

CS 6474/CS 4803 Social Computing: Credibility

Munmun De Choudhury

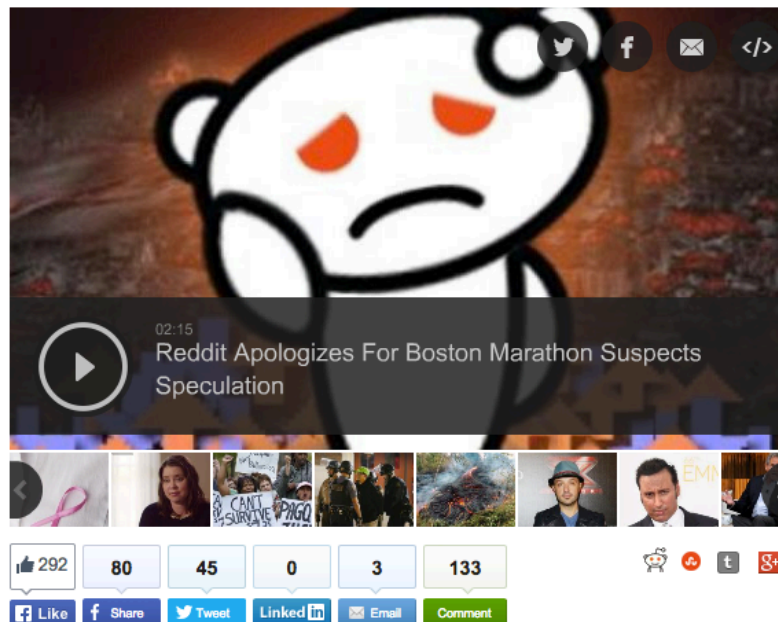
munmund@gatech.edu

Week 9 | October 17, 2018

Reddit Apologizes For Speculating About Boston Marathon Suspects

The Huffington Post | By Katherine Bindley

Posted: 04/22/2013 5:18 pm EDT | Updated: 04/23/2013 7:31 pm EDT

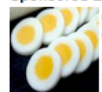


Speculation on social media last week over who was responsible for the [bombing at the Boston Marathon](#) produced its own set of innocent victims: the falsely accused.

Reddit -- which was fiercely criticized for its ["Findthebostonbombers"](#) thread that called out specific people standing near the scene as suspects -- has now issued a [public apology](#) to those whose names were dragged through the mud.

"We all need to look at what happened and make sure that in the future we do everything we can to help and not hinder crisis situations," the statement reads.

Sponsored Links



Odd Trick Fights Diabetes
"Unique" Proven Method To Control Blood Sugar In 3 Weeks. Watch Video.
Smart-Consumer-Lifestyles.com

Know Where You Stand

The message, which was published in a blog post on Monday, continues by mentioning the serious ramifications of the rampant speculation that occurred on the site.

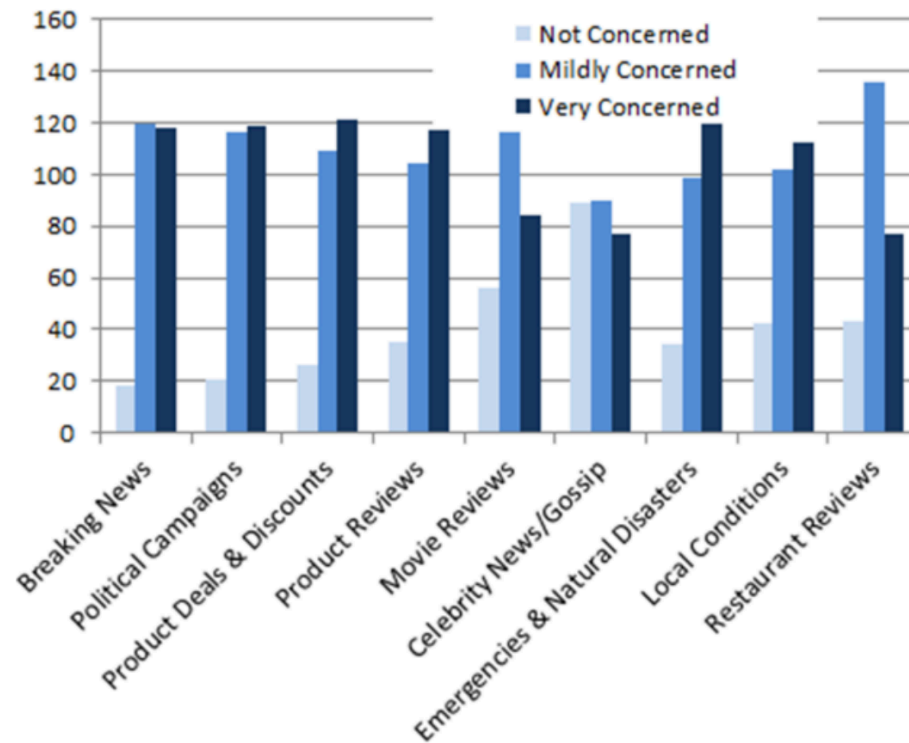
Tweeting is Believing?
Understanding Microblog
Credibility Perceptions

Summary

- The article presents results of a survey study that gauged users' perceptions of tweet credibility.
- 26 features identified to impact credibility assessment based on a pilot sample of five individuals.
- Two controlled experiments to measure the impact of several tweet features (message topic, user name, and user image) on perceptions of message and author credibility
 - Participants recruited from inside Microsoft and through a message board of CMU
- Total of 256 completed surveys, 101 from the corporate group and 155 from the alumni group.
- Main findings:
 - Users are poor judges of tweet credibility based on content alone; reliance on cues such as username and presence/absence of profile pictures.

One limitation of the work is that their current recruitment method does not include certain demographics that consume tweets, like teenagers or adults without a college degree; education may matter

The paper focused on a rather well-educated and specialized group of participants, and that it failed to contrast results of this population and a more general population



Morris et al. focus on assessing credibility of news. Would same observations apply to judging credibility of non-real time information? E.g., health myths

Class Reading – Significance of assessing credibility of anti-vax information

Class Exercise I

Credibility is, after all, a domain-dependent attribute. Take the example of the anti-vax health myth. What additional new feature would you consider, in addition to the ones raised in Morris et al., that could be useful for assessing credibility? How would you factor in end users' bias in perception of credibility?

Which of these pictures of Hurricane Sandy are Real and which ones are fake?



Amazing picture of hurricane #Sandy decending in New York
pic.twitter.com/3mMhCbNq



29 Oct 12

← Reply ↻ Retweet ★ Favorite



DC Maryland Virginia
@DMVFollowers



Follow

McDonalds in Virginia Beach flooded.

pic.twitter.com/FZBoCydM

← Reply ↻ Retweet ★ Favorite



Class Exercise II

Morris et al. examined and studied credibility on Twitter – a primarily text based content system. How would the different cues change if we look at the host of new multimedia sharing social apps (e.g., Instagram) and want to assess credibility of such content?

7 Fake Hurricane Sandy Photos You're Sharing on Social Media



29 Oct 12 Reply Retweet Favorite



DC Maryland Virginia
@DMVFollowers

McDonalds in Virginia Beach flooded.
pic.twitter.com/FZBoCydM
Reply Retweet Favorite



Modulating Video Credibility via Visualization of Quality Evaluations

Nicholas Diakopoulos
School of Communication and Information
Rutgers University
4 Huntington St., New Brunswick, NJ, USA
diakop@rutgers.edu

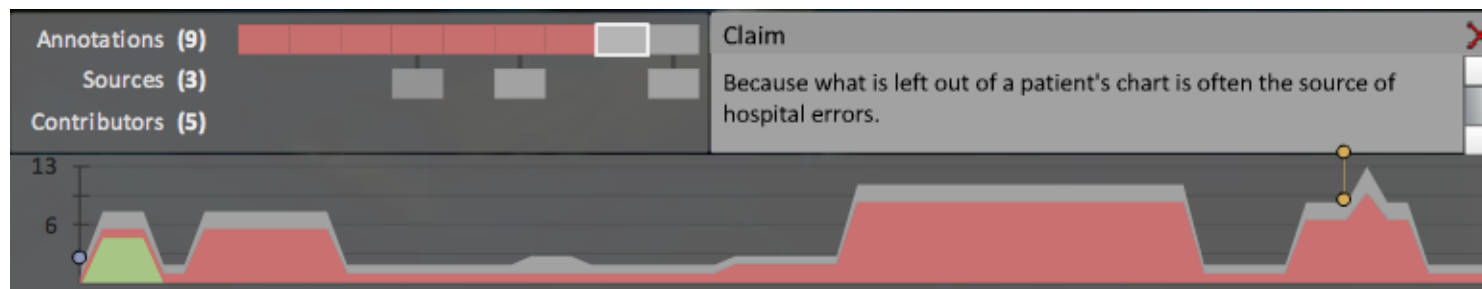
Irfan Essa
School of Interactive Computing
Georgia Institute of Technology
75 5th St. NW, Atlanta, GA, USA
irfan@cc.gatech.edu

ABSTRACT

In this work we develop and evaluate a method for the syndication and visualization of aggregate quality evaluations of informational video. We enable the sharing of knowledge between motivated media watchdogs and a wider population of casual users. We do this by developing simple visual cues which indicate aggregated activity levels and polarity of quality evaluations (i.e. positive / negative) which are presented in-line with videos as they play. In an experiment we show the potential of these visuals to engender constructive changes to the credibility of informational video under some circumstances. We discuss the limitations, and future work associated with this approach toward video credibility modulation.

information quality by combing through the media and engaging in fact-checking and re-contextualization of news and other media reports. For high profile video events such as the State of the Union address given by the president of the U.S., there is a considerable demand for this type of watchdogging activity. For instance, recent coverage by news outlets like PBS included annotated transcripts and video snippets showing analysis from experts and journalists¹. One of the major issues with such analytic presentations as are found on Politifact, Factcheck, and PBS is that, especially for video, the analysis is divorced from the video itself, making the multimedia context difficult to understand in relationship with the textual analysis.

While most methods of watchdogging are labor intensive, another



Finding and Assessing Social Media Information Sources in the Context of Journalism

Nicholas Diakopoulos¹, Munmun De Choudhury², Mor Naaman¹

¹Rutgers University – School of Communication and Information, ²Microsoft Research
diakop@rutgers.edu, munmund@microsoft.com, mor@rutgers.edu

Tottenham Riots
402 sources sharing 551 tweets matching "tottenhamriots" or "tottenham"

Search (enter search terms here) **Search** **Sort** # times retweeted

Show Sources (showing 8 of 10 sources loaded)
All Ordinary People Journalists / Bloggers Organizations Uncategorized Eyewitnesses

Show Tweets All Exclude RTs Images & Videos

1.1 Filters **1.2 Location** **1.3 User Type** **1.4 Eyewitness Icon** **1.5 Friends' Locations** **1.6 Network Sketch** **1.7 Entities**

Daniel Carr, @daniel_carr (2 years, 3 months old)
Myself in 160 characters: Schizophrenic. Also a criminologist
London, United Kingdom
Jouno/Blogger 41 RTed 56 Klout

213 Followers
163 Following

NETWORK SKETCH
FRIENDS' LOCATIONS
London, GB 34.48% Glasgow, GB 5.17% Manchester, GB 3.45%

TOP ENTITIES MENTIONED HISTORICALLY
Bruce Grove, Tottenham Hale, London, BBC, Haha,

31 Tweets
#tottenham #tottenhamriots Fire near Bruce Grove Station, larger one towards Lordship Lane Aug. 6, 2011, 11:27 p.m.
#tottenham #tottenhamriots @MrsCheddies by Bruce Grove I mean north of previous fires, on High Rd towards Lordship Lane Aug. 6, 2011, 11:24 p.m.
@hackneyhive yeah around that area there are 2 fires, one small now, one very large #tottenham #tottenhamriots

Aidan Rowe, @Aidan_Rowe (1 year, 2 months old)
Post-punk, proto-utopian, anarchist, activist, musician, blogger, student, failed comedian.
<http://redwriters1.blogspot.com>
Ordinary Person 23 RTed 49 Klout

215 Followers
395 Following

NETWORK SKETCH
FRIENDS' LOCATIONS
Dublin, IE 43.48% London, GB 4.35% Cork, IE 1.74%

TOP ENTITIES MENTIONED HISTORICALLY
Oslo, BBC, Dublin, Dermot Mulqueen, Johann Hari,

5 Tweets
"Why couldn't the people in #Tottenham just have held a nice dignified protest for us to ignore?" - Liberals #tottenhamriots Aug. 7, 2011, 12:49 a.m.
Any reports of arrests? #tottenham #tottenhamriots Hope everyone is safe. #acab Aug. 7, 2011, 12:06 a.m.
Anyone using the words "mindless", "hooligans" or "thugs" is a racist and an idiot. #tottenham #tottenhamriots

Newsworthiness and Network Gatekeeping on Twitter: The Role of Social Deviance

Nicholas Diakopoulos

Columbia University, School of Journalism
nicholas.diakopoulos@gmail.com

Arkaitz Zubiaga

Dublin Institute of Tech., Applied Intelligence Research Centre
arkaitz@zubiaga.org

Abstract

Publishers of news information are keen to amplify the reach of their content by making it as re-sharable as possible on social media. In this work we study the relationship between the concept of social deviance and the re-sharing of news headlines by network gatekeepers on Twitter. Do network gatekeepers have the same predilection for selecting socially deviant news items as professionals? Through a study of 8,000 news items across 8 major news outlets in the U.S. we predominately find that network gatekeepers re-share news items more often when they reference socially deviant events. At the same time we find and discuss exceptions for two outlets, suggesting a more complex picture where newsworthiness for networked gatekeepers may be moderated by other effects such as topicality or varying motivations and relationships with their audience.

network gatekeepers. Previous literature has found that events with high deviance were more likely to gain coverage in the main stream media (MSM) (P. Shoemaker, Danielian, and Brendlinger 1991), “the role of news media is not to mirror the world as it is, but rather to spotlight and draw public attention to problems and situations that need solutions and repair,” (Pamela Shoemaker 2006). Shoemaker’s theory posits that MSM will select for and favor socially deviant stories and events.

But do network gatekeepers share a similar predilection for selecting socially deviant news items? Is social deviance a professionally imbued newsworthiness criterion, or something that more generally explains interest, attention, and sharing of news? Networked gatekeeping theory explores the idea that every user on Twitter is a gatekeeper, with the discretion to share or not

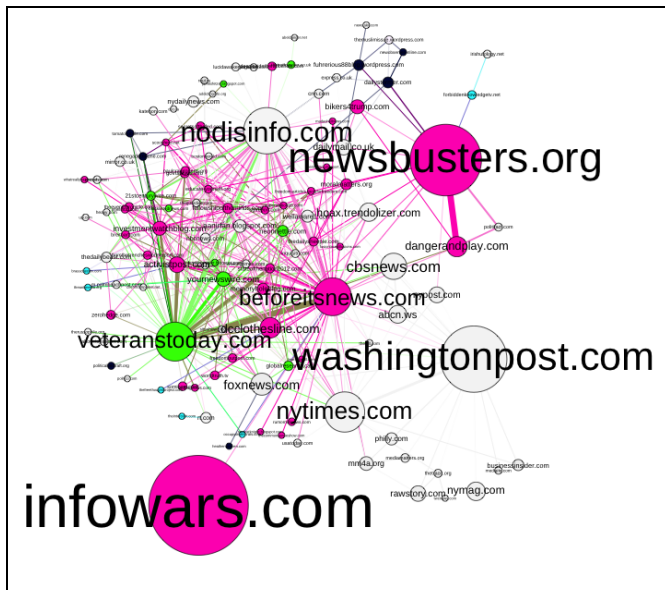
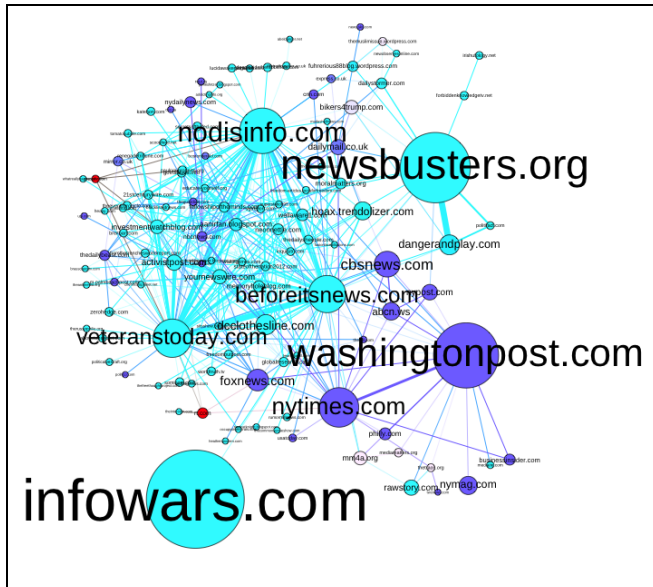
A need for “fact checking systems” that operate outside of the social media ecosystem. But these systems are difficult to build and use. Why?

Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter

Summary (1)

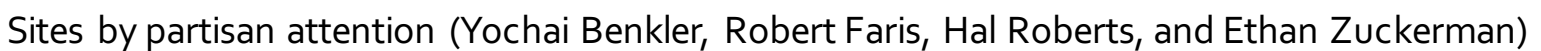
- The paper presents the first study of “fake news”
- The context: in recent years, alternative media outlets have appropriated social media platforms for their perceived economic and political reach and for hosting inaccurate or under-sourced content
- Goals:
 - Provide a systematic lens for exploring the production of a certain type of “fake news”—alternative narratives of man-made crisis events
 - Examine the production of alternate narratives (rumors, conspiracy theories) through Twitter and across the external websites that Twitter users reference as they engage in these narratives

Summary (2)



Leaning	Description
U.S. Alt Right	U.S. focused, anti-mainstream media, pro-Christian, anti-LGBT, anti-feminist, anti-globalist, climate change denying
U.S. Alt Left	U.S. focused, anti-mainstream media, anti-corporatist, critical of police, pro-prison reform, pro-BlackLivesMatter
International Anti-Globalist	Internationally focused, anti-globalist or anti-New World Order/Cabal, anti-corporatist, conspiracy-focused
White Nationalist and/or Anti-Semitic	primarily white-nationalist or anti-Semitic positions
Muslim Defense	primarily challenges mainstream narratives of terrorist attacks by Muslims
Russian Propaganda	primarily supports Russian interests, anti-globalist

- Strong political agendas underlying many of alternative narratives and the domains that hosted them
- More than half of the alternative media sites were coded as primarily motivated by a political agenda—with the conspiracy theories serving a secondary purpose of attracting an audience and reflecting or forwarding that agenda



Role of social media in supporting conspiracy theories

Understanding Anti-Vaccination Attitudes in Social Media

Tanushree Mitra^{1,2}

¹Georgia Institute of Technology

tmitra3@gatech.edu

Scott Counts²

²Microsoft Research

counts@microsoft.com

James W. Pennebaker^{2,3}

³University of Texas at Austin

pennebaker@mail.utexas.edu

Abstract

The anti-vaccination movement threatens public health by reducing the likelihood of disease eradication. With social media's purported role in disseminating anti-vaccine information, it is imperative to understand the drivers of attitudes among participants involved in the vaccination debate on a communication channel critical to the movement: Twitter. Using four years of longitudinal data capturing vaccine discussions on Twitter, we identify users who persistently hold pro and anti attitudes, and those who newly adopt anti attitudes towards vaccination. After gathering each user's entire Twitter timeline, totaling to over 3 million tweets, we explore differences in the individual narratives across the user cohorts. We find that those with long-term anti-vaccination attitudes manifest conspiratorial thinking, mistrust in government, and are resolute and in-group focused in language. New adoptees appear to be predisposed to form anti-vaccination attitudes via similar government distrust and general paranoia, but are more social and less certain than their long-term counterparts. We discuss how this apparent predisposition can interact with social media-fueled events to bring newcomers into the anti-vaccination movement. Given the strong base of conspiratorial thinking underlying anti-vaccination attitudes, we conclude by highlighting the need for alternatives to traditional methods of using authoritative sources such as the government when correcting misleading vaccination claims.

persistent vaccine criticism movement has spread rapidly through social media, a channel often used to disseminate medical information without verification by the expert medical community (Keelan et al. 2010).

Given the increasing reliance on online media for accurate health information and the general growth of social media sites, the attitudes of anti-vaccination advocates risk becoming a global phenomenon that could impact immunization behavior at significant scale (Kata 2010). In fact a controlled study showed that parents opting to exempt children from vaccination are more likely to have received the information online compared to those vaccinating their kids (Salmon et al. 2005). These parents benefit from “herd immunity” in which eradication is achieved by immunizing a critical proportion of the population. However, as internet-fueled misbeliefs drive people to opt out of vaccination, herd immunity is weakened, increasing the chances of a disease outbreak. Thus it is important to understand the underlying characteristics of individuals with anti-vaccination attitudes. What drives people to develop and perpetuate the anti-vaccination movement?

In this paper we explore this question by examining individuals' overt expressions towards vaccination in a social

Class Exercise III

Starbird found that alternative media sites may give the false perception of being exposed to a diverse information diet. Most of them, seemingly different on the surface, harp on the same political (e.g., anti-globalist) agenda.

How can social media platforms ensure information diversity while curbing the negative impact of fake news?



Social bots distort the 2016 U.S. Presidential election online discussion

by Alessandro Bessi and Emilio Ferrara

Abstract

Social media have been extensively praised for increasing democratic discussion on social issues related to policy and politics. However, what happens when this powerful communication tools are exploited to manipulate online discussion, to change the public perception of political entities, or even to try affecting the outcome of political elections? In this study we investigated how the presence of social media bots, algorithmically driven entities that on the surface appear as legitimate users, affect political discussion around the 2016 U.S. Presidential election. By leveraging state-of-the-art social bot detection algorithms, we uncovered a large fraction of user population that may not be human, accounting for a significant portion of generated content (about one-fifth of the entire conversation). We inferred political partisanship from hashtag adoption, for both humans and bots, and studied spatio-temporal communication, political support dynamics, and influence mechanisms by discovering the level of network embeddedness of the bots. Our findings suggest that the presence of social media bots can indeed negatively affect democratic political discussion rather than improving it, which in turn can potentially alter public opinion and endanger the integrity of the Presidential election.

Contents

[Introduction](#)
[Methodology](#)
[Data analysis](#)
[Conclusions](#)

Introduction

Various computational social science studies demonstrated that social media have been extensively used to foster democratic conversation about social and political issues: From the Arab Spring (González-Bailón, *et al.*, 2011; Howard, *et al.*, 2011), to Occupy Wall Street (Conover, *et al.*, 2013a; Conover, *et al.*, 2013b) and many other civil protests (Varol, *et al.*, 2014; González-Bailón, *et al.*, 2013) (Bastos, *et al.*, 2014), Twitter and other social media seemed to play an instrumental role to involve the public in policy and political conversations, by collectively framing the narratives related to particular social issues, and coordinating online and off-line activities. The use of digital media to discuss politics during election times has also been the subject of various studies, covering the last four U.S. Presidential elections (Adamic and Glance, 2005; Diakopoulos and Shamma, 2010; Bekafigo and McBride, 2013; Carlisle and Patton, 2013; DiGrazia, *et al.*, 2013; Wang, *et al.*, 2016), and other countries like Australia (Gibson and McAllister, 2006; Bruns and Burgess, 2011; Burgess and Bruns, 2012), and Norway (Enli and Skogerboe, 2013). Findings that focused on the positive effects of social media such as incrementing voting turnout (Bond, *et al.*, 2012) or exposure to diverse political views (Bakshy, *et al.*, 2015) contributed to the general praise of these platforms as a tool to foster democracy and civil political engagement (Shirky, 2011; Loader and Mercea, 2011; Effing, *et al.*, 2011; Tufekci and Wilson, 2012; Tufekci, 2014; Yang, *et al.*, 2016).

However, as early as 2006, Philip Howard raised concerns regarding the possibility of manipulating public opinion and spreading political misinformation through social media (Howard, 2006). These issues have been later proved true by several studies (Ratkiewicz, *et al.*, 2011a; Ratkiewicz, *et al.*, 2011b) (Metaxas and Mustafaraj, 2012) (El-Khalili, 2013; Ferrara, 2015; Woolley and Howard, 2016; Shorey and Howard, 2016). Of particular concern is the fact social media have been demonstrated effective in influencing individuals (Aral and Walker, 2010). One way to perform such type of manipulation is by using social bots, algorithmically controlled accounts that emulate the activity of human users but operate at much higher pace (*e.g.*, automatically producing content or engaging in social interactions), while successfully keeping their artificial identity undisclosed (Hwang, *et al.*, 2012; Messias, *et al.*, 2013; Ferrara, *et al.*, 2016).

Evidence of the adoption of social media bots to attempt manipulating political communication dates back half a decade: during the 2010 U.S. midterm elections, social bots were employed to support some candidates and smear others, by injecting thousands of tweets pointing to Web sites with fake news (Ratkiewicz, *et al.*, 2011a). The research community reported another similar case around the time of the 2010 Massachusetts special election (Metaxas and Mustafaraj, 2012). Campaigns of this type are sometimes referred to as *astroturf* or *Twitter bombs*. Unfortunately, most of the times, it has proven impossible to determine who's behind these types of operations (Kollanyi, *et al.*, 2016; Ferrara, *et al.*, 2016). Governments, organizations, and other entities with sufficient resources, can obtain the technological capabilities to deploy thousands of social bots and use them to their advantage, either to support or to attack particular political figures or candidates. Indeed, it has become increasingly simpler to deploy social bots, so that, in some cases, no coding skills are required to setup accounts that perform simple automated activities: tech blogs often post tutorials and ready-to-go tools for this purposes [1], [2], [3]. Various source codes for sophisticated social media bots can be found online as well, ready to be customized and optimized by the more technical savvy users (Kollanyi, 2016). We inspected several of these readily available bots and this is a (non-comprehensive) list of the capabilities that they provide: Search Twitter for phrases/hashtags/keywords and automatically retweet them; Automatically reply to tweets that meet a certain criteria; automatically follow any users that tweet something with a specific phrase/hashtag/keyword; Automatically follow back any users that have followed the bot; Automatically follow any users that follow a specified user; Automatically add users tweeting about something to public lists; Search Google (and other engines) for articles/news according to specific criteria and post them, or link them in automatic replies to other users; Automatically aggregating public sentiment on certain topics of discussion; Buffer and post tweets automatically. Most of these bots can run in cloud services or infrastructures like Amazon Web Services (AWS) or Heroku, making it more difficult to block them. Finally, a very recent trend is that of providing Bot-As-A-Service (BaaS): companies like RoboLike (<https://robo-like.com/>) provide "Easy-to-use Instagram/Twitter auto bots" performing certain automatic activities for a monthly price. Advanced conversational bots powered by more sophisticated Artificial Intelligences are provided by companies like ChatBots.io that allow anyone to "Add a bot to services like Twitter, Hubot, Facebook, Skype, Twilio, and more" (<https://developer.pandorabots.com/>).

(Bessi and Ferrara)

- The article studies the presence and activity of social bots involved in the 2016 U.S. Presidential election conversation ongoing on social media
- Data – Twitter data for an extensive period prior to the election that includes all three Presidential debates – 20M tweets
- Employ machine learning techniques to identify social bots
 - Over one thousand features were used, spanning content and network structure, temporal activity, user profile data, and sentiment analysis
 - These produced a score that suggests the likelihood that the inspected account is indeed a social bot
- As many 400K Twitter accounts were social bots

Impact of social bots (1)

- Bots are easily programmable through the Twitter API and can be deployed by just about anyone with preliminary coding knowledge.
- Marginal populations use bots to create an illusion of popularity around fringe issues or political candidates.
 - For instance, the alt-right and white nationalists have used these automated proxies to artificially project hate speech and xenophobia on social-media platforms.
- Hoaxers from 4chan have been accused of using bots to game online polls in favor of Trump.
 - Politicians often refers to online polls as evidence of support.

Impact of social bots (2)

- Politicians, government agencies, hacking collectives, and militaries can use bots to disseminate lies, attack people, and cloud conversation.
- The widespread use of political bots can solidify polarization among citizens.
- Bots can also silence people and groups who might otherwise have a stake in a conversation.
- At the same time they can make some users seem more popular, they make others less likely to speak. This spiral of silence can result in less discussion and diversity in politics.
- Moreover, bots used to attack journalists might cause them to stop reporting on important issues because they fear retribution and harassment.

Should social media platforms censor the “free speech” of harmful bots?

https://www.huffingtonpost.com/entry/mark-zuckerberg-regrets-fake-news-facebook_us_59cc2039e4b05063fe0eed9d

MEDIA 09/27/2017 08:53 pm ET

830



Mark Zuckerberg: 'I Regret' Rejecting Idea That Facebook Fake News Altered Election

He admitted this after Donald Trump claimed that Facebook was “always anti-Trump.”



By Carla Herreria



Facebook CEO [Mark Zuckerberg](#) admitted on Wednesday that he was wrong to dismiss the idea that fake news shared on the giant social network affected last year's presidential election.

Zuckerberg's statement came in response to a tweeted attack from President [Donald Trump](#) hours earlier. Trump claimed that Facebook was “[always anti-Trump](#)” and accused it of colluding with news outlets that the president has deemed to be “fake news.”

Class Exercise IV

Assume you work at Facebook. Design a study to measure if consumption of fake news impacted the 2016 Presidential elections.