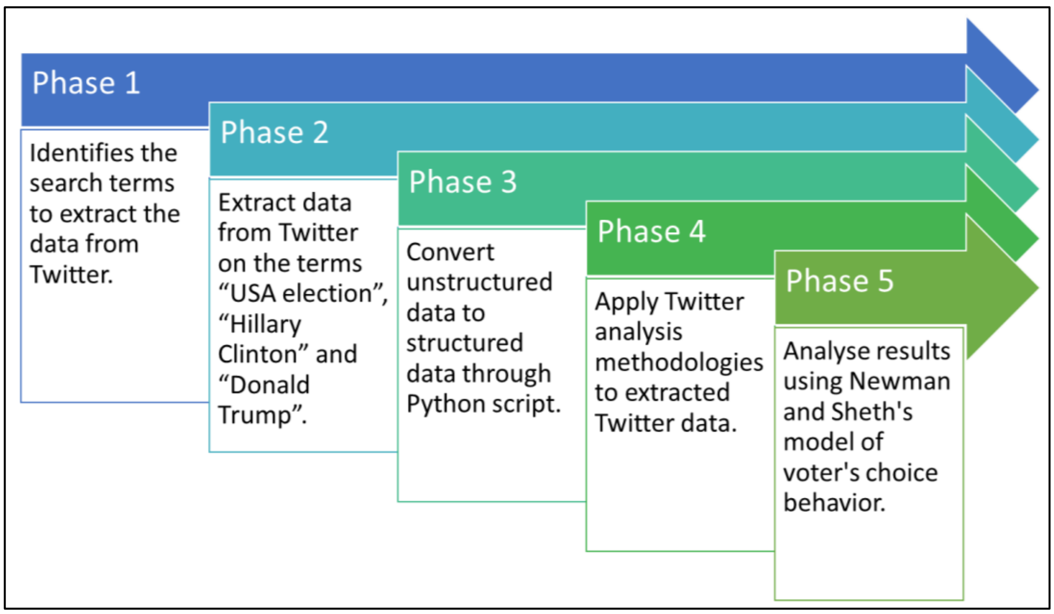

Figure 2. Methodology followed.

5. **Findings and Interpretation**

This section is divided into three sub-sections. Section 5.1 illustrates the way the Twitter handle was used by the presidential candidates. Section 5.2 shows the impact of Twitter users on topics discussed by the presidential candidates using Newman and Sheth's model of voter's choice behaviour. Section 5.3 shows the user communities formed, defined using hashtags.

5.1 Tweet frequency and its impact

To address our first three hypotheses (H1, H2 and H3), all tweets from each candidate's Twitter screen were extracted, providing an overview of the respective campaigns over the election period (13 August - 10 December 2016). We analyzed the screen data in two ways: (1) through hashtag analysis, and (2) by counting the numbers of retweets and likes to estimate user engagement and opinions. Insights derived from tweets are described using the SPIN Framework (Mills, 2012). SPIN frameworks indicate the spreadability and propagativity of tweets among Twitter users.

H1: Higher activity on social media is positively related to higher popularity and engagement among followers.

Spreadability refers to the ease with which campaigns can spread across the Twitter ecosystem. Likes and retweets help a tweet to spread across the various networks within Twitter (Mills, 2012). A descriptive overview of the Twitter activity of the 2016 US presidential candidates is presented in Table 2, which illustrates the degree of spreadability of both candidates Twitter campaigns among Twitter users.

Table 2. Descriptive statistics of activity and engagement.

	Retweet count		Like count	
	Clinton	**Trump**	**Clinton**	**Trump**
Total Tweets	2,400	1,227	2,400	1,227
Minimum activity / tweets	175	1,792	0	0
Maximum activity / tweets	665,370	345,548	1,197,489	634,112
Mean activity / tweets	4619.51	12,439.78	8,617.21	32,749.12
Std. dev. of activity / tweets	16,190.92	14,256.63	31,359.86	37,376.37

From Table 2, it may be inferred that a higher frequency of tweets leads to higher visibility and social presence (from Fig. 11). This is in accordance with existing research. The Clinton campaign was tweeting twice as much as the Trump campaign but went on to lose the election, despite previous research indicating that higher frequency of tweets lead to positive outcomes in elections (Kruikemeier et al., 2016; Larsson and Moe, 2012;). Clinton was exposed to numerous and frequent criticisms over the election campaign which was derived using URL analytics presented in annexure. Prior research has also provided evidence for a detrimental impact of high activity in social media (Karlsen and Enjolras, 2016; Theocharis et al., 2016). Interestingly, the mean *retweet* count of Trump is almost twice that of Clinton, whereas the mean *like* count of Trump is almost 3.8 times that of Clinton. In the following sections, we attempt to explore how this outcome may have occurred.

Propagativity refers to the ease with which tweets can be redistributed, or propagated, among voters, taking into account cycle time, network size (i.e. number of followers), content richness and content proximity (Mills, 2012). 441,261 tweets were collected using the search term 'USA Election', 258,212 tweets were collected using the search term 'Hillary Clinton', and 84,680 tweets were collected with the search term 'Donald Trump'. The difference in the number of tweets collected between campaigns is likely to be because Clinton posted approximately twice the number of the tweets as Trump. Figure 3 shows that the Trump campaign posted more regularly on Twitter, though the buzz created by the Clinton campaign was higher.

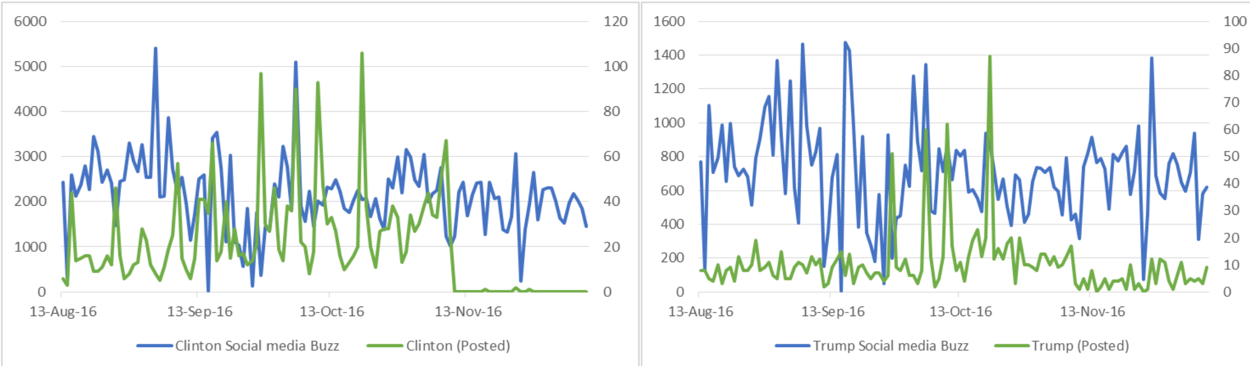

Figure 3. Tweeting frequency vs social media buzz.

The primary axis of Figure 3 represents the social media buzz of the candidate and the secondary axis depicts the number of tweets on the candidates' screen on each day. Trump had 17.6 million followers on Twitter, producing 34,160 tweets over the 120 days, whereas Clinton had 11.7 million followers, totalling 9,838 tweets over the 120 day period. Regression analysis shows that the buzz (Y) may be modelled using regression against user activity (X): **(a) For Clinton Y = 3.122*X + 2089 (b) For Trump Y = 1.989*X + 685.3.** It appears that Hillary Clinton had more reach than Donald Trump.

H2: Less variation in time (greater nexus) between consecutive campaigns is positively related to higher popularity and engagement.

Twitter campaigns are launched with the help of the hashtags. Online campaigns using hashtags are cost-effective for presidential candidates, and the hashtags provide metadata regarding the campaigns (Abascal-Mena et al., 2015). We use hashtags to explore how the respective Twitter campaigns were run by each presidential candidate. Figure 4 presents the frequency of hashtag campaigns used by the presidential candidates, along with the periodicity mean, periodicity standard deviation, retweet (10K), retweet mean (10K), retweet standard deviation (K), favorite sum (10K), favorite mean (10K) and favorite standard deviation (K). In this figure K stands for 1000 in number of retweets and likes (denoted by favorite).

The Trump team consistently incorporated campaign hashtags (#maga; #draintheswamp; #bigleaguetruth) into their Tweets, whereas the Clinton team did not. The use of campaign hashtags in Trump's tweets may have led to the higher campaign polarization among users – and higher voter participation using these hashtags – further propagating the core message of his campaigns.

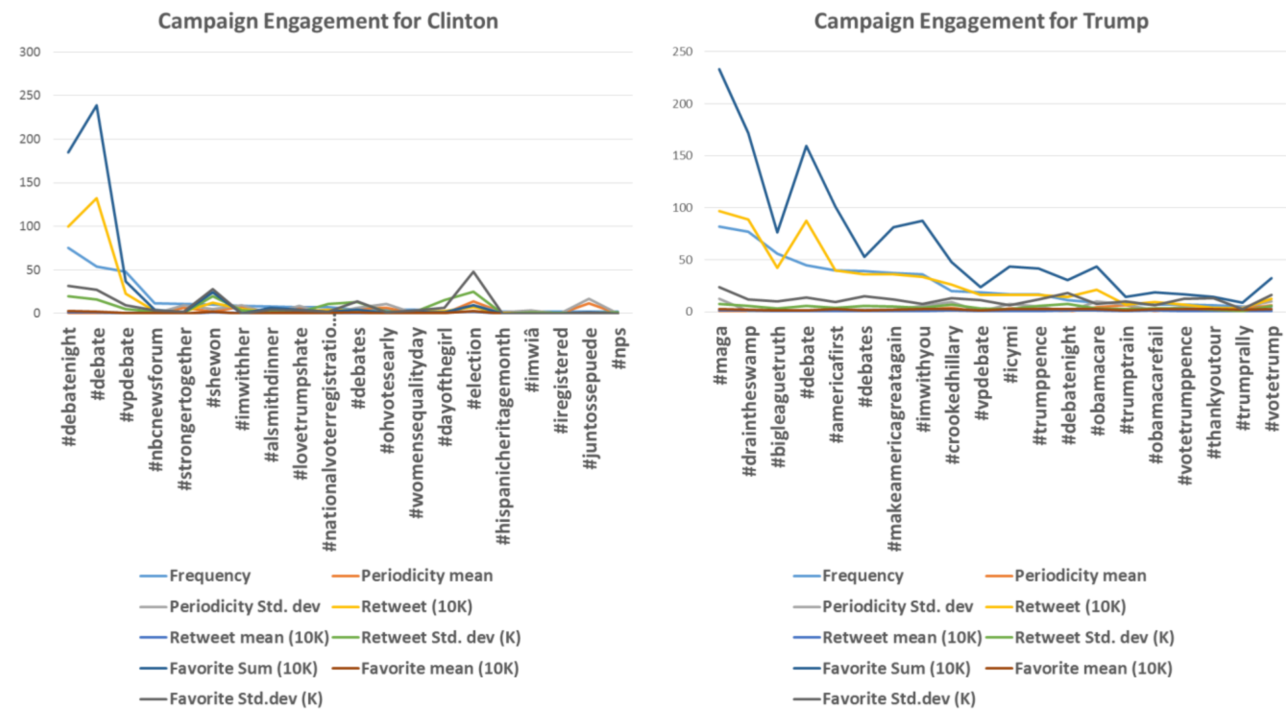

Figure 4. Top hashtags used by Clinton and Trump in their tweets during the election period.

H3: Higher thresholds of sentiments (polarity) within tweets is positively related to higher popularity and engagement among followers.

We subsequently looked to explore whether higher levels of polarity and emotions expressed in tweets have a positive impact in creating social media buzz. Figure 5 shows that, in absolute numbers, the Clinton campaign expressed higher levels of sentiment in tweets. When these statistics are compared by percentage, there is a substantial difference in the 'surprise' sentiment of tweets, with Clinton scoring 49.88% and Donald Trump scoring 25.51%. Clinton appears to have described more *surprises* to users - potentially resulting in the increased social buzz as indicated in Fig. 3. This is in line with existing research (Berger and Milkman, 2012).

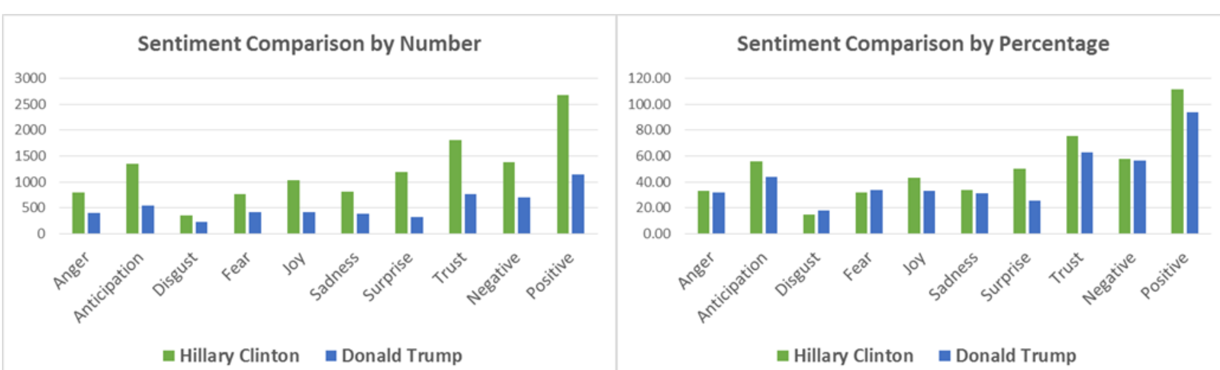

Figure 5. Sentiment analysis of posted tweets - actual numbers vs percentage comparison.

5.2 Twitter discussions surrounding the drivers of voter choice

To explain these trends, we devised a framework for analyzing the discussions surrounding the drives of voter's choice on Twitter, as illustrated in Figure 6. This model maps Twitter analytics to the drivers of voter choice.

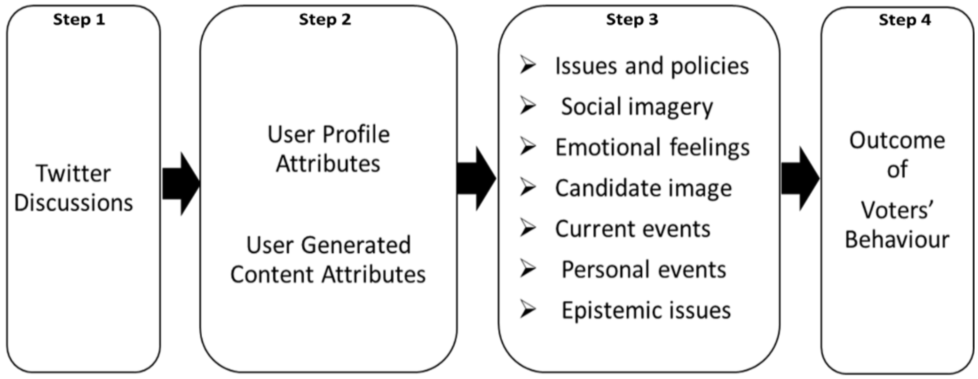

Figure 6: Proposed model for analyzing voter behaviour choice.

H4: Greater levels of social discussion – concerning the components of Newman and Sheth's model of voter's choice behaviour – increase engagement among voters, actively or passively.

Various Twitter functions, such as @mention, reply, and retweet, have been used by candidates for purposes of voter engagement (Borondo et al., 2014; Hosch-Dayican et al., 2016; Jensen, 2017). In the subsequent section, we attempt to explain our data by applying methodologies of Twitter analytics through the framework of Newman and Sheth's model of voter choice (Newman and Sheth, 1985) – detailing seven distinct cognitive domains that drive voter's behaviour.

5.2.1 Issues and policies

Issues and policies concern the economic, foreign and social policies put forward by a candidate during the election period. Key literature highlights that issues and policies are important components in influencing voter's behaviour (Newman and Sheth, 1985).

Economic policy refers to the policies concerned with reducing the level of inflation and budget-balancing. Foreign policies include policies such as those related to defence spending. After extraction from the respective candidate's Twitter screen, tweets were classified into four categories: economy, foreign policy, social issues, and leadership. This was done using content analysis, which was performed on all tweets by both investigators independently. There were 14,508 decision points (2400 tweets from Hillary Clinton, 1227 tweets from Donald Trump and four areas of issues and policies (i.e. economy, foreign policy, social issues and leadership). The two researchers agreed on 13,293 decisions and disagreed on 1,215 decisions, with a coefficient of reliability of 91.62%. This is above the 85% threshold typically used (Kassarjian, 1977). Figure 7 illustrates the tweet counts for both presidential candidates regarding policies and issues.

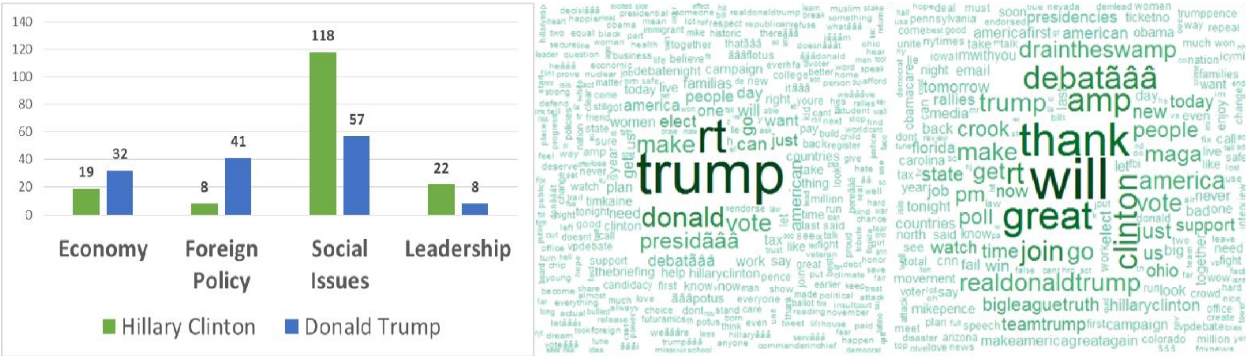

Figure 7. Issues and policies discussed by Clinton (left cloud) and Trump (right cloud).

There were 167 tweets posted by Hillary Clinton with concerning policies. Donald Trump posted only 138. Clinton discussed various social issues, specifically concerning women and children, equality, safety, empowerment, childcare leave, disability, free education, career progression, and mental stability. Clinton's tweets were focused more on social issues (and Trump's policies) whereas Trump focused more on the economy and foreign policy, such as fighting terrorism and crime, immigration, increasing job numbers and easing American business processes. Previous research has suggested that female politicians focus more on women's issues, with a communication style more directed towards attacking the opposing candidate (Evans and Clark, 2016). Our findings are consistent with this.

To investigate how people responded to these issues and policies, tweets identified as explicitly concerning policies were analyzed by aggregating the *retweet* and *like* counts of those tweets. Figure 8 shows that Trump's tweets concerning the economy, foreign policy, and broader social issues received significantly more retweets and likes than Clinton's – signifying that the Republican campaign was able to garner considerable public support in these areas.

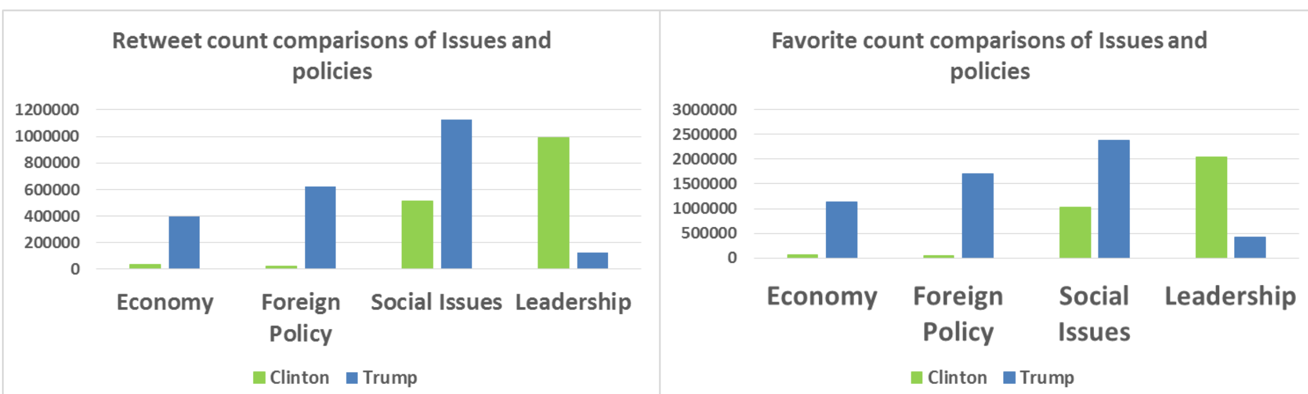

Figure 8. Comparison of the retweet count and favourite (like) count for the issues and policies tweeted by the candidates.

5.2.2 Social imagery

Social imagery refers to the perceived image of the candidate by the voter. A candidate can provoke positive and negative stereotypes of their self-image through an understanding of the socio-economic, cultural, ethical, political, and ideological dimensions of voter demographics.

Figure 9 shows the 30 most popular hashtags over the election period, through which the social images of the candidates can be inferred.

In the run-up to the election, WikiLeaks released over 30 thousand emails and email attachments from Hillary Clinton's private email server (from while she was Secretary of State) – provoking accusations of corruption. Social media discussions presenting the image of Clinton as a corrupt politician, reflected in the hashtags #podestaemails, #wikileaks, and #crookedhillary. However, #iamwithher was also one of the dominant hashtags, indicating a large amount of support for Clinton and opposition to this image.

The hashtags in green boxes reflect a positive image of Hillary Clinton, whereas hashtags in the red boxes purvey a negative image. Hashtags in the blue boxes describe a positive image of Trump; no negative imagery appears among the top 30 hashtags for Trump. The hashtag feature offered by Twitter helps candidates to reach a wider audience and allows voters to engage in the discussions surrounding a particular campaign (Jensen, 2017).

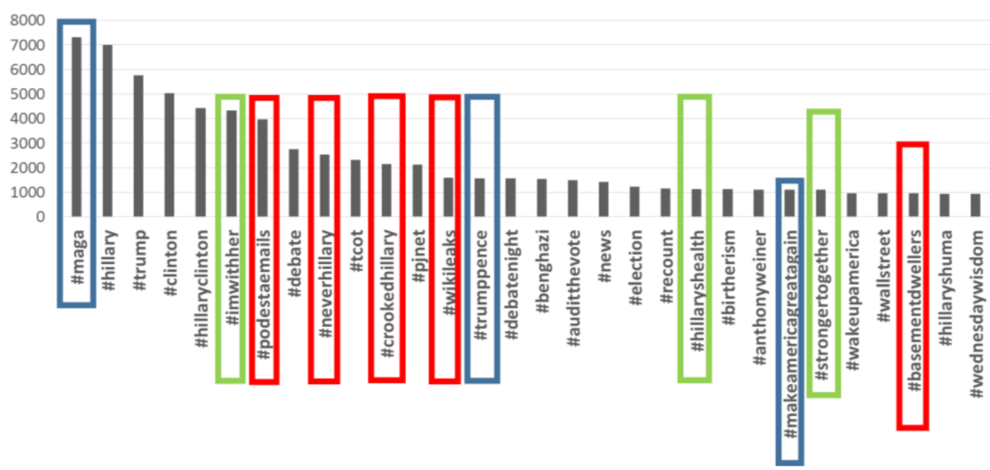

Figure 9. Nature of the imagery used to describe 2016 presidential candidates from the top 30 hashtags used in Twitter discussions.

5.2.3 Emotions

Emotions refer to the personal feelings possessed by voters towards the candidate. A comparative analysis of all discussions surrounding the two candidates was conducted using emotion analysis, as illustrated by Figure 10. The volume of these discussions concerning Clinton – for all sentiments analyzed – was greater than for those concerning Trump. This is also the case in the emotion comparison, in which tweets pertaining to emotions of *trust*, *anger*, *anticipation*, *fear*, and *disgust*, more commonly concerned Clinton. Figure 10 contains two bar charts: the left chart depicting the emotion comparison of presidential candidate's tweets by percentage and the right chart showing the emotion comparison of all tweets identified. From the graph on the left, it can be inferred that users trusted both Clinton and Trump equally, but users posted a greater number of *fear* tweets aimed towards Clinton than towards Trump. In terms of *surprise*, however, the numbers of tweets were similar for both candidates. Different emotions clearly can have different impacts; research has shown that people are more heavily influenced by emotional than cognitive discussions (Song et al., 2016).

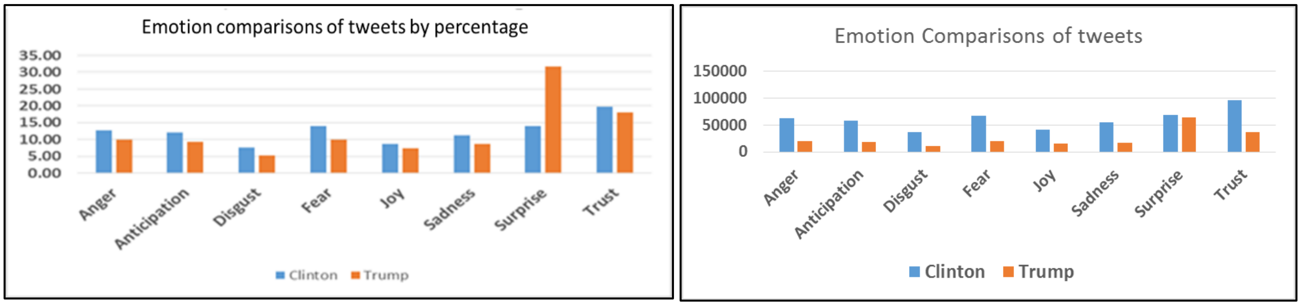

Figure 10. Emotion analysis of tweets concerning candidates Clinton and Trump.

5.2.4 Candidate image

This refers to the salient personality traits of a candidate. Voters may form an opinion the basis of *candidate image* rather than on the basis of campaign issues. As illustrated in Fig. 10, user polarity is somewhat similar in percentage of tweets but there is the difference in the number of tweets surrounding Clinton which can effect polarization of voters towards Clinton.

Figure 11 illustrates the top 30 @mention uses, along with their frequency, over the 18 weeks. Among the 784,153 tweets, there are 32,568 tweets which used the handle @realdonaldtrump (4.15%) and 20,515 tweets using @hillaryclinton (2.61%). The third most popular @mention was @wikileaks, where a lot of debate was took place concerning accusations of corruption of the Clinton campaign. This indicates that the role of WikiLeaks may have been significant in deciding the outcome of the election. Further dominant @mentions concerned news and journalism based sources (CNN, NYTimes, Reuters, FoxNews). Furthermore, the role of opinion leaders like Linda Suhler and Mike Cernovich – who vocally supported Trump – is also highlighted through the popularity of their Twitter handles in the @mention analysis. Prior research has suggested that out-of-party leaders opinions leaders have greater influence in shaping voter opinions than in-group leaders (Nicholson, 2012).

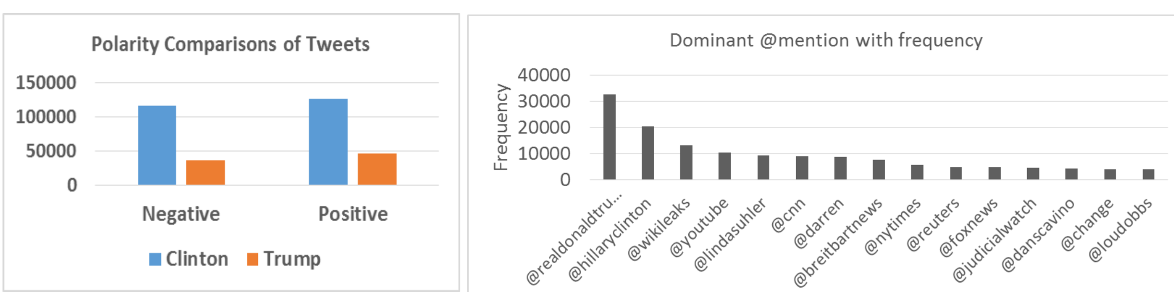

Fig. 11. Polarity analysis and top @mentions in USA election discussions.

5.2.5 Current events

This factor takes into the account the events that occurred over the course of the election, including both domestic and international events with the potential to impact individual voting behaviour. Since topic modelling is highly computationally extensive, our analysis only included days when user sentiments in Twitter fluctuated significantly (i.e. days with tweets polarity ± 2 standard deviations from the mean). This totalled 18 days and allowed construction of a word cloud to illustrate the main concerns during the election periods of enhanced user activity and major fluctuations in sentiments. For the topic modelling, the top 15 topics were identified for each of the 18 days included. Figure 12 illustrates the word cloud created, based on the popularity of 15 topics across 18 days each, to visually represent the hierarchy of topics discussed. Trump had 17.6 million followers on Twitter – producing 34,160 tweets – whereas Hillary Clinton had 11.7 million followers with 9,838 tweets. From this, it can be said that Donald Trump had greater reach than Hillary Clinton. However, Figure 12 indicates that Twitter users were more frequently discussing Clinton. WikiLeaks again appeared to have played a prominent role in the discussions surrounding Clinton. Despite her popularity, the election outcome final may possibly have been impacted by the nature of 'popularity' in such discussions, which may have polarized citizens. Research has shown that increased citizen activity on Twitter around a presidential candidate can be related to negative campaigning or citizen incivility (Hopp and Vargo, 2017). From the word cloud, it can be concluded that Hillary Clinton posted more and was discussed more on Twitter during those election periods that social media discussions increased significantly, potentially due to the emergence of popular news or notable incidents.

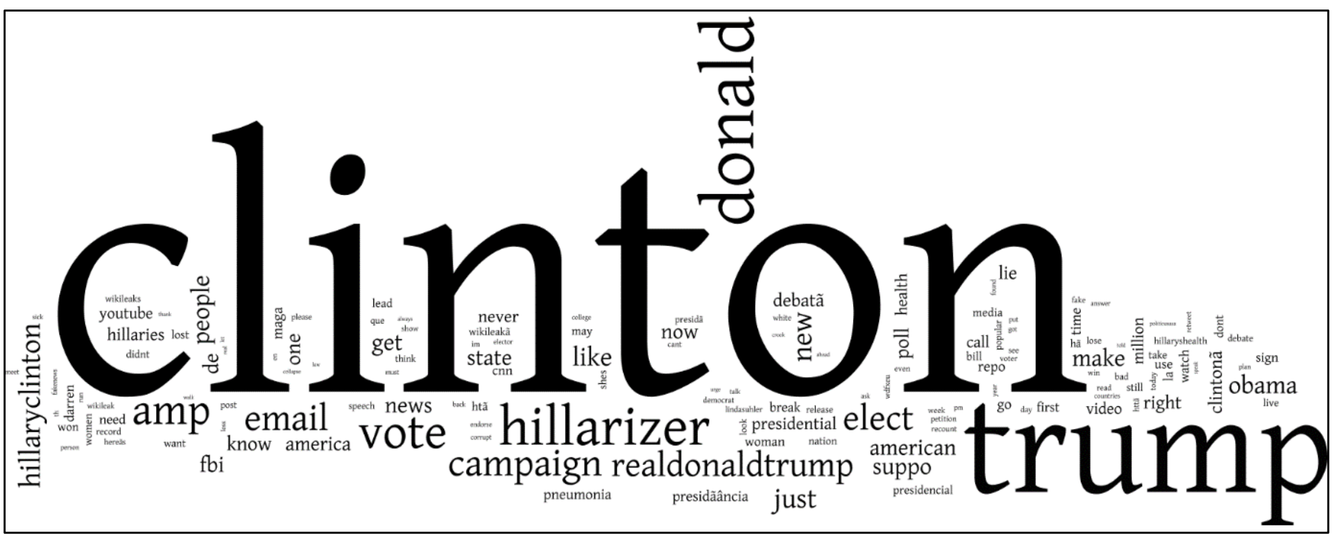

Figure 12. Word cloud on the topics identified from topic modelling of Twitter discussions surrounding the 2016 US election.

5.2.6 Personal events

This is in reference to the historical events from a presidential candidate's past with the potential to cause a voter to change their voting preference. Personal events can influence the voter's decisions positively or negatively. Previous research has emphasized that social media has increased the focus of journalism on a politician’s private life (Ekman and Widholm, 2015). Numerous personal events surrounding the Clinton campaign and were discussed negatively and extensively over Twitter: her deletion of emails using BleachBit; WikiLeaks release of over 30 thousands of her private emails; the FBI releasing detailed interview notes of their investigation into Clinton’s email practices; and many more.

The fact that @WikiLeaks was the 13th most popular hashtag (shown in Fig. 9) gives an estimate of the popularity and potential importance of the Wikileaks story. Trump, in contrast, did not hold a governmental post before winning the election and, as such, did not instil the same kinds of discussions on social media. To analyze the impact of these events, the 10 URLs creating the most buzz in social media discussions were extracted each month (Annexure 1). Each month, we found that the top 10 URLs were centred around Clinton’s personal life – with a negative perspective of her image. Some of the most shared URLs include: a video link posted by Trump, detailing Clinton’s fundraising activities; a video posted by Atlantic, differentiating between Clinton and Trump in terms of ethical disposition; and links posted by WikiLeaks, containing large amount of emails & email attachments sent to and from Clinton's private email server while she was Secretary of State. These events impacted the participants of the Twitter discussions, thereby polarizing them.

5.2.7 Epistemic issues

Epistemic issues refers to the issues raised by the candidates to bring something new in the society. Literature indicates epistemic issues raise the curiosity of the voters (Newman and Sheth, 1985). Figure 9 illustrates that #maga was the most frequently used of all hashtags; an acronym of the nationalist campaign ‘Make America Great Again’. Other campaigns instigated by Donald Trump included ‘Big League Truth’ and ‘Drain The Swamp’. In contrast, #strongertogether, launched by Hillary Clinton with the stated intention of motivating citizens to unite and fight for social issues, had much lower popularity among followers. Figure 7 also illustrates Trump’s campaign received considerable social support, whereas the Clinton campaign received less support in terms of Twitter retweets and mentions.

5.2.8 Overview of presidential candidate engagement through Twitter

Following on from the previous analysis, we looked to explore those who had participated in discussions as *influencers*, and how these individuals were connected within the networks. The top 50 @mention posts were extracted from the candidates’ Twitter screens and were mapped in the @mention network in Figure 13, where the size of the node indicates the frequency of one to one communication directly to a presidential candidate of blogger, celebrities, corporates, institutes, media houses, government officials, social workers and supporters. From Figure 13, it can be derived that media personalities and houses were interacting more with the Clinton campaign using Twitter. This is in line with research that indicates that the more a politician is active on the social media, the more journalists will follow that politician (Rauchfleisch and Metag, 2016).

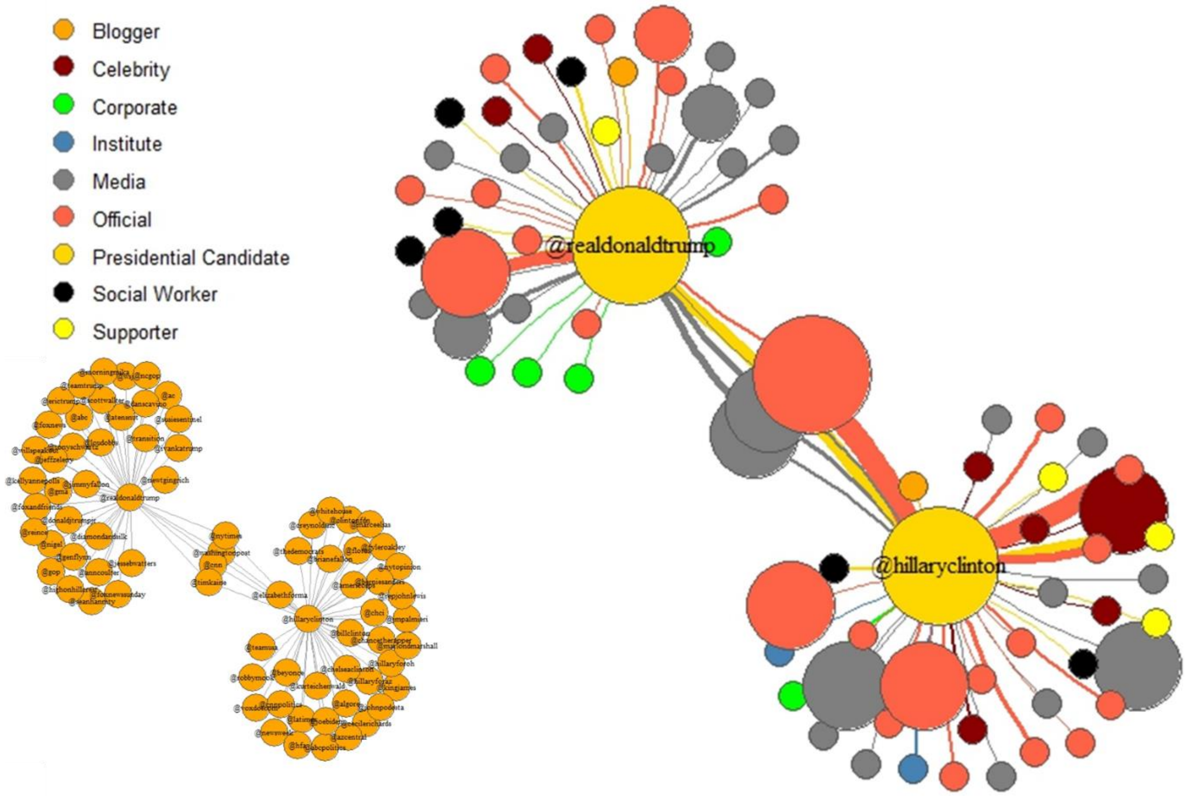

Figure 13. Top 50 @mention network for each candidate including strength of association.

5.3 Acculturation and polarization of users in the online environment

The line between social media and traditional media is becoming increasingly blurred, and social media platforms have been shown to play a significant role in shaping user cultural orientation (Li and Tsai, 2015). Therefore, we hypothesize that hashtag campaigns run on the Twitter have the ability to connect users in different geographical locations and to initiate a process of acculturation among users.

H5: Popular hashtags or campaigns initiate a process of acculturation of ideologies among Twitter users located in different geographical locations.

To explore this, all tweets posted in English (754,109) were extracted. Only 412,767 tweets contained the location of the authors. From these tweets, state names were extracted through content analysis. The final number of tweets included in the analysis was 148,881; posted by 26,386 users. The geographical distribution of the tweets (in red), users (in green), and tweet per user (in blue) is shown in Figure 14. In terms of the volume of tweets surrounding the top 5 hashtag campaigns, the highest contributing states are Tennessee (15815), Arkansas (14359) and Georgia (13283). All these states had a Republican majority in the 2016 election, potentially indicating what impact the popularity of the #MAGA campaign may have had on the outcome of the election.

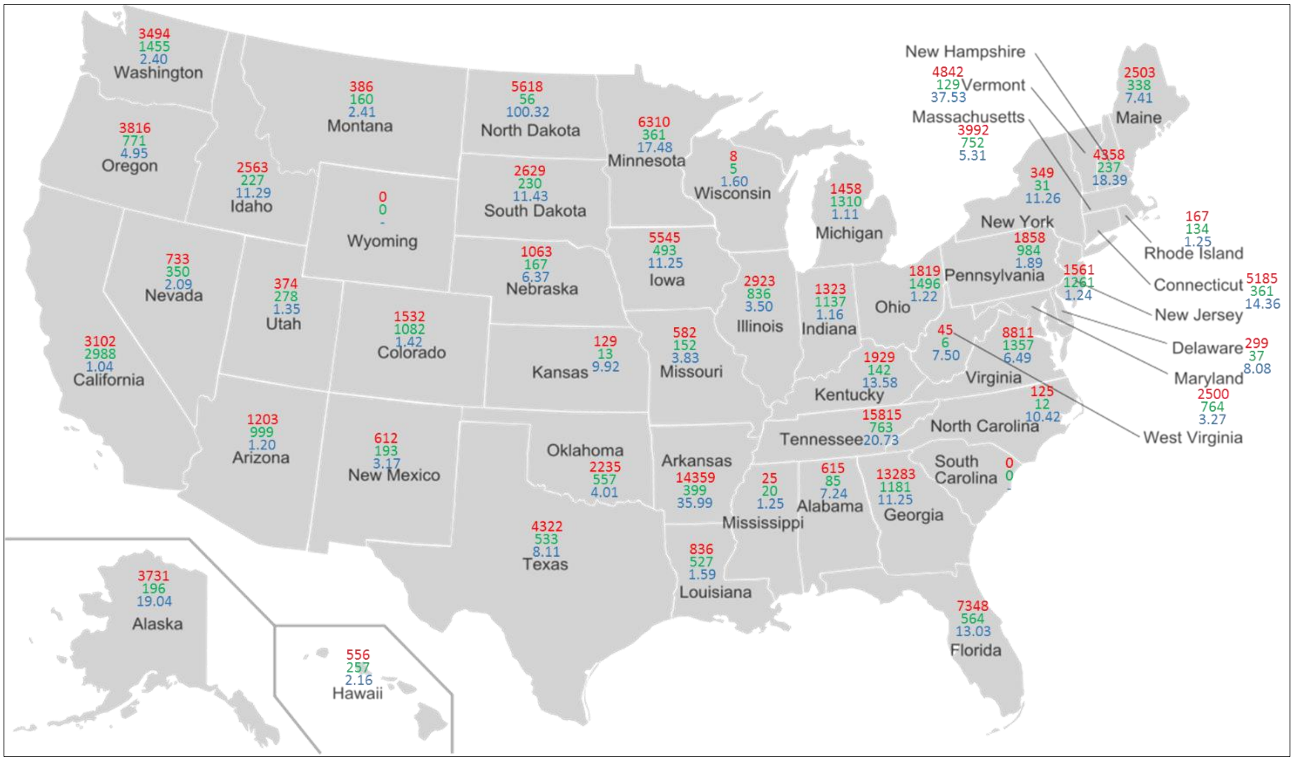

Figure 14. Geographical distribution of tweets of users in reference to the 'USA Election' over the election period.

Figure 15 illustrates the use of the five most popular hashtag campaigns across the states. The highest number uses in our sample occurred in Texas and California; whereas the states Delaware, South Dakota and West Virginia did not contribute to the top five hashtags. 28.7% of the total instances captured for the use of #maga came from the states of Texas (422) and California (328). In California and Texas, Clinton and Trump won respectively; therefore the direct impact of the top hashtag campaigns appears inconclusive.

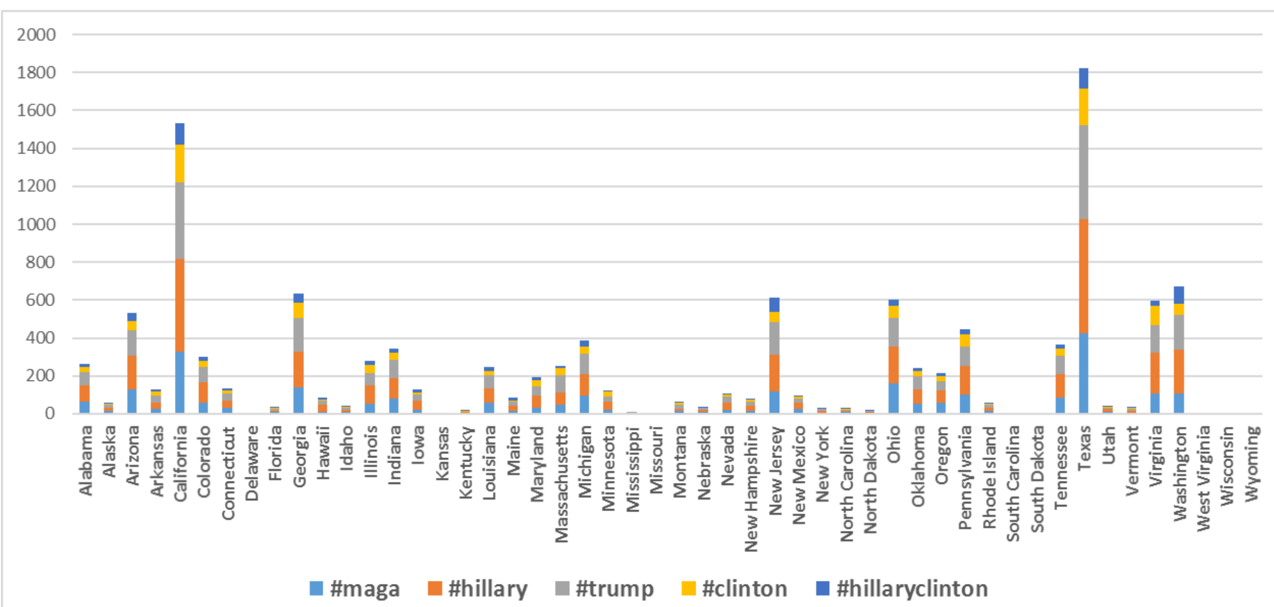

Figure 15. Usage of popular hashtags by geographical location.

Figure 16 shows the distribution of tweets containing the five most popular hashtag campaigns during the 2016 election. Figure 16 illustrates how users from disparate locations can connect through the use of hashtags on Twitter. Therefore, Figures 15 and 16 provide evidence that these campaigns can lead to political integration through the acculturation of ideologies via social media.

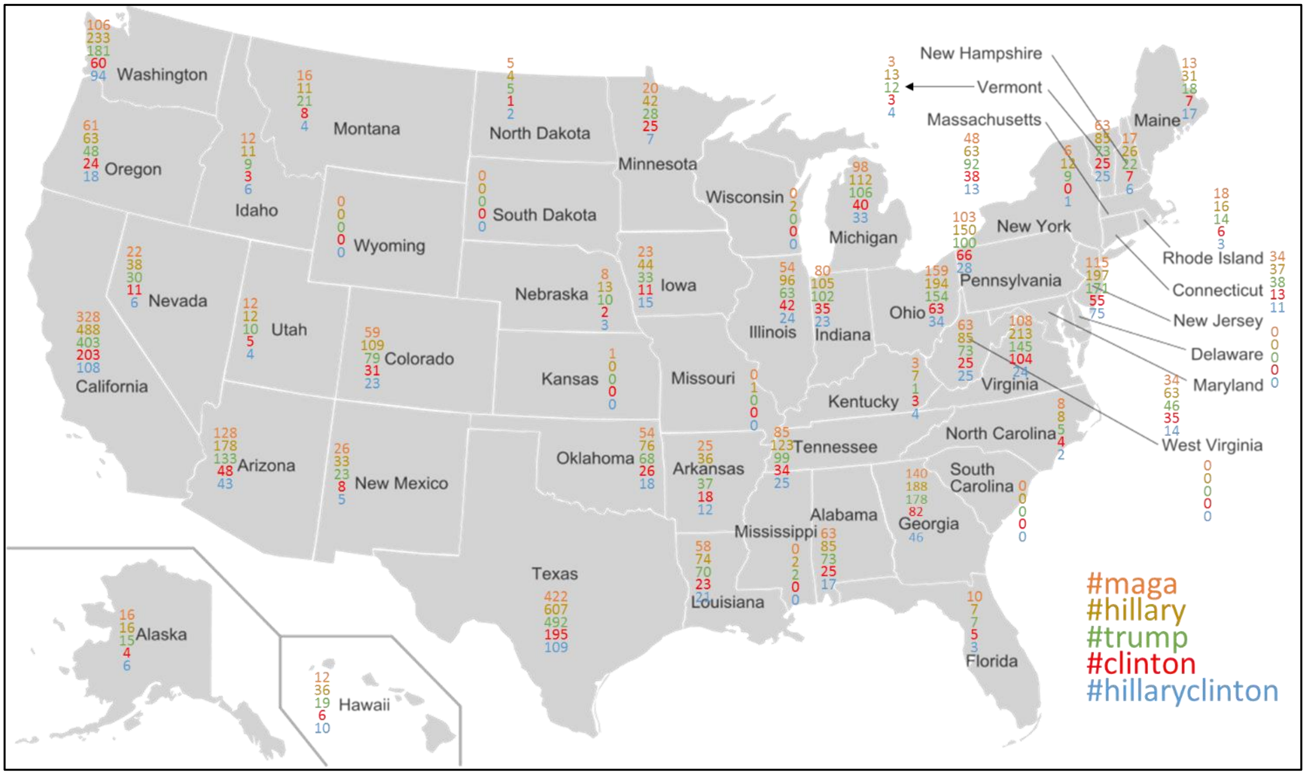

Figure 16. Top 5 hashtag usage by geographical location.

We also attempted to assess whether voter's had undergone polarization in terms of their preferred candidate. In order to address this, the election period was divided into the two phases. For both phases, tweets were categorized into those concerning Clinton or Trump. Sentiment analysis was applied to tweets to identify the polarity of the tweet with respect to that candidate (positive or negative). By comparing the early phase to the late phase, transitions in polarity could be identified. From this, users can be segregated into four groups: (1) users who are positive in the first phase for a candidate and changed their sentiment towards the candidate to negative in the second phase; (2) users who were negative in the first phase and became positive in the second phase; (3) users who were positive in the first phase and remained positive in the second phase; and (4) the users who were negative in the first phase and remained negative in the second phase with respect to the polarity of their sentiment towards the political candidate. This is illustrated in Figure 1 and is described in more detail below.

H6: Political deliberation on social media platform (Twitter) leads to opinion polarization among users.

To test this hypothesis investigate and answer sub-part of research question 3,

What is the nature of this polarization? Do voters undergo transition and polarization of their preferences through Twitter over the course of an election?

The following methodology was adopted:

Step 1: The dataset of tweets collected was divided into two phases of 60 days. Phase 1 was from 13 August - 11 October 2016, and phase 2 was from 12 October - 10 December 2016.

Step 2: For both phases, tweets were separated into those concerning Hillary Clinton and those concerning Donald Trump.

Step 3: The sentiment analysis algorithm (Saif et al., 2013) was applied to the tweets.

Step 4: Users were labelled as 'positive' or 'negative' with respect to their sentiments regarding a candidate. Positive and negative users from Phase 1 and Phase 2 were extracted for both Hillary Clinton and Donald Trump.

Step 5: Users were grouped into one of four groups for both or Hillary Clinton and Donald Trump:

I. Phase 1, Positive Users to Phase 2, Negative Users **(Indicates polarization)**
II. Phase 1, Negative Users to Phase 2, Positive Users **(Indicates polarization)**
III. Phase 1, Positive Users to Phase 2, Positive Users **(No change)**
IV. Phase 1, Negative Users to Phase 2, Negative Users **(No change)**

Table 3 illustrates the number of users in which sentiment transition had occurred during the election period for Trump and Clinton respectively. Previous research had indicated that polarization occurs uniformly across parties (Iyengar and Westwood, 2015). However, our study indicates that higher levels of polarization occurred regarding Clinton than Trump.

Table 3. Impact assessment of polarization of preferences among voters (cells contain number of users and in brackets the number of tweets posted by users).

Highlighted cells indicate polarization from Phase 1 to Phase 2		Hillary Clinton		Donald Trump	
		Phase 2		**Phase 2**	
		Positive	**Negative**	**Positive**	**Negative**
Phase 1	**Positive**	11236 (155640)	10250 (145814)	476 (15185)	309 (3528)
	Negative	10944 (154006)	10243 (147233)	485 (14768)	361 (11057)

H7: Communities are formed among groups of users polarized during social media discussions, around political events such as elections.

Hypotheses 6 and 7 – as well as research question 3 – require the segregation of the user sample into the four groups described above. We further looked to investigate how the top 15 hashtags collected from Twitter were being used by these four groups. Bode et al. (2015) suggested that network clustering has occurred on the basis of the hashtag usage. To look deeper into this concept, we explored how the top 15 hashtags identified in Fig. 8 been used by the four groups described in Table 3; and whether these groups are forming communities with the help of the hashtags. For this, users from Table 3 who had used any of the top 15 hashtags were identified. The number of users in each group is given in Table 4.

Table 4. Polarized and non-polarized users who had used the top 15 hashtags.

Highlighted cells indicate polarization from Phase 1 to Phase 2		Hillary Clinton		Donald Trump	
		Phase 2		**Phase 2**	
		Positive	**Negative**	**Positive**	**Negative**
Phase 1	**Positive**	883	301	267	47
	Negative	4576	1143	98	51

A network graph was plotted showing the usage of the top 15 hashtags, in which each user and hashtag is a node. A user is represented as a circle. The node colour describes the user on the basis of polarization: a green node represents a user who has undergone transition from negative in the first phase to positive in the second phase; a red node represents a user who has undergone a transition from positive in the first phase to negative in the second phase; and a yellow node represents a user who has not undergone any transition. The hashtag is represented as a square node, and the size of the square indicates the frequency of the hashtag use. If the user had used the hashtag, then they fall within the edges of the square. A hashtag usage graph has been drawn for both the presidential candidate's individually (Figure 17). Figure 17 describes that more people were polarized negatively concerning Clinton than Trump, as indicated by the red dots. However, positive polarization was also higher for Clinton in comparison to Trump.

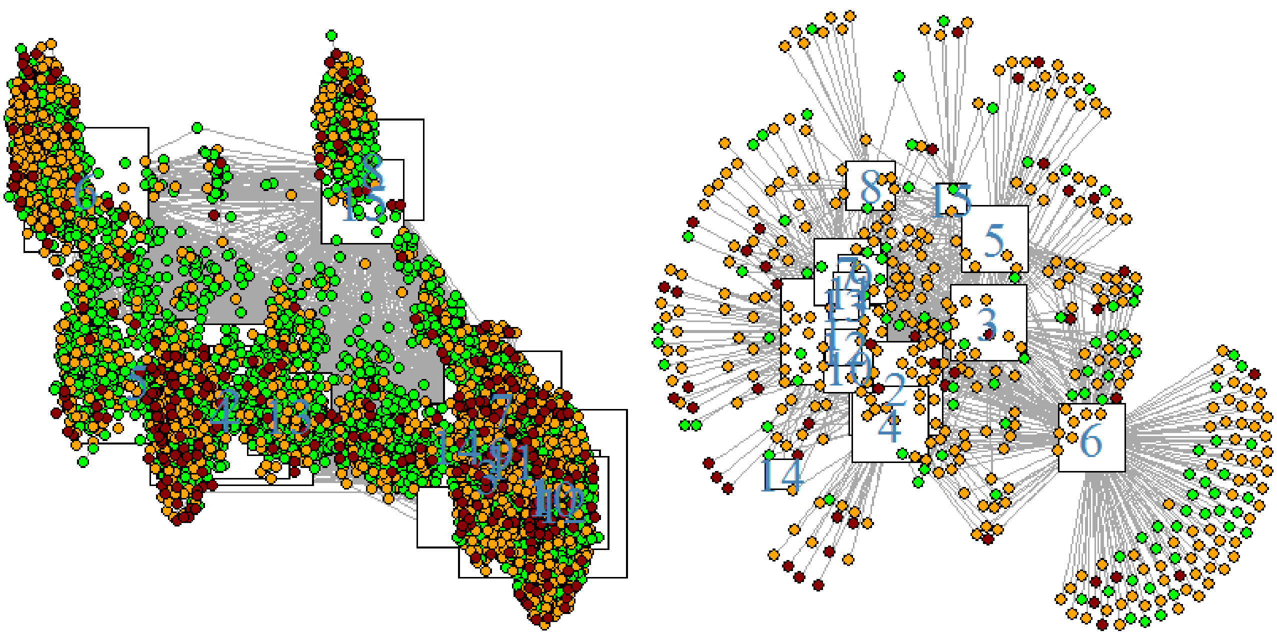

Figure 17. (a) Hashtag usage graph of the users concerning Clinton; (b) Hashtag usage graph of the users concerning Trump. Hashtag Mapping: 1-#maga; 2-#hillary; 3-#trump; 4-#clinton; 5-#hillaryclinton; 6-#imwithher; 7-#podestaemails; 8-#debate; 9-#neverhillary; 10-#tcot; 11-#crookedhillary; 12-#pjnet; 13-#wikileaks; 14-#trumppence; 15-#debatenight.

Using the data depicted in Figure 17, a greedy algorithm of modularity optimization (Fortunato, 2010) was applied to detect communities on the basis of hashtag usage. The communities detected are illustrated in Figure 18 which shows a much higher degree of overlap for Trump campaigns compared to Clinton. From Figure 18, it may be inferred that the users were forming communities on Twitter through the hashtags. With respect to Clinton, the user groups were more disparate and isolated, as depicted in the visualisation of network analysis. In comparison, Twitter users who were discussing Trump exhibited greater synergy among discussed topics and greater participation in discussions surrounding the issues and campaigns highlighted by Trump.

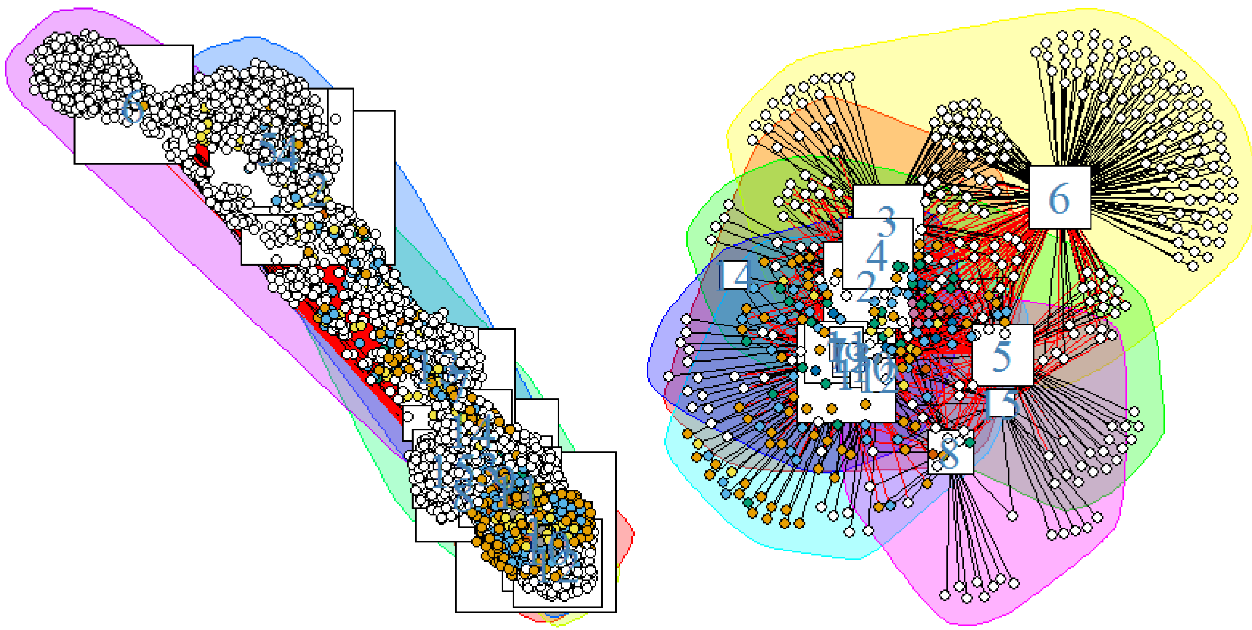

Figure 18. Community detection based on greedy optimization of modularity for Clinton (left) and Trump (right).

6 Discussion

Researchers have used data gathered from surveys, traditional news articles, and now (increasingly) social media for analyzing national events, including elections (DiMaggio et al., 1996; Newman and Sheth, 1985). As data-capture processes differ, the analytical methods applied to data must also differ. Data collected through surveys are typically examined through traditional statistical analyses such as regression, structural equation modelling, ANOVA and many more. The data collected through news articles are often analyzed through methods like exploratory content analysis. The data collected through social media is can be analyzed through social media analytics based on machine learning approaches (e.g. Grover et al., 2018; Kar, 2016; Rathore et al., 2017; Stieglitz and Dang-Xuan, 2013), which can be sub specified to *Twitter analytics*. The study presents a brief overview of Twitter analytical methodology in section 4. The data for this study was extracted from Twitter and analyzed through the use of Twitter analytics and data mining. Data collection in social media has fewer limitations concerning the size of data that can be collected; a restriction typically faced by survey-based research. However, new challenges in the analysis of such large data sets.

This study examines the possible reasons for polarization of voters through Twitter during the US 2016 election. It allows us to identify the popular hashtags, @mentions and the Twitter domains potentially influencing voter's behaviour (section 5.2). High frequency of social media activity can result in increased popularity of a presidential candidate (LaMarre and Suzuki-Lambrecht, 2013; Safiullah et al., 2017); however, in the case of Clinton, it has led to reduced or negative popularity and high levels of criticism and negative media attention (shown in Figure 13). Other studies have also described this phenomenon (Rauchfleisch and Metag, 2016).

Trump was able to maintain a synergy between social media platforms and traditional media outlet and acted as an influencer on Twitter, with campaigns like 'Make America Great Again' and 'Drain The Swamp'; the benefit of which has been previously described (Conway et al., 2015; Karlsen and Enjolras, 2016). The topics of tweets are of high importance during the election period (Figure 8). Research has shown that if the topics being discussed by a presidential election candidate are *liked*, by Twitter users, message promotion is accelerated (Zhang et al., 2016). This was true for the Trump campaign, as depicted in Table 2. The results show that *out-group* leaders such as Linda Suhler and Mike Cernovich played an important role in shaping Trump's public image; Nicholson (2012) having previously described that *out-party* leaders can exert a greater influence on voter opinion in comparison to *in-group* leaders.

Newman and Sheth, in 1985, proposed seven domains that drive voter behaviour. Through this study, we showed that the Twitter discussions concerning these seven domains might have played a significant role in the election outcome through initiating deliberation among geographically dispersed voters. The issues and policies raised by Clinton and Trump (Figure 7) initiated deliberation on Twitter among voters, as illustrated in Figure 14 and Figure 16. The social imagery of the presidential candidates was reflected in the hashtags used by voters (Figure 9). The emotional feelings of Twitter users were analyzed by applying sentiment analysis to social media buzz. In order to examine candidate image, the polarity of the social media buzz along with @mention use was analyzed. Finally, the epistemic issues raised by presidential candidates were identified and analyzed using their popular campaigns, including 'MAGA', 'Big League Truth', 'Drain the Swamp' and 'StrongerTogether'. Our study extends the existing literature regarding these domains of voter behaviour and how manipulation of them through social media may impact the choices of voters.

This study indicates that campaigns on Twitter had been used: (a) by political candidates for spreading information; (b) for influencing voter's political views through acculturation of ideologies among voters, subsequently leading to voter polarization; and (c) for engaging and associating with voters. Through the use of hashtag analysis, @mentions, and word cloud creation, it appears that Clinton's campaigns failed to gain popularity, whereas Trump's campaigns gathered significant support. Surprisingly, Clinton also tweeted more about her Republican rival, in contrast to Trump who focused mainly on his policies and their potential outcomes.

Despite Clinton having much higher visibility, the outcome of the election was affected by the nature of this visibility, and voter resonance with the content of her messages.. Twitter users were to share policies discussed by Trump (Figure 8). However, our analysis highlights that the election outcome may have been strongly polarized by the way the Twitter handles been used by presidential candidates. The number of polarized users for Clinton is higher than that for Trump. This may have been due to the high frequency of tweets by Clinton or the large social media buzz (on Twitter) around Clinton, or a combination of both. Research has previously described polarization as being something uniform across parties (Iyengar and Westwood, 2015), but our study challenges this and shows that outcomes of polarization may be different between parties, and higher engagement leads to a higher number of polarized users. This opens up a research question that can be investigated in future studies.

From the network analyses in Figure 17 and Figure 18, it can be concluded that usage of hashtags had promoted users to forming communities; an observation in keeping with the theory of homophily (Borondo et al., 2014; Himelboim et al., 2016). Polarized users are have been shown previously to form communities among themselves through hashtags (Hollander, 2008; Kim, 2015). Among the top 15 hashtags used over the election period, users with a negatively polarized view of Clinton used the hashtags #podestaemails, #tcot and #pjnet, positively polarized users towards Clinton used the hashtags #hillaryclinton and #imwithher, and non-polarized users used the hashtags #neverhillary and #crookedhillary. With respect to Trump, polarized and non-polarized users were randomly distributed across hashtag usage, and no clear interpretation regarding hashtags usage can be made from the polarized behaviour of users. This may be because of the small user group used in this analysis after filtering. This study supports the idea that Twitter is an extension of off-line interactions between candidates and voters (Miller and Ko, 2015).

6.1 Theoretical Contributions

Methodologically, this study presents a way in which user-generated data (tweets) can be collected from Twitter; and how insights can be derived through the application of Twitter analytics and data mining approaches such as regression analysis and community detection. We present an extensive list of Twitter analytics (descriptive analytics, content analysis, network analysis and geospatial analysis) which can be used to derive insights from the tweets. These methods adopted highlight how the approaches of big data analytics can be applied to social media data to provide innovative insights into complex problem domains.

The findings in our study contribute to the literature surrounding how social ecosystems use social media for conversing on topics across geographically diverse areas. Higher and more consistent frequency of social media activity by a candidate leads to higher popularity and engagement among followers but also higher levels of criticism of the candidate. Consecutive campaigns on social media engender higher popularity and engagement among Twitter users. The study also describes how including strong emotional elements (like surprise) in a tweet can increase the social buzz on social media platforms. Furthermore, greater coverage of the factors described by Newman and Sheth – issues and policies, social imagery, emotional feelings, candidate image, current events, personal events, and epistemic issues – creates more connections with otherwise geographically segregated social communities. Trump's campaign showed more substantial coverage of these factors of voter's choice behaviour compared to Clinton's, which may have impacted the outcome of the election. The study reveals that popular campaigns during the US election connected disparate groups of users on social media and facilitate acculturation of ideologies among users; helping to explain user polarization and the formation of virtual communities on social media platforms.

Results infer in the study can be used for election campaigning and digital communication which will be beneficial in influencing the voters. Furthermore, our research demonstrates how popular frameworks such as Newman and Sheth's model of voter's choice behaviour (Newman and Sheth, 1985) and the SPIN framework (Mills, 2012) can be adopted to analyze communications in virtual communities.

6.2 Implications for practice and policy

The implications of the study for practice and policy are divided into the three sections: (6.2.1) a best practice overview for electoral candidates; (6.2.2) the characteristics of a good election campaign; and (6.2.3) strategies for polarizing voter's behaviour on social media platforms such as Twitter.

6.2.1 Overview of best practices for candidate's standing in an election (Individual level)

Research has shown that political actors are using Twitter to reach out to the public and the media (Shapiro and Hemphill, 2016; Waisbord and Amado, 2017; Vaccari and Valeriani, 2015); as Twitter is multi-directional and offers interactive communication along with message broadcast facilities (Hutchins, 2016; Kruikemeier et al., 2016; Ross and Bürger, 2014; Theocharis et al., 2016). With this in mind, we suggest four *best practices* for an electoral candidate to adopt with respect to social media: (1) The Twitter handle should be responsibly used by the main political actor of the party. The political actor should not respond to every comment made by protestors in the public forum. (2) Candidates should ensure that the wording used in the tweets does not convey negative emotions like anger or disgust. (3) The candidate should strategically handle their engagement over Twitter to act as an influencer on social media platforms. (4) Candidates should be careful with about using information concerning their personal and professional background during the election and should take precautions to contain unflattering information from their pasts. The study illustrates the damaging impact that the release of past governmental information had on the Clinton campaign. (5) Candidates should balance the use of social media platforms and traditional media. Existing literature, in addition to this study, indicates that the more a candidate is active on social media, the more media attention – particularly negative attention – the candidate receives.

6.2.2 Characteristics of good campaigns or hashtags launched during the election period (organizational level)

Campaigns on social media platform are launched through hashtags (Abascal-Mena et al., 2015). The study reveals that campaigns depicts actionable agenda of the candidates; hashtags such as #maga and #draintheswamp used in Trump's tweets led to higher campaign polarity among users, which further helped in propagating the core messages of the campaigns. The study tries to highlight some of the characteristics of successful digital campaigns, firstly a digital campaign should be relevant to a large population emotionally. Secondly, should be capable of holding the voter's attention. Thirdly, a digital campaign should demonstrate their long-term benefits or values to voters.

6.2.3 Strategies for polarizing the voter's behaviour on social media platforms

Political actors have used Twitter for engaging voters (Graham et al., 2013; Purohit et al., 2013; Raynauld and Greenberg, 2014). The connections among users on Twitter can be visually depicted using networks (HerdaĞdelen et al., 2013; Stieglitz and Dang-Xuan, 2013). When political parties design their agendas for elections, two key points should be considered. First, before devising strategies, the party should investigate the issues and policies voters are currently most concerned with. Our study highlighted the concerns of US voters regarding security issues; Trump tweeted more with respect to foreign policy and security issues than Clinton, which increased engagement among voters with his campaigns. Second, campaigns launched during the election period should ensure that they improve the social image of the candidate and the organization among voters.

7 Conclusion

The study supports the notion that social media discussions have the ability to impact the outcome of national elections. This study contributes to the fields of computer-mediated communication and digital politics by shedding light on four key areas. (1) Candidate activity on Twitter – with respect to campaigning, sentiments expressed, and issues and policies discussed during the election period – has been mapped according to voter reaction and responses through: (2) acculturation of ideologies among geographically dispersed voters engaged using hashtags; (3) opinion polarization among voters; and (4) formation of communities. These four areas are depicted in Figure 1.

The study allows us to better understand the dynamics of polarization in the online environment by converting qualitative tweets into quantified data using machine learning algorithms, content analysis, and network analysis. Various factors influencing voter behaviour are highlighted in this study. The study also highlights that social media now plays an important role in the success of election campaigns, as it can facilitate voter engagement, public scrutiny, public harassment and polarize voting outcome. Table 5 summarizes the findings of the study.

Table 5. Summary of findings.

S.No	Hypothesis	Outcome / Result
1	Higher activity on social media is positively related to higher popularity and engagement among followers.	Negative feedback may also increase with higher engagement (as in the case of Clinton's Twitter activity).
2	Less variation in time (greater nexus) between consecutive campaigns is positively related to higher popularity and engagement.	Yes, positively: From the sample collected it seem Trump had less time between consecutive campaigns which may had led to greater engagement and popularity.
3	Higher thresholds of sentiments (polarity) within tweets is positively related to higher popularity and engagement among	Partially. There was very little difference in the percentage of emotional tweets posted between Trump and Clinton except in the case of the 'surprise' emotion.

	followers.	
4	Greater levels of social discussion – concerning the components of Newman and Sheth's model of voter's choice behaviour – increase engagement among voters, actively or passively.	Yes, positively. Greater coverage of all seven factors in campaigns indicated a positive outcome with higher engagement.
5	Popular hashtags or campaigns initiate a process of acculturation of ideologies among Twitter users located in different geographical locations.	Yes. The #maga campaign gained support from citizens across the USA.
6	Political deliberation on social media platforms (Twitter) leads to opinion polarization among users.	Yes. The number of users transitioning from a negative to a positive opinion of a candidate over the election period is higher than for those transitioning from a positive to a negative opinion.
7	Communities are formed among groups of users polarized during social media discussions, around political events such as elections.	Yes. Using hashtag analysis, it is evident that communities are formed around campaigns, which are often overlapping.

This study broadens the literature surrounding social media by presenting how community formation and polarization of voting outcome is feasible based on acculturation of ideologies through social media platforms. This study contributes to various research avenues such the role of influencers in information propagation over a network, the social psychology of online users, best practices in computer-mediated communication, acculturation of ideologies, user polarization and social media usage.

8. Limitations and future work

This study extracted the data set from Twitter, which allows a daily extraction of 4000 to 10000 records. This restriction on the extraction of tweets poses a limitation for this type of study. It is possible that we were unable to track all important events happening on Twitter. The second potential limitation of the study is that, of course, Twitter users may be influenced by other, external events as opposed to solely those related to Twitter discussions. These cannot be mapped or factored into our analyses concerning polarizations in user preferences. Similarly, other popular social media platforms like Facebook have not been considered in this study, due to challenges in accessing such data as well as integration challenges between data sets. A third limitation of the study is that for our analysis of hashtag clustering's of users, we limited our investigations to the top 15 hashtags. If a Twitter user is unaware of a hashtag in popular use, they may not be able to contribute to the discussions concerning that theme. Fourth, most of the analyses involved in social media analytics are based on visualization to draw inferences, future researchers may use statistical test for validating the hypothesis. Lastly, the study cannot track whether tweets had been posted by a human or a bot. Also, we do not attempt to differentiate between tweets made by candidates and those made by a social media marketing company on behalf of the candidate. However, future research could seek to address these limitations and build upon the scope of the study. The limitations highlighted in this study may be explored as future research directions for improving the current theoretical understanding of voting behaviour through social media.

References

Abascal-Mena, R., Lema, R., & Sèdes, F. (2015). Detecting sociosemantic communities by applying social network analysis in tweets. *Social Network Analysis and Mining*, 5(1), 1-17.

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? The *Journal of Politics*, 70(02), 542–555.

Adams, A., and McCorkindale, T. 2013. Dialogue and transparency: A content analysis of how the 2012 presidential candidates used twitter. *Public Relations Review*, 39(4), 357-359.

Ahmed, S., Jaidka, K., & Cho, J. (2016). The 2014 Indian elections on Twitter: A comparison of campaign strategies of political parties. *Telematics and Informatics*, 33(4), 1071-1087.

AlAlwan, A., Rana, N.P., Dwivedi, Y.K., & Algharabat, R. (2017). Social Media in Marketing: A Review and Analysis of the Existing Literature. *Telematics and Informatics,* 34(7), 1177-1190.

Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, 337(6092), 337-341.

Aswani, R., Kar, A.K., Ilavarasan, P.V., & Dwivedi, Y.K. (2018). Search Engine Marketing is not all gold: Insights from Twitter and SEOClerks. *International Journal of Information Management,* 38(1), 107–116.

Aswani, R., Ghrera, S.P., Kar, A.K., & Chandra, S. (2017). Identifying buzz in social media: a hybrid approach using artificial bee colony and k-nearest neighbors for outlier detection. *Social Network Analysis and Mining*, 7(1), 38:1-10.

Aswani, A., Kar, A. K., Aggarwal, S., & Ilavarsan, P.V. (2017, November). Exploring Content Virality in Facebook: A Semantic Based Approach. In *Conference on e-Business, e-Services and e-Society* (pp. 209-220). Springer, Cham.

Attu, R., & Terras, M. (2017). What people study when they study Tumblr: Classifying Tumblr-related academic research. *Journal of Documentation*, 73(3), 528-554.

Barberá, P., Wang, N., Bonneau, R., Jost, J. T., Nagler, J., Tucker, J., & González-Bailón, S. (2015). The critical periphery in the growth of social protests. *PloS one*, 10(11), e0143611.

Barnett, G. A., Ruiz, J. B., Xu, W. W., Park, J. Y., & Park, H. W. (2017). The world is not flat: Evaluating the inequality in global information gatekeeping through website co-mentions. *Technological Forecasting and Social Change*, 117, 38-45.

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, 22(7), 891-893.

Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of Marketing Research*, 49(2), 192-205.

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied psychology*, 46(1), 5-34.

Berry, J. W. (2008). Globalisation and acculturation. *International Journal of Intercultural Relations*, 32(4), 328-336.

Bode, L. (2016). Political news in the news feed: Learning politics from social media. *Mass Communication and Society*, 19(1), 24-48.

Bode, L., Hanna, A., Yang, J., & Shah, D. V. (2015). Candidate networks, citizen clusters, and political expression: Strategic hashtag use in the 2010 midterms. The *ANNALS of the American Academy of Political and Social Science*, 659(1), 149-165.

Borondo, J., Morales, A. J., Benito, R. M., & Losada, J. C. (2014). Mapping the online communication patterns of political conversations. *Physica A: Statistical Mechanics and its Applications*, 414, 403-413.

Boynton, G. R., & Richardson Jr, G. W. (2016). Agenda setting in the twenty-first century. *New media & Society*, 18(9), 1916-1934.

Bruns, A., & Stieglitz, S. (2013). Towards more systematic Twitter analysis: metrics for tweeting activities. *International Journal of Social Research Methodology*, 16(2), 91-108.

Burnap, P., Gibson, R., Sloan, L., Southern, R., and Williams, M. 2016. 140 characters to victory? Using Twitter to predict the UK 2015 General Election. *Electoral Studies*, 41, 230-233.

Burnap, P., Rana, O. F., Avis, N., Williams, M., Housley, W., Edwards, A., and Sloan, L. 2015. Detecting tension in online communities with computational Twitter analysis. *Technological Forecasting and Social Change*, 95, 96-108.

Cao, L., & Zhang, T. (2012). Social networking sites and educational adaptation in higher education: A case study of Chinese international students in New Zealand. *The Scientific World Journal*, 2012.

Ceron, A., Curini, L., Iacus, S. M., and Porro, G. 2014. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, 16(2), 340-358.

Chadwick, A., O'Loughlin, B., & Vaccari, C. (2017). Why people dual screen political debates and why it matters for democratic engagement. *Journal of Broadcasting & Electronic Media*, 61(2), 220-239.

Chae, B. K. (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.

Chatfield, A. T., Scholl, H. J. J., and Brajawidagda, U. 2013. Tsunami early warnings via Twitter in government: Net-savvy citizens' coproduction of time-critical public information services. *Government Information Quarterly*, 30(4), 377-386.

Cleveland, M., Laroche, M., Pons, F., & Kastoun, R. (2009). Acculturation and consumption: Textures of cultural adaptation. *International Journal of Intercultural Relations*, 33(3), 196-212.

Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., and Danforth, C. M. 2015. Climate change sentiment on twitter: an unsolicited public opinion poll. *PloS one*, 10(8), e0136092.

Conway, B. A., Kenski, K., & Wang, D. (2015). The rise of Twitter in the political campaign: Searching for intermedia agenda-setting effects in the presidential primary. *Journal of Computer-Mediated Communication*, 20(4), 363-380.

Croucher, S. M. (2011). Social networking and cultural adaptation: A theoretical model. *Journal of International and Intercultural Communication*, 4(4), 259-264.

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have American's social attitudes become more polarized? *American Journal of Sociology*, 102(3), 690–755.

Dimitrova, D. V., Shehata, A., Strömbäck, J., & Nord, L. W. (2014). The effects of digital media on political knowledge and participation in election campaigns: Evidence from panel data. *Communication Research,* 41(1), 95-118.

Djerf-Pierre, M., & Pierre, J. (2016). Mediatised local government: social media activity and media strategies among local government officials 1989–2010. *Policy & Politics*, 44(1), 59-77.

Domingo, J., & Martos, J. M. (2015). Analysis of political discourse in Spain on school failure on Twitter. *Analytical Archives of Educational Policies*, 24 (70). Http://dx.doi.org/10.14507/epaa.24. 2357

Dwivedi, Y.K., Kapoor, K.K. & Chen, H. (2015). Social Media Marketing and Advertising. *The Marketing Review,* 15(3), 289-309.

Ekman, M., & Widholm, A. (2015). Politicians as Media Producers: Current trajectories in the relation between journalists and politicians in the age of social media. *Journalism Practice*, 9(1), 78-91.

Ems, L. (2014). Twitter's place in the tussle: how old power struggles play out on a new stage. *Media, Culture & Society*, 36(5), 720-731.

Engesser, S., Ernst, N., Esser, F., & Büchel, F. (2017). Populism and social media: How politicians spread a fragmented ideology. *Information, Communication & Society*, 20(8), 1109-1126.

Enli, G. S., & Skogerbø, E. (2013). Personalized campaigns in party-centred politics: Twitter and Facebook as arenas for political communication. *Information, Communication & Society*, 16(5), 757-774.

Ernst, N., Engesser, S., Büchel, F., Blassnig, S., & Esser, F. (2017). Extreme parties and populism: an analysis of Facebook and Twitter across six countries. *Information, Communication & Society*, 1-18.

Evans, H. K., & Clark, J. H. (2016). "You Tweet Like a Girl!" How Female Candidates Campaign on Twitter. *American Politics Research*, 44(2), 326-352.

Ferguson, Y. L., Ferguson, K. T., & Ferguson, G. M. (2017). I am AmeriBritSouthAfrican-Zambian: Multidimensional remote acculturation and well-being among urban Zambian adolescents. *International Journal of Psychology*, 52(1), 67-76.

Forbush, E., & Foucault-Welles, B. (2016). Social media use and adaptation among Chinese students beginning to study in the United States. *International Journal of Intercultural Relations*, 50, 1-12.

Fortunato, S. (2010). Community detection in graphs. *Physics reports*, 486(3), 75-174.

Frame, A., & Brachotte, G. (2015). Le tweet stratégique: Use of Twitter as a PR tool by French politicians. *Public Relations Review*, 41(2), 278-287.

Ganis, M., and Kohirkar, A. 2015. Social Media Analytics: Techniques and Insights for Extracting Business Value Out of Social Media. *IBM Press*.

Gonzalez-Bailon, S., Wang, N., Rivero, A., Borge-Holthoefer, J., & Moreno, Y. (2014). Assessing the bias in samples of large online networks. *Social Networks*, 38, 16-27.

Graham, T., Broersma, M., Hazelhoff, K., & van't Haar, G. (2013). Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign. *Information, Communication & Society*, 16(5), 692-716.

Grover, P., Kar, A. K., Dwivedi, Y.K., & Janssen, M. (2017, November). The untold story of USA presidential elections in 2016 - Insights from Twitter Analytics. In *Conference on e-Business, e-Services and e-Society* (pp. 339-350). Springer, Cham.

Grover, P., & Kar, A. K. (2017). Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature. *Global Journal of Flexible Systems Management*, 1-27. Doi:10.1007/s40171-017-0159-3

Grover, P., Kar, A. K., & Davies, G. (2018). "Technology enabled Health"–Insights from twitter analytics with a socio-technical perspective. *International Journal of Information Management*, 43, 85-97.

Gruzd, A., & Roy, J. (2014). Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet*, 6(1), 28-45.

Gupta, S., Kar, A.K., Baabdullah, A., Al-Khowaiter, WAA. (2018). Big data with cognitive computing: A review for the future. *International Journal of Information Management*, 42, 78-89.

Harris, J. K., Moreland-Russell, S., Choucair, B., Mansour, R., Staub, M., and Simmons, K. 2014. Tweeting for and against public health policy: response to the Chicago Department of Public Health's electronic cigarette Twitter campaign. *Journal of Medical Internet Research*, 16(10), e238.

Heller Baird, C., and Parasnis, G. 2011. From social media to social customer relationship management. *Strategy & Leadership*, 39(5), 30-37.

Henderson, A., & Bowley, R. (2010). Authentic dialogue? The role of "friendship" in a so-cial media recruitment campaign. *Journal of Communication Management*, 14(3), 237-257.

Heo, Y. C., Park, J. Y., Kim, J. Y., & Park, H. W. (2016). The emerging viewertariat in South Korea: The Seoul mayoral TV debate on Twitter, Facebook, and blogs. *Telematics and Informatics*, 33(2), 570-583.

HerdaĞdelen, A., Zuo, W., Gard-Murray, A., & Bar-Yam, Y. (2013). An exploration of social identity: The geography and politics of news-sharing communities in twitter. *Complexity*, 19(2), 10-20.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, 18(2), 40-60.

Himelboim, I., Sweetser, K. D., Tinkham, S. F., Cameron, K., Danelo, M., & West, K. (2016). Valence-based homophily on Twitter: Network analysis of emotions and political talk in the 2012 presidential election. *New Media & Society*, 18(7), 1382-1400.

Hindriks, P., Verkuyten, M., & Coenders, M. (2016). Evaluating Political Acculturation Strategies: The Perspective of the Majority and Other Minority Groups. *Political Psychology*. doi: 10.1111/pops.12356.

Hollander, B. A. (2008). Tuning out or tuning elsewhere? Partisanship, polarization, and media migration from 1998 to 2006. *Journalism & Mass Communication Quarterly*, 85(1), 23–40.

Hong, S., & Nadler, D. (2012). Which candidates do the public discuss online in an election campaign?: The use of social media by 2012 presidential candidates and its impact on candidate salience. *Government Information Quarterly*, 29(4), 455-461.

Hopp, T., & Vargo, C. J. (2017). Does negative campaign advertising stimulate uncivil communication on social media? Measuring audience response using big data. *Computers in Human Behavior*, 68, 368-377.

Hosch-Dayican, B., Amrit, C., Aarts, K., & Dassen, A. (2016). How do online citizens persuade fellow voters? Using Twitter during the 2012 Dutch parliamentary election campaign. *Social Science Computer Review*, 34(2), 135-152.

Hossain, M.A., Dwivedi, Y.K., Chan, C., Standing, S., & Olanrewaju, A-S. (2018). Sharing political content in online social media: A planned and unplanned behaviour approach. *Information Systems Frontiers*, 20(3), 485–501.

Hsu, C. L., & Park, H. W. (2012). Mapping online social networks of Korean politicians. *Government Information Quarterly*, 29(2), 169-181.

Hutchins, B. (2016). The many modalities of social networking: The role of Twitter in greens politics. *Environmental Communication*, 10(1), 25-42.

Ibrahim, N. F., Wang, X., & Bourne, H. (2017). Exploring the effect of user engagement in online brand communities: Evidence from Twitter. *Computers in Human Behavior*, 72, 321-338.

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology*, 50(6), 1141–1151.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690–707.

Jensen, M. J. (2017). Social media and political campaigning: changing terms of engagement?. The *International Journal of Press/Politics*, 22(1), 23-42.

Joseph, N., Kar, A. K., Ilavarasan, P. V., & Ganesh, S. (2017). Review of discussions on internet of things (IoT): insights from twitter analytics. *Journal of Global Information Management*, 25(2), 38-51.

Jungherr, A. (2014). The logic of political coverage on Twitter: Temporal dynamics and content. *Journal of Communication*, 64(2), 239-259.

Kar, A. K. (2016). Bio inspired computing–A review of algorithms and scope of applications. *Expert Systems with Applications*, 59, 20-32.

Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P., Dwivedi, Y.K., & Nerur, S. (2018). Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, 20(3), 531–558.

Kapoor, K.K., & Dwivedi, Y.K. (2015). Metamorphosis of Indian electoral campaigns: Modi's social media experiment. *International Journal of Indian Culture & Business Management*, 11(4), 496–516.

Karlsen, R., & Enjolras, B. (2016). Styles of Social Media Campaigning and Influence in a Hybrid Political Communication System: Linking Candidate Survey Data with Twitter Data. The *International Journal of Press/Politics*, 21(3), 338-357.

Kassarjian, H. H. (1977). Content analysis in consumer research. *Journal of Consumer Research*, 4(1), 8-18.

Kayser, V., & Blind, K. (2017). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*, 116, 208-215.

Kelm, O., Dohle, M., & Bernhard, U. (2017). Social Media Activities of Political Communication Practitioners: The Impact of Strategic Orientation and In-Group Orientation. *International Journal of Strategic Communication*, 11(4), 306-323.

Kim, A. J., and Ko, E. 2012. Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand. *Journal of Business Research*, 65(10), 1480-1486.

Kim, Y. (2015). Does disagreement mitigate polarization? How selective exposure and disagreement affect political polarization. *Journalism & Mass Communication Quarterly*, 92(4), 915-937.

Klinger, U. (2013). Mastering the art of social media: Swiss parties, the 2011 national election and digital challenges. *Information, Communication & Society*, 16(5), 717-736.

Kruikemeier, S., Kruikemeier, S., van Noort, G., van Noort, G., Vliegenthart, R., Vliegenthart, R., ... & H. de Vreese, C. (2016). The relationship between online campaigning and political involvement. *Online Information Review*, 40(5), 673-694.

Lakhiwal, A. and Kar, A.K. 2016. Insights from Twitter analytics: Modeling social media personality dimensions and impact of breakthrough events. *Lecture Notes in Computer Science*, 9844, 533-544.

LaMarre, H. L., & Suzuki-Lambrecht, Y. (2013). Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'. *Public Relations Review*, 39(4), 360-368.

Larsson, A. O. (2017). Going viral? Comparing parties on social media during the 2014 Swedish election. *Convergence*, 23(2), 117-131.

Larsson, A. O., & Ihlen, Ø. (2015). Birds of a feather flock together? Party leaders on Twitter during the 2013 Norwegian elections. *European Journal of Communication*, 30(6), 666-681.

Larsson, A. O., and Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14(5), 729-747.

Lawrence, E., Sides, J., & Farrell, H. (2010). Self-segregation or deliberation? Blog readership, participation, and polarization in American politics. *Perspectives on Politics*, 8(1), 141.

Layman, G. C., Carsey, T. M., & Horowitz, J. M. (2006). Party polarization in American Politics: Characteristics, causes, and consequences. *Annual Review of Political Science*, 9(1), 83–110.

Lee, E. J. (2007). Deindividuation effects on group polarization in computer-mediated communication: The role of group identification, public-self-awareness, and perceived argument quality. *Journal of communication*, 57(2), 385-403.

Lee, J. K., Choi, J., Kim, C., & Kim, Y. (2014). Social media, network heterogeneity, and opinion polarization. *Journal of communication*, 64(4), 702-722.

Li, C., & Tsai, W. H. S. (2015). Social media usage and acculturation: A test with Hispanics in the US. *Computers in Human Behavior*, 45, 204-212.

Llewellyn, C., Grover, C., Alex, B., Oberlander, J., and Tobin, R. 2015. Extracting a topic specific dataset from a Twitter archive. In *International Conference on Theory and Practice of Digital Libraries*. Springer International Publishing (2015, September), 364-367.

Lysenko, V.V., Desouza, K.C. (2011). Moldova's internet revolution: Analyzing the role of technologies in various phases of the confrontation. *Technological Forecasting and Social Change*, doi:10.1016/j.techfore.2011.05.009.

Mao, Y., & Yuxia, Y. (2015). Facebook use and acculturation: the case of overseas Chinese professionals in Western countries. *International Journal of Communication*, 9(1), 2467-2486.

McKelvey, K., DiGrazia, J., & Rojas, F. (2014). Twitter publics: How online political communities signaled electoral outcomes in the 2010 US house election. *Information, Communication & Society*, 17(4), 436-450.

Miller, N. W., & Ko, R. S. (2015). Studying Political Microblogging: Parliamentary Candidates on Twitter During the February 2012 Election in Kuwait. *International Journal of Communication*, 9, 21.

Mills, A. J. (2012). Virality in social media: the SPIN framework. *Journal of public affairs*, 12(2), 162-169.

Mosca, L., & Quaranta, M. (2016). News diets, social media use and non-institutional participation in three communication ecologies: comparing Germany, Italy and the UK. *Information, Communication & Society*, 19(3), 325-345.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, 12(2), 125–135.

Moya-Sánchez, M., & Herrera-Damas, S. (2016). How to Measure Persuasive Potential on Twitter: A Methodological Proposal. *Palabra Clave*, 19(3), 838-867.

Neiger, B. L., Thackeray, R., Van Wagenen, S. A., Hanson, C. L., West, J. H., Barnes, M. D., and Fagen, M. C. 2012. Use of social media in health promotion purposes, key performance indicators, and evaluation metrics. *Health Promotion Practice*, 13(2), 159-164.

Newman, B. I., & Sheth, J. N. (1985). A model of primary voter behavior. *Journal of Consumer Research*, 12(2), 178-187.

Nicholson, S. P. (2012). Polarizing cues. *American Journal of Political Science*, 56(1), 52-66.

Nooralahzadeh, F., Arunachalam, V., & Chiru, C. G. (2013, May). 2012 Presidential Elections on Twitter--An Analysis of How the US and French Election were Reflected in Tweets. In Control Systems and Computer Science (CSCS), 2013 19th International Conference on (pp. 240-246). *IEEE*.

Ogden, D. T., Ogden, J. R., & Schau, H. J. (2004). Exploring the impact of culture and acculturation on consumer purchase decisions: Toward a microcultural perspective. *Academy of Marketing Science Review*, 3(1), 1-22.

Ogola, G. (2015). Social media as a heteroglossic discursive space and Kenya's emergent alternative/citizen experiment. *African Journalism Studies*, 36(4), 66-81.

Panagiotopoulos, P., Barnett, J., Bigdeli, A. Z., & Sams, S. (2016). Social media in emergency management: Twitter as a tool for communicating risks to the public. *Technological Forecasting and Social Change*, 111, 86-96.

Park, S. J., Lim, Y. S., & Park, H. W. (2015). Comparing Twitter and YouTube networks in information diffusion: The case of the “Occupy Wall Street” movement. *Technological Forecasting and Social Change*, 95, 208-217.

Poell, T. (2014). Social media and the transformation of activist communication: Exploring the social media ecology of the 2010 Toronto G20 protests. *Information, Communication & Society*, 17(6), 716-731.

Purohit, H., Hampton, A., Shalin, V. L., Sheth, A. P., Flach, J., and Bhatt, S. 2013. What kind of# conversation is Twitter? Mining# psycholinguistic cues for emergency coordination. *Computers in Human Behavior*, 29(6), 2438-2447.

Rathore, A. K., Kar, A. K., & Ilavarasan, P. V. (2017). Social Media Analytics: Literature Review and Directions for Future Research. *Decision Analysis*, 14(4), 229-249.

Rauchfleisch, A., & Metag, J. (2016). The special case of Switzerland: Swiss politicians on Twitter. *New Media & Society*, 18(10), 2413-2431.

Raynauld, V., & Greenberg, J. (2014). Tweet, click, vote: Twitter and the 2010 Ottawa municipal election. *Journal of Information Technology & Politics*, 11(4), 412-434.

Redfield, R., Linton, R., & Herskovits, M. J. (1936). Memorandum for the study of acculturation. *American anthropologist*, 38(1), 149-152.

Ross, K., & Bürger, T. (2014). Face to face (book) Social media, political campaigning and the unbearable lightness of being there. *Political Science*, 66(1), 46-62.

Rui, J. R., & Wang, H. (2015). Social network sites and international students' cross-cultural adaptation. *Computers in Human Behavior*, 49, 400-411.

Saboo, A. R., Kumar, V., & Park, I. (2016). Using Big Data to Model Time-Varying Effects for Marketing Resource (Re) Allocation. *MIS Quarterly*, 40(4), 911-939.

Safiullah, M., Pathak, P., Singh, S., & Anshul, A. (2017). Social media as an upcoming tool for political marketing effectiveness. *Asia Pacific Management Review*, 22(1), 10-15.

Scherpereel, J. A., Wohlgemuth, J., & Schmelzinger, M. (2017). The Adoption and Use of Twitter as a Representational Tool among Members of the European Parliament. *European Politics and Society*, 18(2), 111-127.

Scott, C. F., Bay-Cheng, L. Y., Prince, M. A., Nochajski, T. H., & Collins, R. L. (2017). Time spent online: Latent profile analyses of emerging adults' social media use. *Computers in Human Behavior*, 75, 311-319.

Shah, D. V., Hanna, A., Bucy, E. P., Wells, C., & Quevedo, V. (2015). The power of television images in a social media age: Linking biobehavioral and computational approaches via the second screen. The *ANNALS of the American Academy of Political and Social Science*, 659(1), 225-245.

Shapiro, M. A., & Hemphill, L. (2017). Politicians and the Policy Agenda: Does Use of Twitter by the US Congress Direct New York Times Content?. *Policy & Internet*, 9(1), 109-132.

Shuai, X., Pepe, A., & Bollen, J. 2012. How the scientific community reacts to newly sub-mitted preprints: Article downloads, twitter mentions, and citations. *PloS one*, 7(11), e47523.

Singh, J.P., Dwivedi, Y.K., Rana, N.P., Kumar, A., & Kapoor, K.K. (2017). Event classification and location prediction from tweets during disaster. *Annals of Operations Research,* DoI: https://doi.org/10.1007/s10479-017-2522-3

Skogerbø, E., & Krumsvik, A. H. (2015). Newspapers, Facebook and Twitter: Intermedial agenda setting in local election campaigns. *Journalism Practice*, 9(3), 350-366.

Social Media Fact Sheet. PEW RESEARCH CENTER, 12 Jan. 2017. Web. Retrieved on 21 July 2017 from http://www.pewinternet.org/fact-sheet/social-media/.

Song, Y., Dai, X. Y., & Wang, J. (2016). Not all emotions are created equal: Expressive behavior of the networked public on China's social media site. *Computers in Human Behavior*, 60, 525-533.

Stephens, M., & Poorthuis, A. (2015). Follow thy neighbor: Connecting the social and the spatial networks on Twitter. *Computers, Environment and Urban Systems*, 53, 87-95.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217-248.

Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining*, 3(4), 1277-1291.

Stirland, S. L. (2008). Propelled by internet, barack obama wins presidency. *Wired Magazine*, 4.

Thackeray, R., Neiger, B. L., Hanson, C. L., and McKenzie, J. F. 2008. Enhancing promo-tional strategies within social marketing programs: use of Web 2.0 social media. *Health promotion practice*, 9(4), 338-343.

Theocharis, Y. (2013). The wealth of (occupation) networks? Communication patterns and information distribution in a Twitter protest network. *Journal of Information Technology & Politics*, 10(1), 35-56.

Theocharis, Y., Barberá, P., Fazekas, Z., Popa, S. A., & Parnet, O. (2016). A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil Twitter Use When Interacting With Party Candidates. *Journal of Communication*, 66(6), 1007-1031.

Theocharis, Y., Lowe, W., van Deth, J. W., & García-Albacete, G. (2015). Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society*, 18(2), 202-220.

Vaccari, C., & Valeriani, A. (2015). Follow the leader! Direct and indirect flows of political communication during the 2013 Italian general election campaign. *New Media & Society*, 17(7), 1025-1042.

Vaccari, C., Chadwick, A., & O'Loughlin, B. (2015). Dual screening the political: Media events, social media, and citizen engagement. *Journal of Communication*, 65(6), 1041-1061.

Vaccari, C., Valeriani, A., Barberá, P., Bonneau, R., Jost, J. T., Nagler, J., & Tucker, J. A. (2015). Political Expression and Action on Social Media: Exploring the Relationship Between Lower-and Higher-Threshold Political Activities Among Twitter Users in Italy. *Journal of Computer-Mediated Communication*, 20(2), 221-239.

Van Kessel, S., & Castelein, R. (2016). Shifting the blame. Populist politicians’ use of Twitter as a tool of opposition. *Journal of Contemporary European Research*, 12 (2), pp. 594-614.

Volkova, S., & Bachrach, Y. (2015). On predicting sociodemographic traits and emotions from communications in social networks and their implications to online self-disclosure. *Cyberpsychology, Behavior, and Social Networking*, 18(12), 726-736.

Vromen, A., Xenos, M. A., & Loader, B. (2015). Young people, social media and connective action: From organisational maintenance to everyday political talk. *Journal of Youth Studies*, 18(1), 80-100.

Waisbord, S., & Amado, A. (2017). Populist communication by digital means: presidential Twitter in Latin America. *Information, Communication & Society*, 20(9), 1330-1346.

Wu, A. X. (2014). Ideological polarization over a China-as-superpower mindset: An exploratory charting of belief systems among Chinese Internet users, 2008-2011. *International Journal of Communication*, 2014(8), 2650–2679.

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, 30(5), 316-327.

Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-Scale Network Analysis for Online Social Brand Advertising. *MIS Quarterly*, 40(4), 849-868.

Zhang, Y., Moe, W. W., & Schweidel, D. A. (2017). Modeling the role of message content and influencers in social media rebroadcasting. *International Journal of Research in Marketing*, 34(1), 100-119.

Zhu, D. H. (2013). Group polarization on corporate boards: Theory and evidence on board decisions about acquisition premiums. *Strategic Management Journal*, 34, 800–822.

Annexure

Top URL across the month along with their descriptions (Annexure)				
August				
Rank	**URL**	**Description**	**Count**	**Polarity towards Hillary Clinton**
1	https://t.co/D0MeBJXBwN	Hillary Clinton Deleted Emails using BleachBit which intended to prevent recovery of deleted emails	259	Negative
2	https://t.co/ubS4OTxGbg	According to Marine Le Pen, leader of the National Front in France ” For France, anything is better than Clinton”. Clinton will bring "war," "devastation" and "instability" as the president.	248	Negative
3	https://t.co/CQTSo2ETJF	According to USA, WTFM Hillary Clinton as an insider threat because she had sent classified information using her personal server.	229	Negative
4	https://t.co/MEcH3u2uT2	Expose Hillary	228	Negative
5	https://t.co/b2hFO1RlIQ	Huma Abedin, Hillary Clinton's top aide, was assistant editor of an Islamic journal published an article accusing Jews of 'working the American political system'.	201	Negative
6	https://t.co/MJQp0rcnzH	Hillary Clinton needs to address the racist undertones of her 2008 campaign.	200	Negative
7	https://t.co/fFpvl62RMB	Election promotion	191	-
8	https://t.co/XJBZ59Rzb2	Hillary Clinton had claimed that Mexico’s corruption and scandal-plagued President Enrique Peña Nieto is America’s friend	189	Negative
9	https://t.co/hNfvE9Bau4	Dr. Ben Carson reaction on granting special “access” and “favors” to Clinton Foundation donors by Hillary Clinton during her State Department tenure.	171	Negative

10	https://t.co/uewPloyyoH	According to The New York Post, Clinton continued to email classified information even after she resigned as Secretary of State in 2013. According to Raj Shah because of this Hillary Clinton can't be trusted for nation's security.	167	Negative
September				
Rank	**URL**	**Description**	**Count**	**Polarity towards Hillary Clinton**
1	https://t.co/QZ8BpcZk2l	WikiLeaks – Releasing the information regarding the governance of Hillary Clinton	587	-
2	https://t.co/9dreUeDhZ9	WikiLeaks	587	-
3	https://t.co/YcjQUb83qr	Steph Curry being asked Hillary or Trump? Curry responded: "Hillary" Steph Curry is a basketball player of the National Basketball Association.	368	Positive
4	https://t.co/sBHOHU5dYn	Steph Curry Chooses Hillary Clinton Over Donald Trump For President	368	Positive
5	https://t.co/c1zs5DStuN	Hillary Clinton career flashback	257	Positive
6	https://t.co/vznTnFelwu	National Poll results: Donald Trump and Hillary Clinton essentially going to tied over presidential election	255	-
7	https://t.co/tOg4KIAvVA	New Batch of Hillary Clinton Emails showing Clinton Foundation contacts to cope with crises facing the U.S. government overseas.	254	Negative
8	https://t.co/oCVHoPvNHM	FBI had released detailed interview notes of investigation of Hillary Clinton's email practices.	240	Negative
9	https://t.co/BIZvlAPHew	Clinton was facing criticism of not holding a news conference for the months but had able to raise the $50 million from 22 fund-raising events, averaging around $150,000 an hour.	215	Negative
10	https://t.co/so5MCo2TVK	According to Clinton, America should treat cyber attacks like any other attack	210	Negative
October				
Rank	**URL**	**Description**	**Count**	**Polarity towards Hillary**

				Clinton
1	https://t.co/uKh5sCFfrv	The video posted by Donald Trump on Twitter showcasing the activities done by Hillary Clinton to raise the fund	1131	Negative
2	https://t.co/bUUkzgOA2E	TowsonU is a manager for the best DJ in Maryland and tweeted that he will not vote for Hillary Clinton	990	Negative
3	https://t.co/9ZcbSAmD0j	The article by Atlantic, differentiating between the Hillary Clinton and Donald Trump at the end of the article saying Trump is unfit for the office and declaring him as a demagogue, a xenophobe, a sexist, a know-nothing, and a liar person.	933	Positive
4	https://t.co/S7tPrl2QCZ	Wikileaks	712	Negative
5	https://t.co/lcG6u02Kgv	The Atlantic posted video supporting Hillary Clinton and pointing out bad things against Donald Trump	588	Positive
6	https://t.co/qy2EQBa48y	Wikileaks	556	Negative
7	https://t.co/b5HqsGrc7N	Flashback on Hillary Clinton decisions and their results is failure when it comes to national security and international relations	497	Negative
8	https://t.co/3cBNYjl5CD	Wikileaks had thrown the lights on the money raised by Hillary Clinton by leaking the emails.	482	Negative
9	https://t.co/0aHB7pV7u3	Wikileaks	443	Negative
10	https://t.co/QKOqtwFgwM	Wikileaks	401	Negative
November				
Rank	**URL**	**Description**	**Count**	**Polarity towards Hillary Clinton**
1	https://t.co/86uLziQXC4	A Thanksgiving message from President-elect Donald J. Trump.	1471	-
2	https://t.co/ZTh5cuY26Z	Justification for nominating Tom Price as Chairman of the House Budget Committee Congressman	1102	-
3	https://t.co/VvtB0z3L0G	Video posted on Twitter saying not to make fun of Hillary Clinton in front of the females	382	Positive
4	https://t.co/d7ueOJvlvT	Clinton leading	305	Positive
5	https://t.co/qcaDTsF8c7	Choice for Secretary of State	293	-
6	https://t.co/mDMYLSrGTn	Tweet by Twitter handle @America_1st_ saying voting for Hilliary Clinton is like supporting crime	281	Negative

7	https://t.co/tvPFZ73o30	Clinton leading	273	Positive
8	https://t.co/kUKaLrlQzw	Clinton leading	273	Positive
9	https://t.co/6NAY9dm5G1	Policy plans for First one hundred days	272	-
10	https://t.co/VbisTkUE3A	Clinton has won popular vote with substantial margin	266	Positive
December				
Rank	**URL**	**Description**		
1	https://t.co/puZVWYs9b4	TIME's Person of the Year for 2016	507	-
2	https://t.co/bzCbt0iaXD	Clinton Ignored the Working Class	281	Negative
3	https://t.co/MRUAYv1DkE	Electoral College petition to make Hillary Clinton as a President.	270	Positive
4	https://t.co/Mcc74kwzKa	Thank you, tour 2016 Cincinnati, Ohio	247	