# Analysis of political discourse on twitter in the context of the 2016 US presidential elections

Ussama Yaqub[a], Soon Ae Chun[b,*], Vijayalakshmi Atluri[a], Jaideep Vaidya[a]

[a] Rutgers Business School, 1 Washington Park, Newark, NJ 07102, USA
[b] City University of New York, Staten Island, NY 10314, USA

## ABSTRACT

Social media now plays a pivotal role in electoral campaigns. Rapid dissemination of information through platforms such as Twitter has enabled politicians to broadcast their message to a wide audience. In this paper, we investigated the sentiment of tweets by the two main presidential candidates, Hillary Clinton and Donald Trump, along with almost 2.9 million tweets by Twitter users during the 2016 US Presidential Elections. We analyzed these short texts to evaluate how accurately Twitter represented the public opinion and real world events of significance related with the elections. We also analyzed the behavior of over a million distinct Twitter users to identify whether the platform was used to share original opinions and to interact with other users or whether few opinions were repeated over and over again with little inter-user dialogue. Finally, we wanted to assess the sentiment of tweets by both candidates and their impact on the election related discourse on Twitter. Some of our findings included the discovery that little original content was created by users and Twitter was primarily used for rebroadcasting already present opinions in the form of retweets with little communication between users. Also of significance was the finding that sentiment and topics expressed on Twitter can be a good proxy of public opinion and important election related events. Moreover, we found that Donald Trump offered a more optimistic and positive campaign message than Hillary Clinton and enjoyed better sentiment when mentioned in messages by Twitter users.

## 1. Introduction

Online social media networks, such as Facebook and Twitter, have enabled people to not only use the platform for interaction with one another but also to read and share news, discuss important events and engage in political discussions. Additionally, proliferation of smart phones has further facilitated the use of this medium, allowing citizens to communicate without any limitation on time or location.

This rich medium of communication, presented by social media, has been recognized by politicians and political parties globally (Romero, Meeder, & Kleinberg, 2011). The potential of social media in political campaigns was first highlighted during the US Presidential elections of 2008. Twitter played an important part in the campaign of Barack Obama. The Obama campaign made effective use of Twitter to post campaign updates along with informing followers of opportunities to volunteer (Baumgartner et al., 2010). In 17 months starting from April 2007 to Election Day November 5th 2008, the Obama campaign posted 262 twitter messages and gained approximately 118,000 new followers (Glassman, Straus, & Shogan, 2009). In light of this successful Twitter

campaign, all major candidates and political parties now have some form of presence on social media.

Social media also allows government institutions to have candid communication with their citizens (Lorenzi et al., 2014), potentially increasing openness and transparency into the working of their organizations (Bertot, Jaeger, & Grimes, 2010). Studies have shown that from civic services to police departments, information sharing and public engagement through Twitter can lead to greater transparency and more confidence of citizens on their state and local institutions (Heverin & Zach, 2010).

As of 1st quarter of 2017 Twitter has an average of 328 million monthly active users (Statista, 2017) providing political actors with a massive user base to share their message quickly and cheaply without going through the traditional media briefings and news conferences (Romero et al., 2011). Thus it is not surprising that in the recently concluded US Presidential Elections of 2016, Twitter played a very important role in the dissemination of information regarding various policy points for both serious presidential contenders, Hillary Clinton and Donald Trump. Both candidates had millions of followers on

Twitter and their tweets closely monitored by public and mainstream media. Although it is hard to quantify the role Twitter played in the 2016 elections, majority agree that it was significant. This means that political players cannot ignore the role of social media as a communication channel to not only share their own political agenda but also as a real-time two-way channel to continuously monitor and measure public reactions. Overall, social media presents an exciting avenue of opportunity for politicians, campaigners and political activists to not only broadcast their message but also to engage in dialogue with proponents of competing political ideas and ideologies.

This increase in political discourse on Twitter has also led to an increase in research of Twitter analytics in terms of election prediction and candidate popularity (Tumasjan, Sprenger, Sandner, & Welpe, 2010; Yaqub, Chun, Atluri, & Vaidya, 2017). Some studies have even suggested that sentiment analysis of tweets can potentially be used as a substitute for traditional polls monitoring consumer confidence and political approval ratings (O'Connor, Balasubramanyan, Routledge, & Smith, 2010).

In this paper, we investigate citizen participation in the political discourse that took place on Twitter during the US Presidential Elections of 2016 by analyzing the citizens' sentiment and behavior. Our goal is to test 1) if the sentiments, and contents expressed in the political discourse on Twitter are indicative of the citizen opinion and the topics reflect the real-life events of importance; 2) gauge the sentiments of each candidate's campaign messages on Twitter and assess their impact on the sentiment of the overall election related discussion on the platform; and 3) if the social media communication behaviors promotes citizen to citizen interactions with abundant exchange of original ideas and opinions.

To test these assumptions we employed data analytics approach. We collected over 3.1 million tweets for 21 days consecutively, starting from 29th of October, up until 18th of November 2016, downloading 150,000 tweets per day on average. This data was then used to analyze citizen behavior and sentiment during this period. Furthermore we also analyzed the tweets posted by Hillary Clinton and Donald Trump, both of whom used Twitter actively for electioneering. All tweets made by both candidates starting from 29th of October until Election Day (8th of November), were considered for this study.

Specifically, we present the following analyses of the Twitter data set:

- Topic and sentiment analysis of Twitter dataset: The aim here was to detect if there existed a significant correlation between the sentiment and topics discussed on Twitter with the actual citizen opinion and real world events and breaking news of the period. This relationship between sentiment and popular trends on Twitter with real-life events suggested that citizen tweets can be used as a good predictor of the importance of certain topics and of public opinion regarding the elections and the two presidential candidates.

- Sentiment and impact analysis of tweets by the two presidential candidates: The purpose of this analysis was to utilize candidate tweets in-order to evaluate sentiment of messages posted by both Hillary Clinton and Donald Trump in the last days of their campaign. We wanted to assess the characteristics of message propagated by each candidate along with evaluation of the impact of candidate tweets on the sentiment of discussions taking place on Twitter during this period.

- Analysis of social media users' behavior: The social media usage behavior analysis aims to identify how actively Twitter users were using the online forum to speak their mind and engage with one another. Usage behavior analysis gauged whether there was diversity of opinions and open interaction between citizens, or were few opinions repeated over and over again as retweets with little one-to-one interaction.

The rest of the paper is structured as follows. In Section 2, we review the related work in the literature while in Section 3 we discuss our methodology and data. In Section 4, we develop hypotheses that are tested through data analysis. Section 5 presents the results of our data analysis and hypothesis testing. In Section 6 we discuss these results while in Section 7, we conclude our paper.

## 2. Literature review

The role of Internet and communication technologies (ICT) in modern society cannot be understated. Individuals and institutions around the world are trying to increase public engagement by utilizing Web 2.0 (Bertot et al., 2010; Lorenzi et al., 2014). This provides a quick and cost effective platform to political actors and state institutions to communicate quickly and directly with public (Heverin & Zach, 2010). For example, Twitter is now being used by city governments to benefit their populations by raising information awareness in a simple, low cost fashion. The idea is to enhance the responsiveness of different branches of local governments that deal primarily in performing tasks on behalf of the citizens and interacting with them (Lorenzi et al., 2014).

Along with governance, Twitter sentiment analysis is also used in vast array of areas related with governance and public trust ranging from predicting resentment against government policies to predicting general election results (Calderon et al., 2015; Tumasjan et al., 2010). Various models have been developed that try to understand the user behavior and retweeting on Twitter (Broersma & Graham, 2012). The emerging field of techno-social systems aims to comprehend and predict this behavior. Although this area of study is still evolving and generating a lot of enthusiasm, nonetheless, a debate on the efficacy of using Twitter sentiment analysis to predict elections and other real world events still continues (Avello, Daniel, & Mustafaraj, 2011; Metaxas, Mustafaraj, & Gayo-Avello, 2011). Important questions such as how representative Twitter users are of general population remain to be answered. These issues become acute when these analyses are conducted on data obtained from developing countries where a relatively small percentage of population has access to internet.

Another aspect is the varying levels of citizen activity. Some users are far more active online than others and thus have a greater 'weight' to their opinions. There also exists much noise on Twitter in the form of automated activity and spam, which exploit trending topics to advertise various unrelated products or content. Different solutions have been proposed to differentiate between human activity and that generated by bots (Chu, Gianvecchio, Wang, & Jajodia, 2012).

Other studies have looked at how the information spreads on social networks and what role sentiment plays in diffusion (Ferrara & Yang, 2015b). Most agree that sentiment does play an important role in information diffusion on Twitter. Some have gone as far as saying that there exists a positivity bias in information spread and that positive tweets are retweeted more and reach a wider audience than negative tweets (Ferrara & Yang, 2015a; Ferrara & Yang, 2015b).

With regards to the use of Twitter in politics, researchers have examined the ways in which Twitter influences communication of mainstream news and journalism. Recent research shows that social media in general and Twitter in particular are playing an important role for mainstream media as a news source. This can be in form of a quote or policy issue outlined through Twitter messages by politicians or other political actors such as news commentators and observers (Parmelee, 2013). Twitter is now ever more used as news agenda building tool for mainstream media (Jungherr, 2014; Wallsten, 2014). This was a very commonly observed phenomenon during the recently concluded US elections of 2016.

Utilization of Twitter by politicians and their campaigns is a popular subject of study. Usage of Twitter during the campaign cycle of 2008 in USA by Barack Obama generated interest in understanding Twitter's role in political campaigns (Abroms & Craig Lefebvre, 2009; Baumgartner et al., 2010). Similar research was also conducted in analyzing Twitter activity of US Congress members during their