The Text and Images of the GOPDebate: What the Public is "Talking" About on Instagram

Sikana Tanupabrungsun¹, Jeff Hemsley¹
¹School of Information Studies, Syracuse University

Abstract
The richness of Instagram data makes it possible to tell a compelling story about the public’s “talk” on Instagram. Our work focuses on the use of Instagram by citizens to express their thoughts on the 2016 GOP presidential election. We collected Instagram posts with the hashtags #GOPDebate and #GOPDebates for five months. Using topic-modeling and sentiment analysis techniques to analyze both textual (post captions) and visual (images) attributes we are able to illustrate the topical network of political discussions and actors whom the public talks about on each topic. Our work contributes to literature on the role of social media, and specifically Instagram, in the political domain. The methodology also demonstrates how textual and visual attributes can be used together to categorize photo content.

Keywords: Instagram; politicians; topic modeling; image processing
doi: 10.9776/16526
Copyright: Copyright is held by the authors.
Contact: stanupab@syr.edu, jjhemsle@syr.edu

1 Introduction
With the increased usage of social network sites is an analogous increase in the digitization of and transformation of social practices. For example, politicians’ use of social media is both similar to the ways they have used other media (Golbeck, Grimes, & Rogers, 2010; Hemphill, Otterbacher, & Shapiro, 2013) and different (Graham, Broersma, Hazelhoff, & Haar, 2013; Grant, Moon, & Grant, 2010). Certainly, public figures whose careers heavily depend on public’s impressions, such as pop celebrities, media personalities and politicians utilize social media to stay in the public eye without relying on, or having their messages mediated by, the main stream media. As such, social network sites become an important source of information providing insights into not only politicians’ strategies but also the opinions of those in the greater public.

While most of the existing scholarly works have looked primarily at either Facebook or Twitter, looking at either politicians (Bruns & Highfield, 2013; Golbeck et al., 2010) or citizens (Hoang, Lim, Achananuparp, Jiang, & Zhu, 2011; Lee, Ahn, Oh, & Ryu, 2015; Stiglitz & Dang-Xuan, 2012), little is known about Instagram. Instagram, however, is the third most popular social network site and is the fastest growing image-centric platform. Since its launch in 2010, Instagram has grown to over 400 million active users who post an average of 80 million photos per day (Instagram, 2015). According to a Pew study, the proportion of Americans using Instagram rapidly jumped from 13% in 2012 to 26% in 2014, while Twitter is used by only 23% of Americans. The demographics of Instagram users also make it compelling to study online politics especially when considering the “Obama coalition”. This coalition is made up of women, young people and minorities. The majority of Instagram users are women; user’s age skews young with the majority adults aged between 18 and 29; and, finally, racial minorities are more likely to use Instagram than whites: for example, 38% of African Americans and 34% of Latinos versus only 21% of whites are on Instagram (Pew Research Center, 2015). If the three legs of this coalition are any indication like they were in 2008 and 2012, Instagram can potentially be an indicator of the election’s outcome.

Instagram is a mobile application for editing and sharing photos. Similar to Twitter, Instagram users have a network of followers whom they can share a photo or 15-second video with. Users can put a post caption limited to 2,200 characters with an option to @mention others to get their attention and/or embed hashtags to participate in public conversations. As Pew suggests, photos and videos have become the key social currencies online (Rainie, Brenner, & Purcell, 2015) Instagram data is a potentially rich source of information for studying the political sphere.

In this work, we are interested in the use of Instagram by the public in a political context. Our work looks at the public’s “talk” on Instagram about the 2016 GOP presidential debates by summarizing

¹ http://techcrunch.com/2014/01/21/instagram-is-the-fastest-growing-social-site-globally-mobile-devices-rule-over-pcs-for-social-access/
the hot topics as well as topical sentiments. We collected Instagram posts with the hashtags #GOPDebate and #GOPDebates. Since May 10th, 2015, when we started collecting data with our two hashtags, we have 18,361 Instagram posts.

In particular, our research questions are:

What are the “hot topics” on Instagram and who are the “hot actors” on each topic, related to the 2016 GOP presidential election debates?

What are public’s opinions on the debates in general and, specifically, on each topic?

We answer the questions with mixed-methods approach starting with a qualitative content analysis on both textual and visual attributes. To do this we transform images to descriptive text: e.g., the name of person, text or objects in images. The transformation can be done either manually or with image processing techniques such as Optical Character Recognition (OCR) and pattern recognition. Then, we use topic-modeling techniques to cluster post captions and image descriptions by topics or themes of conversation (RQ1). We analyze the sentiments on post captions and image description to summarize the sentiment of the public on the debates as a whole and organized by topics (RQ2).

2 Proposed Methodology

Topic modeling is a clustering technique to group similar documents together in order to discover the abstract topics of each group. Sentiment analysis identifies and extracts subjective information from data, such as the positive or negative sentiment of the message. Many works have performed sentiment analysis on text data in a variety of situations: e.g. product reviews and movie reviews (Cui, Mittal, & Datar, 2006; Naveed, Gottron, Kunegis, & Alhadi, 2011; Pang, Lee, & Vaithyanathan, 2002). Both topic modeling and sentiment mining on social network data are certainly not new but most studies are conducted on textual information alone (Hu, John, Seligmann, & Wang, 2012; Hu, Wang, & Kambhampati, 2013; Pak & Paroubek, 2010; Yajuan, Zhum, & Furu, 2012; Yang et al., 2014). However, not many works actually use visual data and, as far as we know, none make use of both textual and visual data. Wang, Wang, Tang, Liu, & Li, (2015) discuss that the main challenge of image-centric sentiment analysis is the semantic gap between visual features, e.g. color histogram, brightness and the presence of objects, and higher-level image sentiments. In other words, visual attributes do not provide sufficient semantic meanings. The lack of annotated social media images is another challenge. To overcome the challenges, they developed an Unsupervised SENTiment Analysis (USEA) framework. Their method makes use of both textual attributes, e.g. captions, emoticons and comments, and visual attributes. The model was tested on Flickr and Instagram posts, the results show that their model has successfully overcome the challenges and is capable of predicting sentiment of posts.

We will apply OCR and Face recognition to transform images to descriptive texts. The results from an analysis on both caption and descriptive texts of Instagram posts will promote a better understanding of the public’s political opinions toward the 2016 GOP presidential election candidates during the debates and the role of Instagram in the political domain. We will present the results with topical network plots to not only illustrate the topical network of political discussions but also to show the “hot users” whom the public talks about on each topical discussion on Instagram.

The main essence of this work is around the examination of Instagram posts as a gauge of public’s opinions towards the 2016 GOP presidential election. The findings can be used along with or instead of traditional polls for the segments of the populations noted above. Electoral campaigns will also benefit from the findings, as they are direct feedbacks from the public. In addition, the findings will provide the understanding about the role of social media in the political sphere and contribute to the online political communication literature. The methodology also makes a meaningful contribution in terms of demonstrating how visual and textual attributes can be used together in content analysis approaches.

3 References


