Twitter Influencers and Increased Polarization during two U.S. Presidential Elections

James Flamino, Alessandro Galeazzi, Stuart Feldman, Michael W. Macy, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Alexandre Bovet, Hernan A. Makse, Boleslaw K. Szymanski

Defining polarization

• Formal definition: the state of having two opposite or contradictory tendencies, opinions, and aspects
• Polarization is commonly used when discussing politics, especially within the topic of social media
• Political scientists refine the definition into four types:
  • Policy polarization
  • Partisan polarization
  • **Ideological polarization**
  • Geographic polarization

Figure ref: McCarty, Nolan. *Polarization: What everyone needs to know*. Oxford University Press, 2019.
The four types of polarization can be refined into levels of affect:

- Elite polarization
- Media polarization
- Voter polarization

Traditionally, the degrees of influence and interaction between these levels were more hierarchical, with a high barrier of entry to disseminate information.

With the introduction of decentralized and interactive social media, the dynamics of information spread between these user types have changed.
Defining level of affect in polarization

- Social media lowers the barrier of interactions with all levels of individuals and provides inexpensive and direct channels of communication with potentially millions of people.
- This widens the range of entities that can politically influence general audiences.
- This has been established, but how do top political influencers change over time now that social media is commonplace?
- Importantly, does this facilitate an increase in polarization?
Our goals

- To quantify the types of news propagated on Twitter
- To identify “influencers” and their political biases based on the information they propagate
- To use influencers and their user bases to quantify ideological polarization
- To compare the above observations between two U.S. Presidential Elections to see how they change
The data

• We collected Twitter data on the 2016 and 2020 U.S. presidential elections
• For 2016 we collected data from June 1\textsuperscript{st} to November 8\textsuperscript{th} and for 2020 we collected from June 1\textsuperscript{st} to November 2\textsuperscript{nd}
• 171 million tweets were generated by 11 million users in 2016, and 702 million tweets were generated by 20 million users in 2020
Identifying news biases

- We extracted the domain names of each Tweet containing a URL and classified all news media outlet (e.g. cnn.com) links.
- We grouped news outlet links into 8 categories of political alignment based on the linked news media outlet’s biases:
  - Fake news
  - Extreme bias right
  - Right
  - Lean right
  - Center
  - Lean left
  - Left
  - Extreme bias left
- Classifications were derived from allsides.com and mediabiasfactcheck.com

Figure ref: https://www.allsides.com/media-bias/media-bias-chart
URL links in Twitter

A: Proportion of tweets

B: Proportion of users

C: User main category

D: Links category

Proportion of links

2016

2020
Identifying influencers – network generation

- For each news category (URL link classification) for each year we assemble a retweet network
- Retweeting often indicates interest, agreement, and trust in the content of a tweet, therefore representing endorsement of URL
- A node in the retweet network is represented by a user from our Twitter data
- An edge from node $v$ to node $u$ represents that node $u$ has retweeted some content from node $v$ that contains a classified URL
- In other words the edge indicates that $u$ is endorsing and propagating $v$’s news content
Identifying influencers – centrality measure

- Given our network topology, the higher is a node's out-degree, the greater is its local influence.
- But just using local properties (e.g. out-degree centrality) is not a comprehensive way to identify influencers (local properties will incorrectly classify users that went viral once, or just happen to produce a lot of content).
- Instead, we use the Collective Influence (CI) algorithm.
The CI algorithm finds the minimal set of nodes that can collectively reach the entire network when information diffuses according to a linear threshold model.

The CI algorithm considers influence as an emergent collective property (not local). From this, the CI algorithm can rank “super-spreaders” (influencers) of information in social networks.

Identifying influencers – centrality measure

Node $v$ (influencer) has content retweeted by Node $u$ (author).
Top influencer comparisons

2016

Fake News
1. @PrisonPlanet
2. @RealAlexJones
3. @zerohedge
4. @DRUDGE_REPORT
5. @realDonaldTrump

Right News
1. @FoxNews
2. @realDonaldTrump
3. @dcexaminer
4. @DRUDGE_REPORT
5. @realDonaldTrump

Right Leaning News
1. @WSJ
2. @WashTimes
3. @RT_com
4. @realDonaldTrump
5. @RT_America

Extreme Bias Right News
1. @realDonaldTrump
2. @DailyCaller
3. @BreitbartNews
4. @wikileaks
5. @DRUDGE_REPORT

2020

Fake News
1. @seanhannity
8. @OANN
13. @zerohedge
16. @PrisonPlanet
20. @realDonaldTrump

Right News
1. @DonaldTrumpJr
2. @marklevinshow
3. @jsolomonReports
5. @FoxNews
15. @realDonaldTrump

Right Leaning News
1. @nypost
2. @WSJ
3. @DonaldJTrumpJr
4. @EricTrump
5. @realDonaldTrump

Extreme Bias Right News
1. @DonaldTrumpJr
2. @BreitbartNews
3. @dbongino
4. @marklevinshow
5. @realDonaldTrump
# Top influencer comparisons

## 2016

### Center News
1. @CNN
2. @thehill
3. @politico
4. @CNNPolitics
12. @DRUGE_REPORT

### Left News
1. @HuffPost
2. @TIME
3. @thedailybeast
4. @RawStory
5. @HuffPostPol

### Left Leaning News
1. @nytimes
2. @washingtonpost
3. @ABC
4. @NBCNews
25. @HillaryClinton

### Extreme Bias Left News
1. @Bipartisanism
2. @PalmerReport
3. @peterdaou
4. @crooksandliars
5. @BoldBlueWave

## 2020

### Center News
1. @CNN
2. @thehill
3. @politico
4. @CNNPolitics
12. @DRUGE_REPORT

### Left News
1. @thehill
2. @AP
3. @Reuters
4. @kylegriffin1
5. @JonLemire

### Left Leaning News
1. @nytimes
2. @washingtonpost
3. @ABC
4. @NBCNews
25. @HillaryClinton

### Extreme Bias Left News
1. @DearAuntCrabby
2. @funder
3. @ImpeachmentHour
4. @MeidasTouch
5. @TheDemCoalition
Analyzing top influencer affiliations

- As previously discussed, there are three levels of affect traditionally in polarization (elites, news, and voters)
- We look at the top 25 influencers of each news category and investigate their accounts to identify which of the three levels they affiliate with IRL
- We also further refine the “voters” level by splitting it into two subcategories:
  - Independent (Twitter users with no affiliation to media or political entities)
  - Other (users with no identification or identifying information whatsoever)
Influencer affiliation results

• Unaffiliated (independent and other) influencers are more common in the fake and extreme bias categories
• Affiliated (media and political) influencers are more common in the other news categories
• Media influencers have a greater presence in the left, left-leaning, and center news categories compared to their counterparts
Influencer affiliation results

• The number of media influencers within most of the categories decreases from 2016 to 2020
• The extreme right-bias and fake news categories, increased in media-affiliated influencers,
• The extreme right-bias category also increased in politically-linked influencers
Measuring polarization with influencer propagation

- Influencers represent politically biased modalities
- We used them to measure if there are any changes in **ideological polarization**
- Polarization increases as users segregate themselves while propagating information from influencers
- Example: users will either retweet influencers from the left or the right, with little overlapping retweeting in between
Assembling influencer propagation similarity network

• For each year, we gathered the top 100 influencers across each news category and removed all news category labels, creating two sets
• For each set we generated an undirected, complete graph where a node is an influencer, and the edges are weighted by how similar the linked nodes are in the users that retweet their content
• We ran community detection on the two networks to see how influencers are grouped by the users that retweet them
Similarity network analysis

- Two communities were found, and the news category labels were re-assigned
- One community had left-biased influencers and one had right-biased influencers
- We quantified community cohesion and separation using modularity and normalized cut

**Modularity results:**
- 0.365 (SE = 0.007) in 2016
- 0.39 (SE = 0.01) in 2020

**Normalized cut results:**
- 0.36 (SE = 0.04) in 2016
- 0.128 (SE = 0.005) in 2020

Figure ref: shorturl.at/fH148
Measuring polarization with ideological inference

- For robustness, there are other established ways to measure ideological polarization.
- We can use a method developed in [1] which infers the ideology of users from the ideological alignment of the political actors they follow.
- This produces ideological estimates of the members of the U.S. Congress that are highly correlated with ideological estimates based on roll call voting similarity such as DW-NOMINATE.

Figure ref: https://en.wikipedia.org/wiki/NOMINATE_(scaling_method)
Assembling influencer-user bipartite network

• We can use influencers as the political actors used to infer the followers’ ideology, and we can use retweets as our representation of following
• We use the same influencer sets as before, complemented with a set of all Twitter users that have retweeted at least three of these influencers
• We create an adjacency matrix where element $a_{ij}$ is the number of times user $i$ has retweeted influencer $j$
• This creates a bipartite network of users retweeting influencers that is then projected on to a one-dimensional scale using SVD decomposition
• We can distribute both influencers and users along this dimension of ideology in a histogram
Analyzing ideological distribution

- To quantify if polarization has increased between 2016 and 2020 using the ideological scale, we used Hartigan’s dip test.
- In our distribution, this indicates increasing separation from ideological consensus, resulting in polarization.

- **User results:**
  - $D = 0.11$ in 2016
  - $D = 0.15$ in 2020

- **Influencer results:**
  - $D = 0.18$ in 2016
  - $D = 0.21$ in 2020
Conclusions

- 70% of the collective top 100 influencers in 2020 were not present in 2016
- Extreme biased right and fake news categories received increasing endorsement from officially-affiliated influencers from 2016 to 2020
- But the relative volume of extremist and fake news on Twitter has decreased
- Regardless, multiple analyses confirm a robust pattern of increasing division into opposing Twitter echo chambers with growing in-group density and out-group separation
- These analyses also confirm growing ideological separation of both users and influencers on Twitter
Thank you!