# The Ebb and Flow of Controversial Debates on Social Media

**Kiran Garimella**[1]
Aalto University
Helsinki, Finland
kiran.garimella@aalto.fi

**Gianmarco De Francisci Morales**
QCRI
Doha, Qatar
gdfm@acm.org

**Aristides Gionis**
Aalto University
Helsinki, Finland
aristides.gionis@aalto.fi

**Michael Mathioudakis**[1]
Aalto University & HIIT
Helsinki, Finland
michael.mathioudakis@aalto.fi

[1]Contact authors; contributed equally.

**Abstract**

We explore how the polarization around controversial topics evolves on Twitter – over a long period of time (2011 to 2016), and also as a response to major external events that lead to increased related activity. We find that increased activity is typically associated with increased polarization; however, we find no consistent long-term trend in polarization over time among the topics we study.

## 1 Introduction

We study how online discussions around controversial topics change as interest in them increases and decreases – or "ebbs and flows". We are motivated by the observation that interest in enduring controversial issues is re-kindled by major related events. The gun control debate in U.S., which is revived whenever a mass shooting occurs, is one such example. The occurrence of such an event commonly causes an increase in collective attention, as reflected in the volume of related activity in social media. Moreover, motivated by the common perception that political polarization has risen recently, we track the polarization of controversial topics over the span of multiple years.

Specifically, our study is based on Twitter data that cover five years (2011 to 2016). We track four popular controversial topics of discussion in the U.S. that are recurring and attracted considerable attention during the 2016 U.S. election cycle. Given a controversial topic, we build a *endorsement* network from the retweet information on Twitter for each day of activity, and thus obtain a time series of endorsement networks. Following (Garimella et al. 2016b), we measure the polarization reflected in each network instance by using the `Random Walk Controversy` (RWC) measure.

Our analysis then seeks to answer two questions: firstly, how polarization varies with increased volume of activity (i.e., the number of users who actively discuss the topic) in a given day; and secondly, how polarization of each topic evolves over a long period of time. With respect to the former, our findings suggest that increased volume of activity is associated with increased polarization. With respect to the latter, our findings suggest that there is no consistent long-term trend over time.



## 2 Related Work

A few studies exist on the topic of controversy in online news and social media. In one of the first papers, (Adamic and Glance 2005) study linking patterns and topic coverage of political bloggers, focusing on blog posts on the U.S. presidential election of 2004. They measure the degree of interaction between liberal and conservative blogs, and provide evidence that conservative blogs are linking to each other more frequently and in a denser pattern.

These findings are confirmed by a more recent study of (Conover et al. 2011), who focus on political communication regarding congressional midterm elections. Using data from Twitter, they identify a highly segregated partisan structure (evidenced in retweets but not replies), with limited connectivity between left- and right-leaning users.

In another recent work, (Mejova et al. 2014) consider discussions of controversial and non-controversial news over a span of 7 months. They find a significant correlation between controversial issues and the use of negative affect and biased language.

More recently, (Garimella et al. 2016b; 2016a) show that controversial discussions on social media have a well-defined structure, when looking at the *endorsement* network. They propose a measure based on random walks (RWC), which is able to identify controversial topics, and *quantify* the level of controversy of a given discussion on social media by the structure of its endorsement network.

Unlike the aforementioned works, here we are interested in dynamic aspects of polarized networks. Previous studies with similar focus includes (Lehmann et al. 2012), which examines spikes in the frequency of hashtags and identifies a classification scheme that predicts whether the hashtags correspond to endogenously or exogenously driven topics. Other related studies works include (Morales et al. 2015), who study polarization over time for the death of Hugo Chavez, and (Andris et al. 2015), who study the partisanship of the U.S. congress over a long period of time.

Moreover, we are particularly interested in the network response to external stimuli that lead to increased collective attention in the controversial topic – an issue that only very recently has seen some attention in the literature (Romero, Uzzi, and Kleinberg 2016).

Table 1: Keywords for the controversial topics.

| Topic | Keywords | #Tweets | #Users |
|---|---|---|---|
| Obamacare | obamacare, #aca | 866 484 | 148 571 |
| Abortion | abortion, prolife, prochoice, anti-abortion, pro-abortion, planned parenthood | 1 571 363 | 327 702 |
| Gun Control | gun control, gun right, pro gun, anti gun, gun free, gun law, gun safety, gun violence | 824 364 | 224 270 |
| Fracking | fracking, #frack, hydraulic fracturing, shale, horizontal drilling | 2 117 945 | 170 835 |

## 3 Dataset

Using the repositories of the Internet Archive,[1] we collect a $1\%$ sample of tweets from September 2011 to August 2016,[2] for four topics of discussion related to 'Obamacare', 'Abortion', 'Gun Control', and 'Fracking'. These topics constitute long-standing controversial issues in the U.S.[3] and have been used in previous work (Lu, Caverlee, and Niu 2015). For each topic, we use a keyword list as proposed by (Lu, Caverlee, and Niu 2015) (shown in Table 1), and extract a base set of tweets containing at least one topic-related keyword. To enrich this original dataset, we use the Twitter REST API to obtain all tweets of users who have participated in the discussion at least once.[4] Table 1 shows the final statistics for the dataset.

We process the dataset to build *retweet networks* — i.e., directed networks of users, where there is an edge between two users ($u \rightarrow v$) if $u$ retweets $v$. Polarized networks, particularly the ones considered here, can be broadly characterized by two opposing *sides*, which express different opinions on the associated topic. It is commonly understood that retweets indicate *endorsement*, and endorsement networks for controversial topics have been shown to have a bi-clustered structure (Conover et al. 2011; Garimella et al. 2016b; 2017), i.e., they consist of two well-separated clusters that correspond to the opposing points of view on the topic.

In this paper, we build upon this observation to reveal the opposing sides around a topic. In particular, following an approach from previous work (Garimella et al. 2016b), we collapse all retweets contained in the dataset of each topic into a single large static retweet network — and use the METIS clustering algorithm (Karypis and Kumar 1995) to identify two clusters that correspond to the two opposing sides. We evaluate the sides by manual inspection of the top retweeted users, URLs, and hashtags. The results are fairly consistent and accurate, and can be inspected online.[5]

[1] `https://archive.org/details/twitterstream`

[2] To be precise, we have data for 57 months from that period, as our data source has no data available for three months.

[3] According to `http://2016election.procon.org`.

[4] Up to 3200 due to limits in the API.

[5] https://mmathioudakis.github.io/polarization/

Given the traditional daily news reporting cycle, we build a time series of retweet networks with the same granularity. At this granularity level, we can easily identify spikes of interest in the topics, as seen in Figure 1. Such spikes commonly correspond to external newsworthy events (Mathioudakis and Koudas 2010). Figure 1 is manually annotated with major events associated with the observed spikes. These results support the observation that Twitter is used as an *agorá* to discuss the daily matters of public interest (De Francisci Morales, Gionis, and Lucchese 2012).

**Notation** The set of retweets that occur within a single day $d$ gives rise to one retweet network $N_d^{rt}$. Each user associated with a retweet is represented with one node in the network. There is a directed edge from user $u$ to user $v$ only when user $u$ has retweeted at least one tweet authored by user $v$. In addition, each node $u$ in the network is associated with a binary attribute $s(u) \in \{1, 2\}$ that represents the side the node belongs to.

## 4 Analysis

We quantify the polarization of a network $N_d$ via the random walk controversy (RWC) score introduced in previous work (Garimella et al. 2016b). Intuitively, the score captures whether the network consists of two well-separated clusters. The higher the value of the score, the clearer the separation between the two sides.

### Polarization vs Collective Attention

We explore how RWC varies with the number of users in the networks, which is a proxy for the amount of collective attention the topic attracts. We sort the time series of networks by volume of active users, and partition them into ten quantiles (each having an equal number of days), so that days of bucket $i$ are associated with smaller volume than those of bucket $j$, for $i < j$. For each bucket, we report the mean and standard deviation of the values for each measure, and observe the trend from lower to higher volume.

Note that RWC is carefully defined so that its expected value does not depend on the volume of underlying activity (i.e., number of network nodes).

We observe a significant pattern in the relationship between RWC and interest in the topic. Figure 2 shows RWC as a function of the quantiles of the network by retweet volume. There is a clear increasing trend, which is consistent across topics. This trend suggests that increased interest in the topic is correlated with an increase in controversy of the debate, and increased polarization of the retweet networks for the two sides.

**Non-controversial topics.** For comparison, we perform measurements over a set of non-controversial topics, defined by the hashtags *#ff*, standing for 'Follow Friday', used every Friday by users to recommend interesting accounts to follow; *#nfl*, used to discuss about American football; *#sxsw*, used to comment on the *South-by-South-West* conference; *#tbt*, standing for 'Throwback Thursday', used every Thursday by users to share memories (news, pictures, stories) from the past. We find that RWC remains in low value ranges, even as the volume of activity spikes (Figure 3).

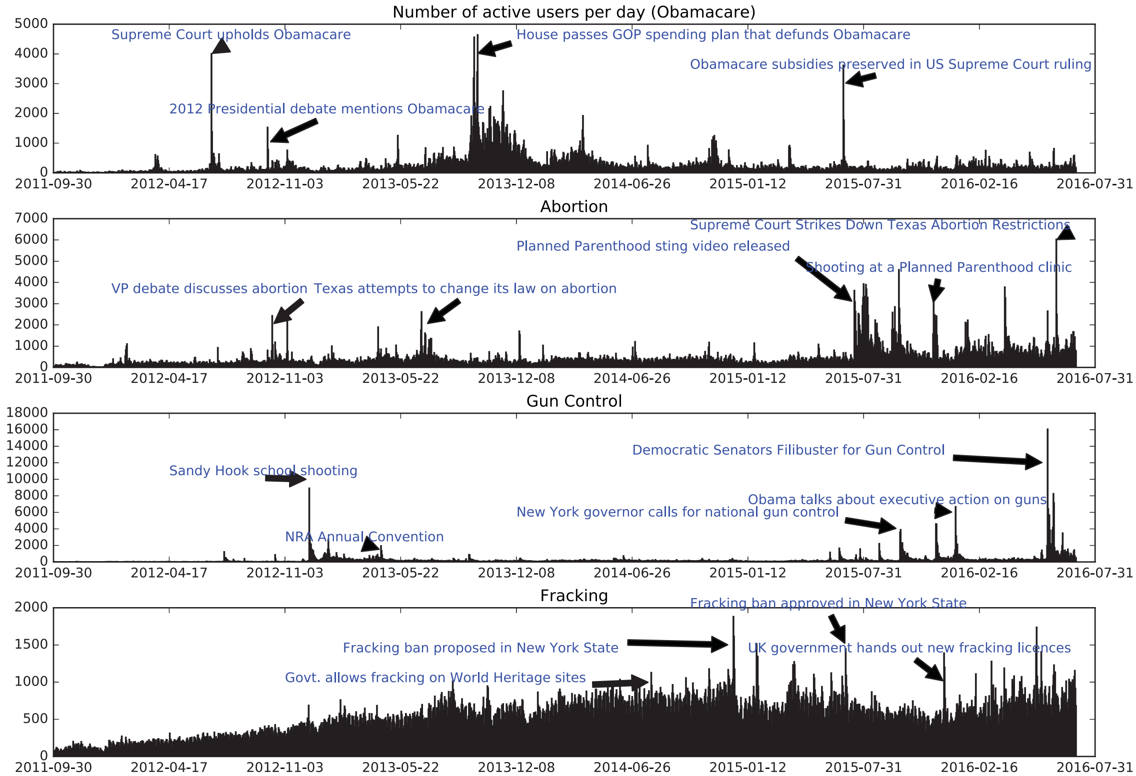


Figure 1: Daily trends for number of active users for the four controversial topics under study. Clear spikes occur at several points in the timeline. Manually chosen labels describing related events reported in the news on the same day are shown in blue for some of the spikes.

### Long-term Evolution of Polarization

Let us now focus on the long-term evolution of RWC. A common point of view holds that social media is aggravating the polarization of society and exacerbating the divisions in it (Benkler 2006). At the same time, the political debate itself (in U.S.) has become more polarized in recent years (Andris et al. 2015). However, we do not find supportive evidence for this argument in our analysis.

Figure 4 shows the long-term trends of the RWC measure for the four topics. The trend is downwards for 'abortion' and 'fracking' and slightly upwards for 'obamacare' and 'gun control'. One could argue that the latter topics are more politically linked to the current administration in U.S., and for this reason have received increasing attention with the approaching elections. However, the only safe conclusion that can be drawn from this dataset is that there is no clear signal. The figure suggests that social media, and in particular Twitter, are better suited at capturing the 'twitch' response of the public to events and news. In addition, while our dataset spans a quite long time range for typical social media studies, it is still shorter than ones used typically in social science (coming from, e.g., census, polls, congress votes). This limit is intrinsic to the tool, given that social media have risen in popularity only relatively recently (Twitter is 10 years old).

## 5 Conclusions and Future Work

We analyzed four controversial topics of discussion on Twitter for a period of five years. By examining their endorsement networks, we found that spikes in collective attention correspond to an increase in the controversy of the discussion. However, while instantaneous temporary increase in controversy happens in relation to external events, we did not find evidence of long term increase in polarization.

In future work, we plan to extend our analysis to include other network-structure and content-based measures. Equally of interest is whether the observations made in this study translate to other social media beside Twitter, for instance, Facebook or Reddit. Finally, while we did not find any consistent long-term trend in the polarization of the discussions, it is worth continuing this line of investigation, as the effects of increased polarization might not be easily discoverable from social-media analysis alone.

### Acknowledgements.

This work was supported by the Academy of Finland project "Nestor" (286211) and the EC H2020 RIA project "SoBigData" (654024).

### References

Adamic, L. A., and Glance, N. 2005. The political blogosphere and the 2004 us election: divided they blog. In

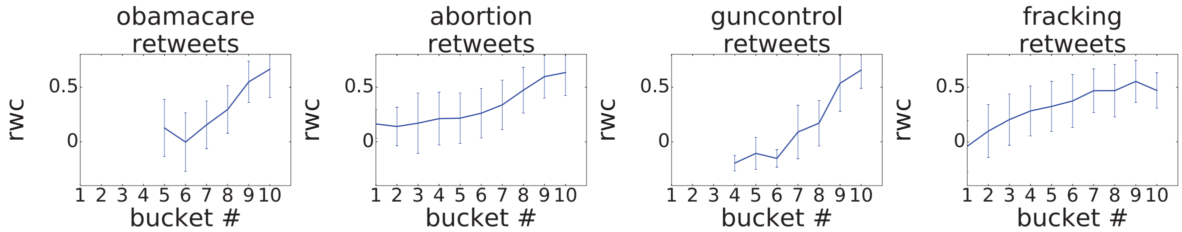


Figure 2: RWC score as a function of the activity volume in the retweet network. An increase in interest in the controversial topic corresponds to an increase in the controversy score of the retweet network.

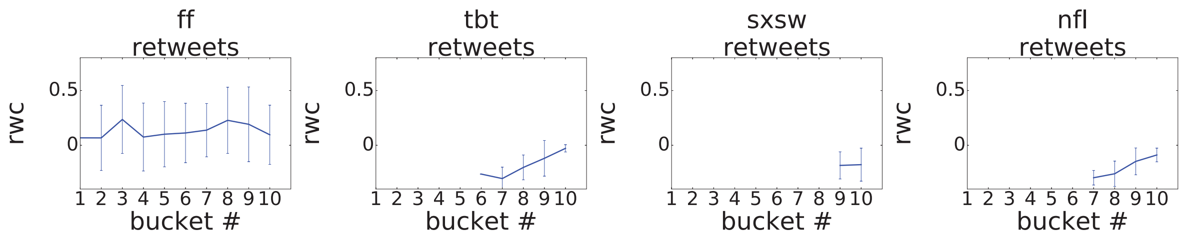


Figure 3: Non-controversial topics: RWC score as a function of the activity in the retweet network. In this case, increase in interest does not affect the controversy score of the networks.

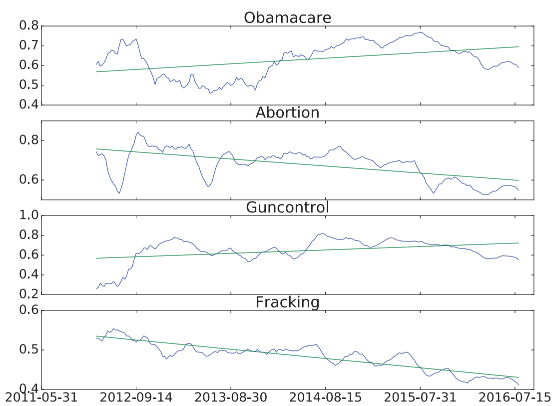


Figure 4: Long-term trends of RWC (controversy) score in our dataset. The plots show no consistent trend among the controversial topics under consideration.

*LinkKDD*, 36–43.

Andris, C.; Lee, D.; Hamilton, M. J.; Martino, M.; Gunning, C. E.; and Selden, J. A. 2015. The rise of partisanship and super-cooperators in the us house of representatives. *PloS one* 10(4):e0123507.

Benkler, Y. 2006. *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.

Conover, M.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Menczer, F.; and Flammini, A. 2011. Political Polarization on Twitter. In *ICWSM*.

De Francisci Morales, G.; Gionis, A.; and Lucchese, C. 2012. From Chatter to Headlines: Harnessing the Real-Time Web for Personalized News Recommendation. In *WSDM*, 153–162.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2016a. Exploring Controversy in Twitter. In *CSCW [demo]*.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2016b. Quantifying Controversy in Social Media. In *WSDM*, 33–42.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2017. Reducing Controversy by Connecting Opposing Views. In *WSDM*.

Karypis, G., and Kumar, V. 1995. METIS - Unstructured Graph Partitioning and Sparse Matrix Ordering System.

Lehmann, J.; Gonçalves, B.; Ramasco, J. J.; and Cattuto, C. 2012. Dynamical Classes of Collective Attention in Twitter. In *WWW*, 251–260.

Lu, H.; Caverlee, J.; and Niu, W. 2015. BiasWatch: A Lightweight System for Discovering and Tracking Topic-Sensitive Opinion Bias in Social Media. In *CIKM*, 213–222.

Mathioudakis, M., and Koudas, N. 2010. TwitterMonitor: Trend Detection over the Twitter Stream. In *SIGMOD*, 1155–1158.

Mejova, Y.; Zhang, A. X.; Diakopoulos, N.; and Castillo, C. 2014. Controversy and sentiment in online news. *Symposium on Computation + Journalism*.

Morales, A.; Borondo, J.; Losada, J.; and Benito, R. 2015. Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos* 25(3).

Romero, D. M.; Uzzi, B.; and Kleinberg, J. 2016. Social networks under stress. In *WWW*, 9–20.