Big Bird, Binders, and Bayonets: Humor and live-tweeting during the 2012 U.S. Presidential Debates

Kevin Driscoll
University of Southern California
US
kedrisco@usc.edu

Mike Ananny
University of Southern California
US
ananny@usc.edu

François Bar
University of Southern California
US
fbar@usc.edu

Kristen Guth
University of Southern California
US
kguth@usc.edu

Abe Kazemzadeh
University of Southern California
US
kazemzad@usc.edu

Alex Leavitt
University of Southern California
US
aleavitt@usc.edu

Kjerstin Thorson
University of Southern California
US
kjerstin.thorson@usc.edu

Abstract

During the 2012 U.S. election cycle, social media analytics services were eager to demonstrate the efficacy of their tools to capture public opinion on Twitter. Graphics they produced to illustrate user sentiment regarding the candidates and issues were later reproduced by major news organizations. To better understand the particular practices that undergird such summary representations, we collected 35,247,043 tweets during the three televised presidential debates, nearly half of which were made up of retweets. Using a combination of quantitative content analysis and software-assisted close textual analysis, we examined the use of humor and sarcasm, 'astroturfing' by campaigns and other strategic actors, and the prevalence of retweeting 'bots.' Although sentiment analysis systems rarely disclose their methodologies, the diverse practices we encountered in this corpus makes clear that large-scale computational methods must account for the local contexts within which tweets are produced if they are to report meaningful statistics.

Keywords

humor; politics; sentiment analysis; social media; content analysis

Introduction

Researchers in both industry and the academy are attempting to analyze sentiment on Twitter, assigning numerical scores to tweets through algorithmic analysis. Simultaneously, debates about the value of measuring public opinion through this medium have questioned the interpretive use of such techniques to summarize a vast array of opinions (Sifry, 2012; Metaxas et al., 2011; Gayo-Avello, 2011). During the 2012 U.S. election cycle, providers such as Crimson Hexagon, Topsy, and Radian6 partnered with news organizations that were eager to find ways to report on the millions of Tweets and Facebook posts about the campaign and candidates (Petulla, 2013). Dashboards illustrating support for Obama and Romney were broadcast on CNN; Politico tried to predict the outcome of the Republican primaries using sentiment on Facebook. These representations were often offered as a complement to traditional polling as a more instantaneous assessment of public mood. Sentiment analysis became an integral component of the campaign coverage.

Sentiment-based quantifications of public opinion are contested constructs, not unlike the emergence of polling data as a technology of public opinion in the first half of the twentieth century (Igo, 2008). Technical questions about Twitter sentiment analysis remain: Do sentiment algorithms accurately capture the intended meanings of tweet authors? When Twitter users represent only 16 percent of the population (Pew Internet & American Life Project, 2012), what publics does sentiment analysis

capture? Although not widely recognized within news media framings, these technical debates are intertwined with the social practices of tweeting politics.

This paper contributes to this debate through a multi-faceted analysis of tweets posted during each of the three 2012 presidential debates. While some studies have documented the practice of live-tweeting during televised media events (Bruns & Burgess, 2011; McPherson et al., 2012; Shamma, Kennedy, & Churchill, 2010) and others have explored the technical complexities of sentiment algorithms (Mejova, Srinivasan, & Boynton, 2012; Skoric et al., 2012; Metaxas et al., 2012), these literatures remain relatively disconnected. Our aim is to straddle these literatures, showing how different meanings of large-scale methodological tools like sentiment analysis depend upon understanding both online, networked social practices and assumptions underlying 'big data' collection and interpretation. We focus on retweeting and livetweeting practices, in particular as they relate to humor, as we suspect prevalent sentiment analysis approaches may overemphasize their importance in political discourse. We examine practices such as the use of humor and sarcasm (Mejova, Srinivasan, & Boynton, 2013), 'astroturfing' by campaigns and other strategic actors (Ratkiewicz et al., 2011), and retweeting 'bots' that distort the implicit "one tweet, one opinion" assumption behind many sentiment analysis algorithms (Metaxas et al., 2012).

Data Collection & Methods

During each 2012 presidential debate, we collected tweets matching an evolving set of 427 debate-related keyword rules using the commercial data provider Gnip PowerTrack. Our observation period began one hour before each debate's start and lasted until four hours after the conclusion of the televised broadcast. The full corpus consists of 35,247,043 tweets produced by 5,833,227 unique user accounts. The size of this collection is comparable with statistics reported by journalists and social media analytics firms in the days following the debates (e.g., Sharp, 2012). Our analysis focuses on the 17,579,576 retweets within these data. Retweets accounted for 49.9% of the total corpus of tweets we collected during the three debates. To investigate livetweeting practices and their relationships to sentiment analysis, we used both quantitative content analysis and software-assisted textual analysis.

Identifying Comedic Accounts

First, we conducted a content analysis on the top quartile of most-tweeted accounts (N=197).This small group produced 18,425 tweets that were retweeted 4,388,446 times collectively during (and in the hours following) the three presidential debates. We found that 62 (31%) of the top quartile of the most retweeted accounts were primarily used for humor, comedy, parody and sarcasm. These 62 humorous accounts were responsible, directly or indirectly via retweet, for 1,370,537 (4%) of all the tweets in our collection. To put this concentration of comic accounts in context, we also coded a random sample of the remaining three quartiles and found that the number of accounts clearly used only for humor was negligible.

Emergent Live-Tweeting Practices

The results of the content analysis identified accounts that were exclusively comic, however, it was clear that many of the most retweeted tweets were humorous – even if the accounts that sent them were not. To investigate the use of humor by accounts that are not strictly comedic, we are in the midst of identifying thematic clusters of tweets in the top quartile of most retweeted tweets and conducting close readings of each thematic cluster. This analysis indicates a set of live-tweeting practices that have yet to be addressed by the existing literature.

Several of the most retweeted jokes appear to be direct, unattributed copies of other widely-circulated tweets – a sort of "joke plagiarism." Jokes appear to have significant currency as a means of boosting one's visibility on Twitter (as measured by retweet count and number of followers), providing Twitter users with an incentive to plagiarize successful jokes. A related phenomenon that appears less

frequently involves comedic accounts tweeting the exact same message during multiple debates. These practices suggest that the authors of these accounts are tracking the circulation and reception of their tweets.

Finally, several of the phenomena identified above might be attributed to semi-autonomous software agents ("bots"). In particular, we suspect that certain bots systematically retweet jokes that have been widely circulated in order to gain followers. Although we can only speculate about the motives for creating such bots, a likely motivation is that highly visible accounts are valued more highly on the grey market by spammers who intend to use the accounts to circulate phishing links.

Conclusion

Our findings highlight the need to take into account social and cultural contexts, and the practices of production and re-circulation when using computational methods to interpret large bodies of political expression (boyd & Crawford, 2011). These results indicate that many of the most widely circulated tweets are, in fact, jokes produced by Twitter users who exclusively trade in humor, suggesting an important revision of our initial questions regarding the representativeness of Twitter for purposes of measuring public sentiment. Although Twitter seems unlikely to represent the opinions of voters or citizens, it nevertheless presents a valuable opportunity for scholars of political communication and public deliberation to observe the negotiation of conflict through the use of humor. Furthermore, few of the social media dashboard providers publicly disclose their expertise at interpreting humor, sarcasm, or other unique social behaviors. Yet, the prevalence of humor makes clear that such systems must contend with sarcasm, satire, and context-dependent political humor if they are to report meaningful statistics.

Acknowledgments

References

boyd, d. & Crawford, K. (2011). Six Provocations for Big Data. Oxford Internet Institute's "A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society" on September 21, 2011.

Bruns, A., & Burgess, J. E. (2011). New methodologies for researching news discussion on Twitter. The Future of Journalism 2011, 8 - 9 September 2011, Cardiff University, Cardiff, UK.

Gayo-Avello, D. (2011). Don't turn social media into another 'Literary Digest' poll. *Communications of the ACM*, 54(10), 121-128.

Igo, S. E. (2007). *The averaged American: Surveys, citizens, and the making of a mass public*. Cambridge, MA: Harvard University Press.

McPherson, K., Huotari, K., Cheng, F., Humphrey, D., Cheshire, C., & Brooks, A. L. (2012, February). Glitter: a mixed-methods study of twitter use during glee broadcasts. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion, 167-170.

Mejova, Y., Srinivasan, P., & Boynton, B. (2013, February). GOP primary season on twitter: popular political sentiment in social media. In Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, 517-526.

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011, October). How (not) to predict elections. In Privacy, Security, Risk and Trust , IEEE Third International Conference on Social Computing, 165-171.

Metaxas, P.T., & Mustafaraj, E. (2012). Social Media and the Elections. Policy Forum (October 26, 2012). Retrieved from http://www.lawnet.gr/assets/files/somed.pdf.

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010, May). From tweets to polls: Linking text sentiment to public opinion time series. In Proceedings of the International AAAI Conference on Weblogs and Social Media, 122-129.

Petulla, S. (2013, January 23) Feelings, nothing more than feelings: The measured rise of sentiment analysis in journalism. *Neiman Journalism Lab*. Retrieved from: http://www.niemanlab.org/2013/01/feelings-nothing-more-than-feelings-the-measured-rise-of-sentiment-analysis-in-journalism/.

Pew Internet & American Life Project. (2012, February). *What Users Do Online* [Data file]. Retrieved from: http://www.pewinternet.org/Static-Pages/Trend-Data-(Adults)/Online-Activites-Total.aspx.

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media. In Proceedings of the ICWSM, 297-304.

Shamma, D. A., Kennedy, L., & Churchill, E. F. (2010). Conversational shadows: Describing live media events using short messages. In Proceedings of ICWSM, 331-334.

Sharp, A. (2012, October 4). Dispatch from the Denver debate. *Official Twitter blog*. Accessed on February 28, 2013: http://blog.twitter.com/2012/10/dispatch-from-denver-debate.html

Sifry, M. (2012, January 13). Politico-Facebook Sentiment Analysis Will Generate "Bogus" Results, Expert Says. *Tech President*. Retrieved from: http://techpresident.com/news/21618/politico-facebook-sentiment-analysis-bogus

Skoric, M., Poor, N., Achananuparp, P., Lim, E. P., & Jiang, J. (2012). Tweets and Votes: A Study of the 2011 Singapore General Election. In Proceedings of the 45th Hawaii International Conference on System Science (HICSS), 2583-2591.

License