

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; Stieglitz & 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 platform¹. 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

- Bruns, A., & Highfield, T. (2013). Political networks on Twitter. *Information, Communication & Society*, 16(5), 667–691. <http://doi.org/10.1080/1369118X.2013.782328>
- Cui, H., Mittal, V., & Datar, M. (2006). Comparative experiments on sentiment classification for online product reviews. *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI'06) - Volume 2*, 1265-1270
- Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, 61(8), 1612–1621. <http://doi.org/10.1002/asi.21344>

- Graham, T., Broersma, M., Hazelhoff, K., & Haar, G. van 't. (2013). Between broadcasting political messages and interacting with voters. *Information, Communication & Society*, 16(5), 692–716. <http://doi.org/10.1080/1369118X.2013.785581>
- Grant, W. J., Moon, B., & Grant, J. B. (2010). Digital Dialogue? Australian Politicians' use of the Social Network Tool Twitter. *Australian Journal of Political Science*, 45(4), 579–604. <http://doi.org/10.1080/10361146.2010.517176>
- Hemphill, L., Otterbacher, J., & Shapiro, M. (2013). What's congress doing on twitter? In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 877–886). ACM.
- Hoang, T.-A., Lim, E.-P., Achananuparp, P., Jiang, J., & Zhu, F. (2011). On Modeling Virality of Twitter Content. In C. Xing, F. Crestani, & A. Rauber (Eds.), *Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation* (pp. 212–221). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-24826-9_27
- Hu, Y., John, A., Seligmann, D., & Wang, F. (2012). What Were the Tweets About? Topical Associations between Public Events and Twitter Feeds. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4658>
- Hu, Y., Wang, F., & Kambhampati, S. (2013). Listening to the crowd: automated analysis of events via aggregated twitter sentiment (pp. 2640–2646). Presented at the Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, AAAI Press.
- Instagram. (2015). Press News. Retrieved from <https://instagram.com/press/>
- Lee, J., Ahn, J., Oh, J. S., & Ryu, H. (2015). Mysterious Influential Users in Political Communication on Twitter: Users' Occupation Information and Its Impact on Retweetability. *iConference 2015 Proceedings*. Retrieved from <https://www.ideals.illinois.edu/handle/2142/73456>
- Naveed, N., Gottron, T., Kunegis, J., & Alhadi, A. C. (2011). Bad News Travel Fast: A Content-based Analysis of Interestingness on Twitter (pp. 1–7). Presented at the ACM WebSci'11, Koblenz, Germany. Retrieved from http://www.websci11.org/fileadmin/websci/Papers/50_paper.pdf
- Pak, E., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. In *In Proceedings of the Seventh Conference on International Language Resources and Evaluation* (Vol. 10).
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs Up?: Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10* (pp. 79–86). Stroudsburg, PA, USA: Association for Computational Linguistics. <http://doi.org/10.3115/1118693.1118704>
- Pew Research Center. (2015, October 2). Social Networking Fact Sheet. Retrieved from <http://www.pewinternet.org/fact-sheets/social-networking-fact-sheet/>
- Rainie, L., Brenner, J., & Purcell, K. (2015, October 2). Photos and Videos as Social Currency Online. Retrieved from <http://www.pewinternet.org/2012/09/13/photos-and-videos-as-social-currency-online/>
- Stieglitz, S., & Dang-Xuan, L. (2012). Political Communication and Influence through Microblogging—An Empirical Analysis of Sentiment in Twitter Messages and Retweet Behavior. In *2012 45th Hawaii International Conference on System Science (HICSS)* (pp. 3500–3509). <http://doi.org/10.1109/HICSS.2012.476>
- Wang, Y., Wang, S., Tang, J., Liu, H., & Li, B. (2015). Unsupervised Sentiment Analysis for Social Media Images. *International Joint Conference on Artificial Intelligence*.
- Yajuan, D., Zhum, C., & Furu, W. E. (2012). *Twitter Topic Summarization by Ranking Tweets Using Social Influence and Content Quality*. Proceedings of COLING 2012.
- Yang, Y., Jia, J., Zhang, S., Wu, B., Chen, Q., Li, J., ... Tang, J. (2014). How Do Your Friends on Social Media Disclose Your Emotions? (Vol. 14, pp. 1–7). Presented at the Proc. AAAI.