The Ideological Landscape of U.S. Twitter Elites

Subhayan Mukerjee[1], Kokil Jaidka[2], and Yphtach Lelkes[1]

[1]Annenberg School for Communication, University of Pennsylvania
[2]Nanyang Technological University, Singapore
*corresponding author: Yphtach Lelkes (yphtach.lelkes@asc.upenn.edu)

Abstract

Social media platforms such as Twitter are becoming increasingly important as a political forum for campaigning, propaganda, and mobilization. As a result, there are mounting concerns that at least in the United States, the information environment on Twitter is fragmented along ideological lines with users ensconced into one of two polarized camps with limited exposure to cross-cutting views. However, there is little quantitative evidence for this claim, typically relying on methods that use small populations of political elites for assessing the extent of political polarization on the platform. This study makes several advancements over the existing body of literature by first, putting political elites in the larger context of the American Twittersphere, and second, by weighting the ideologies of Twitter elites by their tweeting activity. We find little evidence of polarization in how information is produced on Twitter at an elite level. Finally, we use a network analytic technique to assess the level of selective exposure on Twitter by building a coexposure network, and find that the Twitter consumption landscape is characterized by the presence of ideological echo-chambers.

Keywords— political ideology, polarization, social media, network analysis

1 Introduction

The extent of political polarization in the United States has long been a subject of scholarly research [1, 2, 13, 14, 15, 23]. This has been a particularly hot-button issue in recent years owing to the rise of digital media platforms, and their increasing prominence in American politics. Pew surveys have shown that not only do a majority of Americans use social media [34], but that social media platforms also increasingly mediate their access to online news [25].

This has in turn has led scholars to question the relationship between political polarization and digital platforms [35, 16], often yielding mixed or conflicting findings [19, 10, 33, 3].

Twitter has been a useful tool for researchers trying to understand political communication on digital platforms, and it has gained particular prominence in recent years owing to its increasing use as a means of communication by political leaders and media elites alike. Empirical work using Twitter data has further advanced this debate on polarization [8, 21, 5, 38, 12]. Barberá's study [4] that introduced a Bayesian ideal point technique for inferring the ideologies of Twitter users has perhaps been the most influential in this regard. In this technique Barberá identified elites apriori, based on the public offices they hold or the party tickets they run on, and estimated the latent ideal points of a Twitter account based on which of these predetermined elites they followed.

While his estimation method is widely applied and considered to be robust, he focuses on handpicked 'political' elites, whereas many people may be using twitter for reasons outside of politics. The implication is that it is not known whether Barberá's elites are important for polarization – either by being the most followed, or by being the most vocal. While it is probable that the political elites alone are driving the overall polarization in the overall Twittersphere, it is also possible that their influence matters little in the larger, noisier picture.

In other words, given that the Twittersphere is likely skewed towards discussing entertainment rather than politics, is it fair to assess the polarization of Twitter in general, by simply looking at the polarization of political elites? Moreover, what are the implications for the ordinary users who follow them – are they exposed to cross-cutting viewpoints, or are they restricted to echo chambers of singular ideologies? These questions point to empirical gaps that prevent us from getting a clearer picture of the landscape of ideological polarization among American elites on Twitter.

In this study, we outline a new approach to measure the extent of polarization on Twitter. To ensure that our approach identifies actual American Twitter elites (hereon referred to as "American elites"), we first filter the accounts followed by most American users [1], randomly selected from the Twitter 1% firehose. Next, in order to adjust for the actual production of polarized content, we weight the estimated ideology of each American elite by their tweeting frequency. We also analyze subset the American elites according to their profession and examine the ideological distribution across the different genres.

Finally, to understand the ideological fragmentation of American Twitter *audiences*, we use a network analytic technique to identify the level of ideological sorting in so far as Twitter following relationships are concerned. This helps us better understand the extent to which the Twitter consumption landscape is characterized by selective exposure in how ordinary Americans follow elites, and the existence of potential echo-chambers.

[1]identified as Twitter users who geo-tagged their Tweets with a location in the United States following a method proposed in [31]

2 Measuring Political Polarization on Twitter

The question 'Is Twitter polarized?' has been proposed ever since the introduction of Twitter as a platform for political discourse and a tool for communication research. While Barberá's ideal-point estimation technique [4] remains the most popular in estimating ideologies of its users and assessing the extent of political polarization on Twitter [4, 6], there have been several other attempts that seek to do the same. Some of the earlier studies that tried to answer this question were undertaken by computer scientists. [11] for instance, used the structure of Twitter conversations to understand whether polarization manifested itself in who people spoke to. They ran a community detection algorithm on an American Twitter conversation network and found that two distinct communities emerged - which they identified as Democratic and Republican respectively. [9] found that one could reliably predict the party a British Twitter user supported based on how many times they mentioned the British political parties in their tweets. Other studies have demonstrated the utility of machine learning techniques that use the network properties of a user's Twitter network as input features in inferring their political ideologies [28, 40].

Elsewhere, there have been studies that have sought to measure and quantify polarization, albeit around very specific issues. The study by [39] explored the divide between pro-life and pro-choice advocates following the shooting of a late-term abortion doctor by analyzing a Twitter conversation network. In another study of Twitter in Egypt, researchers characterized secular versus Islamist polarization based on the retweet network [37]. The study by [22] demonstrated the use of linguistic features of a Twitter conversation centered around French marriage reform for exploring polarization. A similar approach was followed to study the Twitter conversation about the late Venezuelan president Hugo Chávez [26] and identify the emergence of polarization in interactions. Finally, the authors of [18] used network analysis of Twitter conversation networks to find evidence of clustering around shared political views during the 2011 Canadian Federal Election.

3 Data and Methods

3.1 Data collection

Our first improvement over existing studies that look at elite polarization on Twitter begins with our choice of elites. We do not decide on the elites based on who they are, or what position they occupy in real-life. Instead, we begin with a sample of Tweets from a three-month time period (January - March 2019) from the Twitter 1% firehose. Next, we filtered the tweets to retain only those which were geo-tagged with a location in the United States, or were posted by users who reported their location as a place in the United States, following a set of heuristics established in previous work [31]. We then sampled 10,000 Twitter accounts and discarded any accounts which were no longer available (either suspended, deleted, or changed to a private setting). Finally, we were left with 9959 Twitter accounts, henceforth known as 'ordinary American Twitter users', who followed a total of 393,919 unique accounts.

3.1.1 Identifying American elites

We divided the sample into ten random sub-samples, and obtained the thousand most followed handles that each sub-sample of ordinary Twitter users followed. We denoted as 'American elites' the union of these ten sets of thousand handles. In other words, if any Twitter account was among the top thousand most followed accounts in any of the ten sub-samples of ordinary American Twitter users, we deemed it to be an elite. The final number of elites that we ended up with was 1822.

3.1.2 Assigning weighted ideologies to American elites

After finalizing the set of elites, we used the aforementioned ideal-point estimation algorithm of Barberá [4] to infer their ideologies. This method works under the assumption that "Twitter users prefer to follow political actors (politicians whose position on the latent ideological dimension are similar to their" because decisions to follow are "costly signals about users' perceptions of both their ideological location and that of political account." In other words, this method works by first identifying Twitter handles of well known liberal and conservative actors at two ends of the latent ideological spectrum, and then interpolating the latent ideal points of a random user based on which of these known actors they follow.

To then characterize the extent to which information production at the elite level is polarized on Twitter, we weight the inferred ideology of each elite account with their normalized tweeting frequency, and observe the distribution of the weighted ideologies. This captures the polarization of information production on Twitter better than previous attempts that use unweighted distributions since it takes into consideration one, the actual Twitter elites, who are the most followed by our samples of American Twitter users, and two, the rate at which ideologically slanted tweets are created.

3.2 Analysis

3.2.1 Ideological distribution by activity

To account for the long tail effects of inactive social media users, we further unpack elite polarization by examining the ideological distribution of elites within each decile of tweeting frequency. This analysis helps us understand the relationship between polarization and "loudness". In other words, are the elites who are the loudest (and tweet the most), the ones who drive the polarization? We then manually code each elite as entertainment personality, public figure, political figure, sports personality, brands, and so on and observe the ideal point distribution within each sector to compare the variation in elite polarization between these sectors.

3.2.2 Ideological sorting in the Twitter Co-Exposure Network

Finally, we use a network analytic technique to assess the extent to which the Twitter landscape is characterized by the existence of ideologically sorted communities of information exposure. To do this, we build an audience overlap network of the elites following a method developed to understand audience fragmentation in news consumption [27], that has since seen application in the context of Twitter [36, 24]. To do this, we first create a bi-partite 'following network' that captures all the follower-following relationships between the elites and the ordinary American users (the non-elites) in our sample. Next we extract the elite-projection of this bi-partite network. This gives us the audience overlap network (also called a coexposure network; see [17]), where each node is an elite, and the weight of edge between every pair of nodes captures the number of shared followers. The mechanism of network construction is demonstrated in Figure 1.

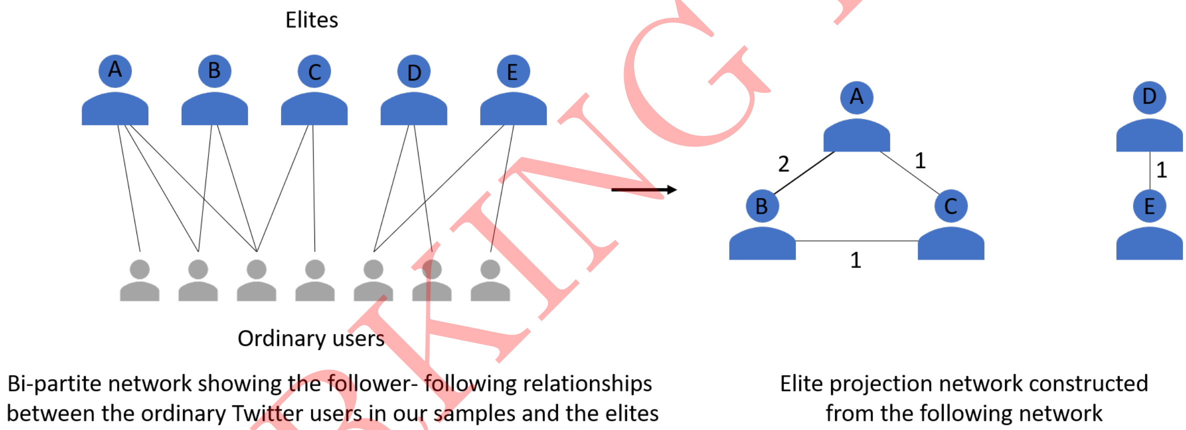

Figure 1: The Logic of Network Construction

In this elite network, we use a community detection algorithm to find groups of elites that are characterized by heavy overlap within themselves than without. This algorithm captures the manner in which the ordinary users in our samples distribute their attention across these groups of elites. The existence of ideologically sorted communities of elites would point to a polarized consumption landscape since it means that a substantial group of ordinary users follow elites who are ideologically similar. In other words, this analysis would help characterize the degree to which ordinary users demonstrate selective exposure to partisan information on Twitter.

4 Results

4.1 Overall Distributions

The twenty most popular elites as followed by our sample of American users are listed in Table 1.

Rank	Elite Twitter Handle	Elite Name
1	BarackObama	Barack Obama
2	RealDonaldTrump	Donald Trump
3	TheEllenShow	Ellen DeGeneres
4	Drake	Drake
5	KanyeWest	Kanye West
6	POTUS	President Trump
7	Rihanna	Rihanna
8	JColeNC	J. Cole
9	POTUS44	President Obama
10	SportsCenter	SportsCenter
11	KingJames	LeBron James
12	JimmyFallon	Jimmy Fallon
13	NYTimes	The New York Times
14	HillaryClinton	Hillary Clinton
15	AOC	Alexandria Ocasio-Cortez
16	ArianaGrande	Ariana Grande
17	CNNBRK	CNN Breaking News
18	ESPN	ESPN
19	chancetherapper	Chance The Rapper
20	MichelleObama	Michelle Obama

Table 1. The list of the 20 elites that are most followed by our sample of ordinary Americans

This list includes popular politicians, mainstream media outlets, and sports and TV celebrities. Of the 1822 elites that we obtained, 22 of those no longer had functional Twitter accounts. We applied Barberá's ideal-point estimation technique to infer the ideologies on the remaining 1800. Since 590 of these 1800 did not follow any political actors (that Barberá's technique requires), they were therefore coded as 'Neutral' with an ideal-point of 0. We repeated the exercise with our original sample of ordinary Twitter users (N = 9959), and found that 5976 (i.e. 60% of them) did not follow any political actors, and had to therefore be coded as 'Neutral' as well. Figure 2 shows the distribution of the ideal points of both elites and ordinary Twitter users.

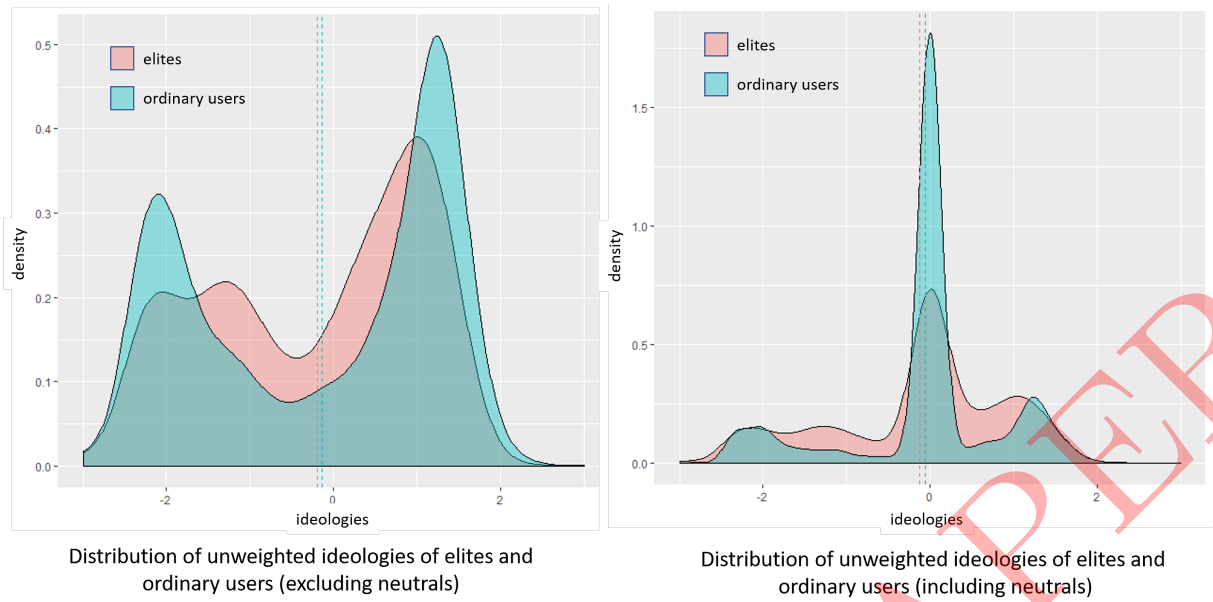

Figure 2: Ideological distribution of elites and ordinary Twitter users

4.2 Ideological Distributions by Tweeting Activity

Next we weight the elites' ideal points by their tweeting activity. Figure 3 compares the distribution of the elites' unweighted ideal points with their weighted equivalent. This suggests that the apparent ideological polarization hitherto found among the elites disappears when their tweeting activities are accounted for.

To further unpack the relationship between polarization and tweeting activity, we plot the distribution of ideal points for elites within each decile of tweeting activity (Figure 4). In other words, we divide the set of elites into ten roughly equal sized groups after sorting them by tweeting activity. The first decile has the elites who tweeted the least, while the tenth decile has the elites who tweeted the most. By plotting and comparing the ideal point distributions we find that while the density of liberals roughly remains constant at all levels of tweeting activity, the density of conservatives gradually increases. In the tenth (or the most active) decile, conservatives dominate the distribution, rising above even the neutrals.

4.3 Ideological Distributions by Elite Genre

Our next analysis seeks to compare different genres of elites. To do this, we manually coded each of the 1800 elites into one (or more) of ten genres. The genres were as follows: "hard news", "media outlets", "political figures", "political pundits", "public figures", "entertainment", "sports", "brands", "organizations", and "meme accounts." Their relative frequencies are shown in Figure 5. Since an elite can belong to more than one genres, the number of elites in each genre do not add up to the total number of elites in our set. The entertainment genre is

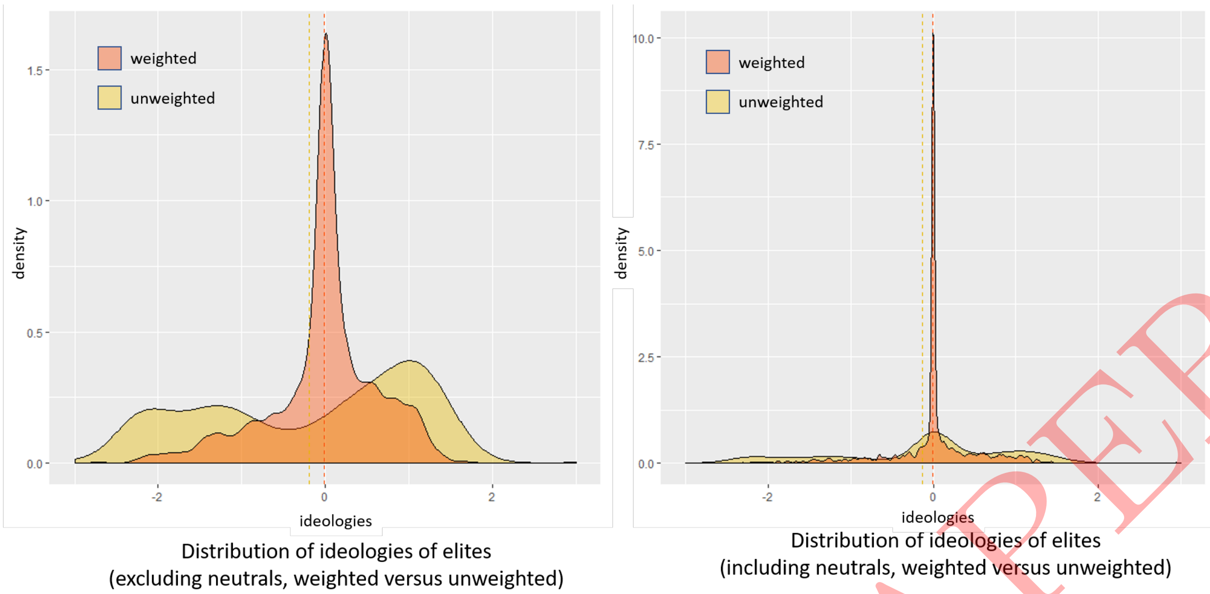

Figure 3: Ideological distribution of elites: weighted versus unweighted

by far the most popular, with sports and public figures at distant second and third places. The prominence of non-political actors in the set of elites is possibly the reason behind the apparent lack of overall polarization that we find. This also shows that being so relatively few in number, prior studies, by just focusing on political elites had potentially overestimated the true extent of political polarization on the platform.

In Figure 6 we show the (weighted) ideal point distribution for each genre of elites. While most of the categories (particularly the popular ones like entertainment and sports) do not exhibit polarization, media outlet accounts show a distinct bi-modal distribution. This seems to suggest that there are relatively more media outlets on the conservative end of the latent ideological spectrum than there accounts are in the overall distribution. Similarly, the distributions of both, hard news accounts and political pundits show a fat tail to the right, signalling the presence of relatively more conservative accounts in these genre than in the overall distribution.

4.4 Analysis of Coexposure Twitter Network

For our final analysis we build a coexposure network (of audience overlap) as described in the Methods section. The nodes in this network correspond to our elites, and the edges between them encode the number of shared followers they have. Before we analyze the network we follow previous studies [30, 27] in first considering the presence of statistical noise that can potentially confound our findings. We remove noisy edges by a method known as dyadic thresholding. This technique uses the traditional ϕ coefficient to estimate the extent to which the weight of an edge is greater than what is expected due to chance, before proceeding to

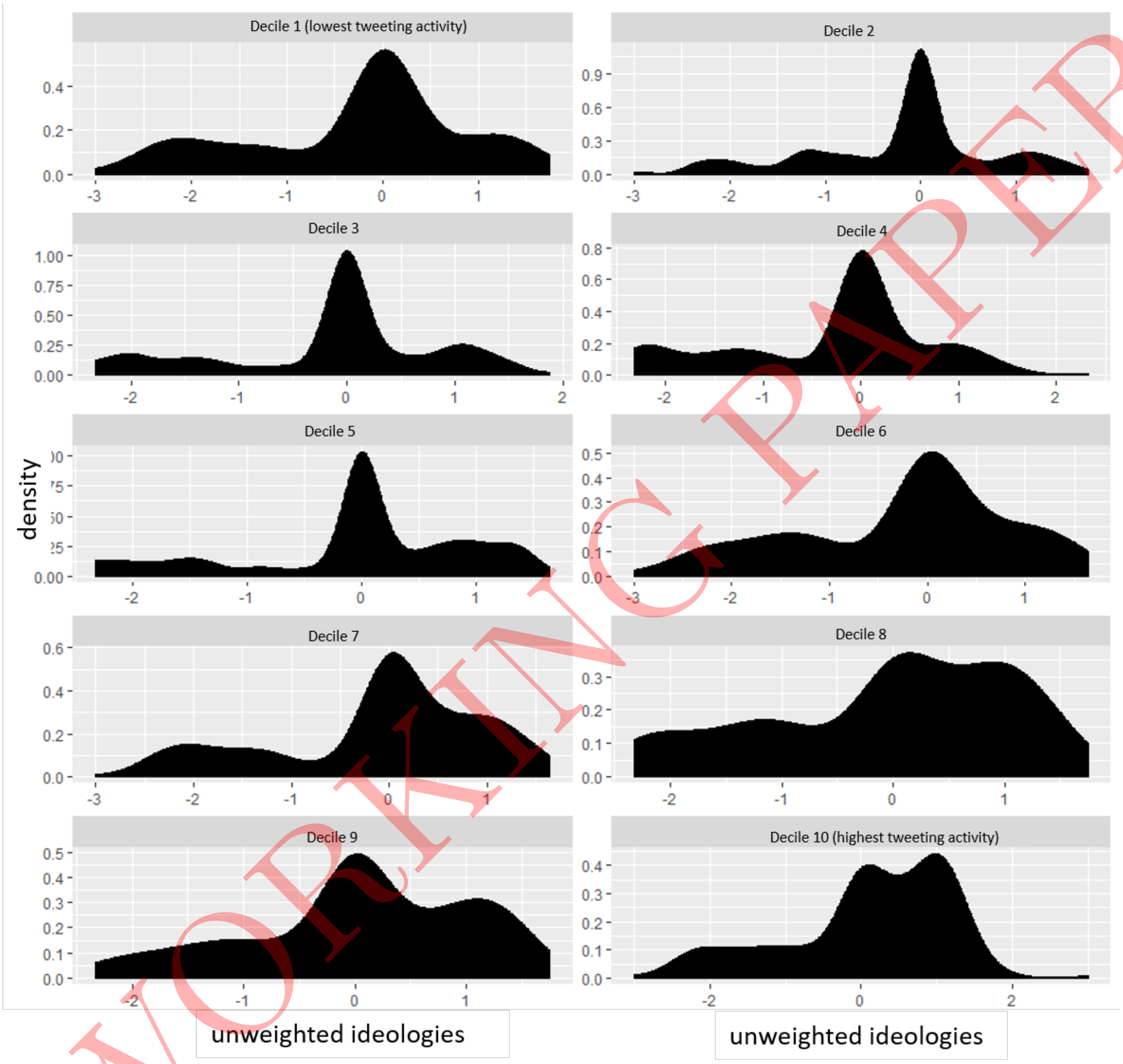

Figure 4: Ideological distribution of elites by decile of tweeting frequency

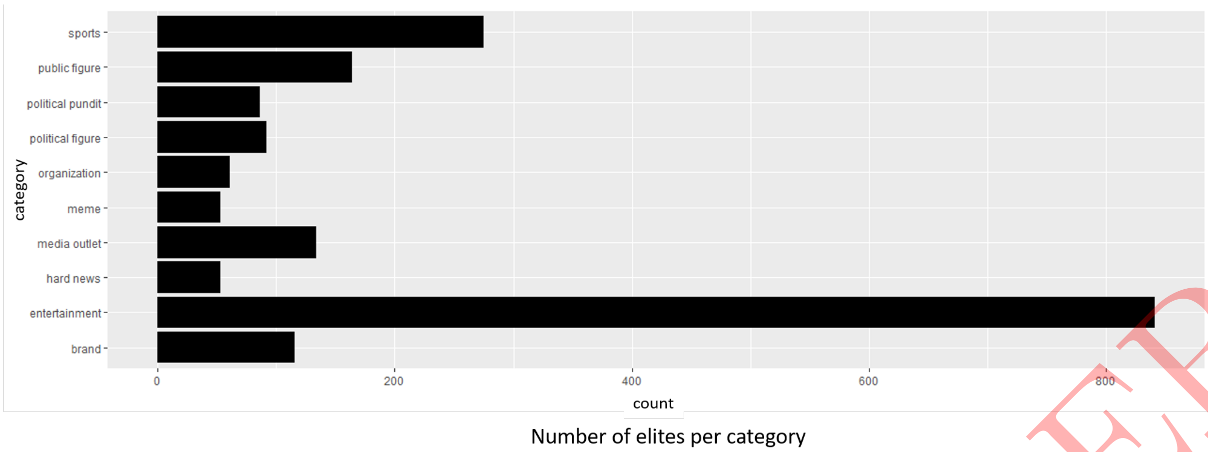

Figure 5: Elite count by category

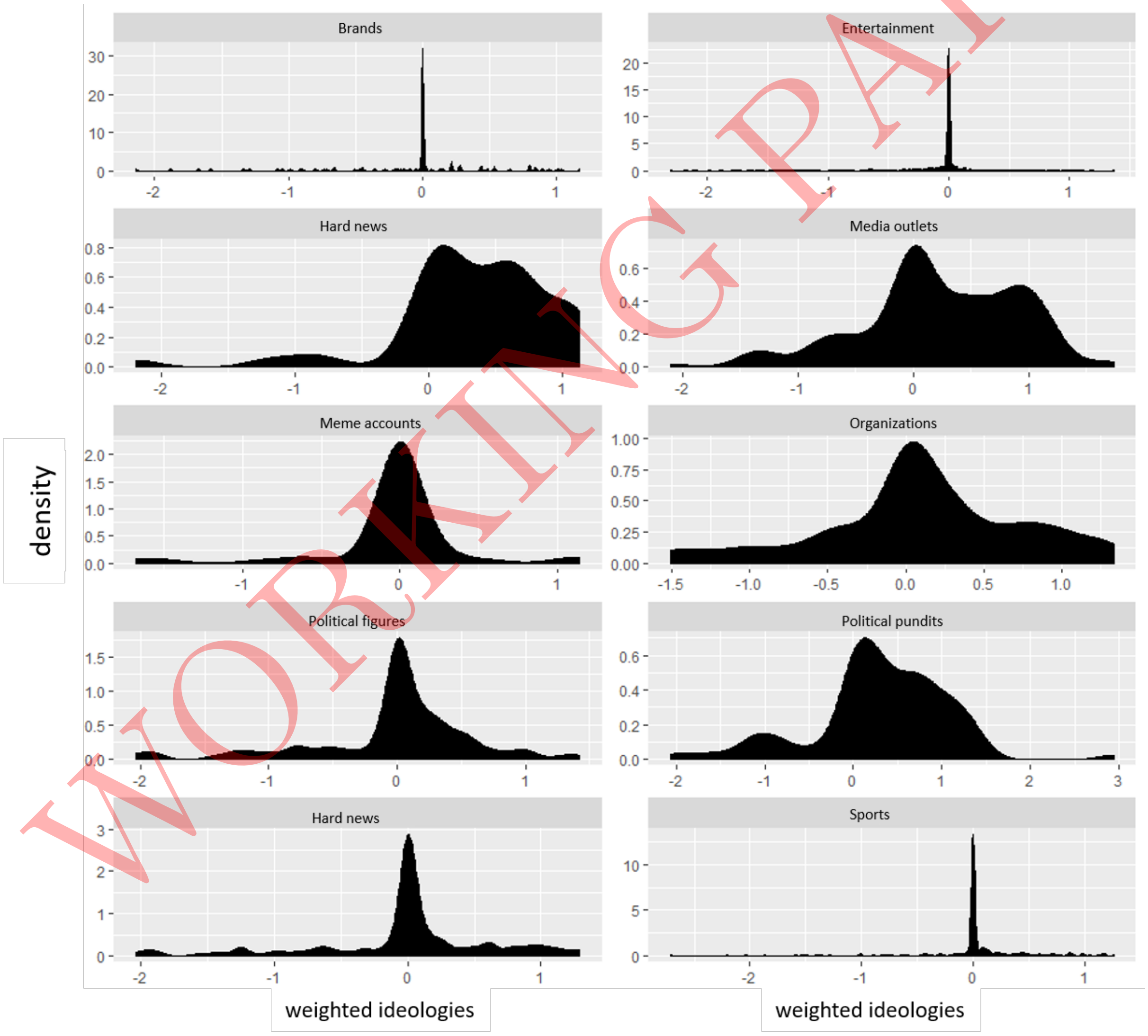

Figure 6: Ideological distribution of elites by category

eliminate those edges that do not meet the required criteria for significance as assessed through a t-value. We use a strict threshold, eliminating all ties with $|t| > 2.58$ (which corresponds to a p-value of < 0.01). We then run a WalkTrap community detection algorithm on this filtered network. The WalkTrap algorithm [29] operates under the intuition that random walks along the network edges over a sufficiently long period of time will tend to get trapped within groups of nodes that share heavy overlap between themselves than with those they don't.

The community detection algorithm returns a partitioning of the network into 5 communities. The network properties for the 5 community induced subgraphs are tabulted in Table 2.

Community	Size	Density	Degree Centralization
1	552	0.443	0.459
2	512	0.605	0.348
3	413	0.880	0.116
4	68	0.994	0.001
5	255	0.755	0.245

Table 2. Network properties of the communities.

The ideal-point distributions (both unweighted and weighted by tweeting activity) of the elites in each of these communities is shown in Figure 7.

While the unweighted distribution does show the communities being distinct in so far as the ideal points of the elites in them are concerned, the weighted distribution collapses them and makes them nearly indistinguishable. However, two-sided t-tests of the means of the ideal points reveal a clear ideological divide, as depicted in Figure 8.

Community 1 (unweighted μ = -0.24, 95% CI: [-0.31, -0.17]; weighted μ = -0.04, 95% CI: [-0.07, -0.01]) and community 2 (unweighted μ = -0.35, 95% CI: [-0.44, -0.25]; weighted μ = -0.05, 95% CI: [-0.09, -0.01]) have a liberal slant. Community 4 has a strong conservative slant (unweighted μ = 1.41, 95% CI: [1.31, 1.51]; weighted μ = 0.64, 95% CI = [0.51, 0.79]). Communities 3 (unweighted μ = -0.86, 95% CI: [-0.20, -0.03]; weighted μ = -0.03, 95% CI = [-0.10, 0.04]) and 5 (unweighted μ = -0.06, 95% CI: [-0.05, .17]; weighted μ = 0.02, 95% CI = [-0.04, 0.08]) have means that are not significantly different from 0, and can thus be classified as neutral.

These findings suggest some evidence of ideological fragmentation in how ordinary Americans divide their attention among elites on Twitter. The presence of ideologically sorted communities, even after factoring in tweeting activity, lends credence to the fact that information consumption on Twitter is characterized by selective exposure, in so far as following decisions of ordinary Americans are concerned.

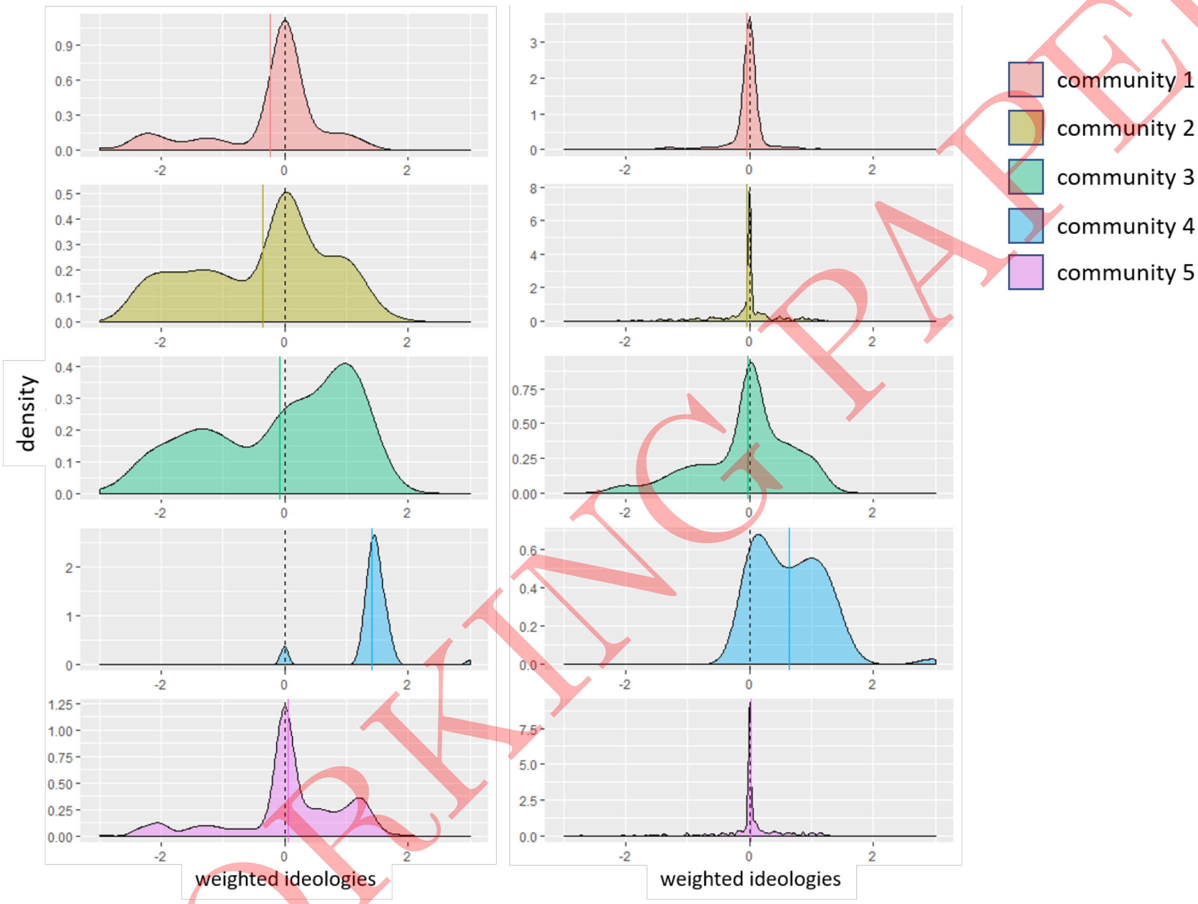

Figure 7: Ideological distribution of elites by community

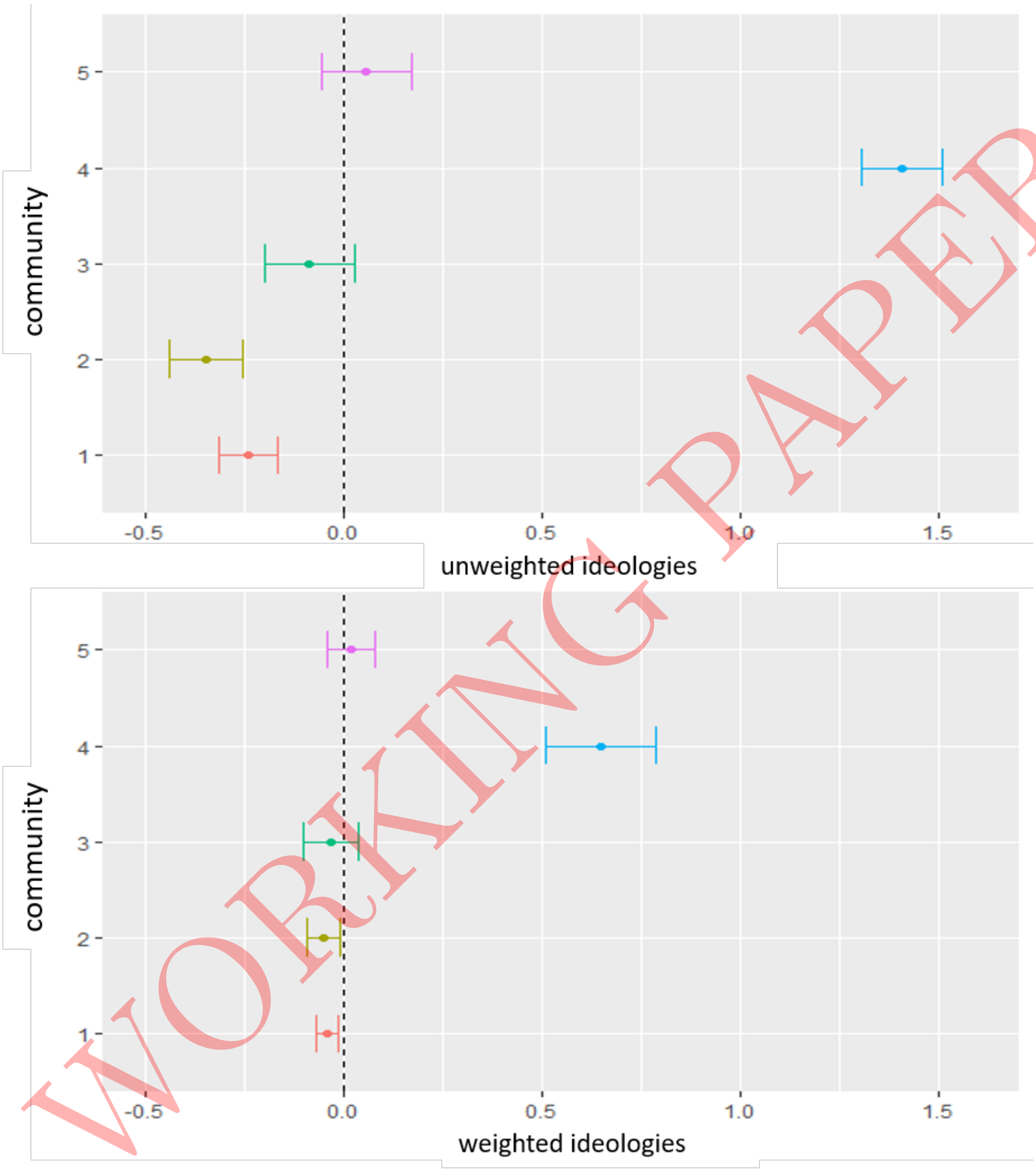

Figure 8: Means and 95% confidence-intervals of means of ideologies by communities

5 Discussion

Many theories of democracy consider the need for a common space for deliberation by an informed citizenry to be of paramount importance for the functioning of the democratic process [7]. The potential of social media platforms to provide this common space needed for public deliberation has long provided fodder for speculation among scholars - lending itself to the concept of the Habermasian ideal of the public sphere [20]. Many studies that seek to understand the extent of political polarization on Twitter are often motivated by this idea: is Twitter, by virtue of being increasingly prominent in our political discourse, conducive to the open exchange of ideas and political viewpoints that characterizes a public sphere? Or is the information environment on Twitter largely polarized, and characterized by echo-chambers instead? This study sheds light on the extent of political polarization on Twitter in the US by making several advancements over the existing literature.

First, we use a set of elites that is informed by an empirical distribution of the number of followers they have in our random sample of American Twitter users, instead of choosing them based on extraneous attributes.

Second, it demonstrates the extent to which previous studies have overstated the level of political polarization on Twitter by focusing solely on political actors. Political figures and political pundits comprise less than 10% of the elites we find to be the most popular as indicated by who our sample of ordinary Americans follow. The percentage rises to less than 20% when combined with hard news and media outlets as Figure 5 demonstrates. Entertainment (44.9%) is by far the most popular category, with sports personalities (14.7%) and public figures (9.74%) following far behind. These findings further call to question the reliability of self-reports used in surveys which find that a majority of American adults who use Twitter get news on Twitter as well [32]. While it is true that the Twitter algorithm can show users tweets from accounts they don't follow, popular interest in politics, as measured by people's decisions to follow hard news and political accounts is very low.

Third, our study makes the distinction between the ideological distribution of elites and the distribution of actual information production on Twitter. An information environment like Twitter is polarized to the extent to which the information on it is polarized. The presence of a number of ideologically polarized but rarely active elites would not make Twitter a polarized platform. By weighting the estimated ideologies of the Twitter elites by their tweeting activity, we find a more pointed unimodal distribution, implying that information production on Twitter is even less polarized than what the presence of ideological elites would imply. Future studies in this area could employ longitudinal designs to identify any temporal trends in these findings.

Finally, we find that while information production may not be polarized, the same cannot necessarily be said of information exposure. Our network analysis reveals a an ideologically sorted community substructure of elites, that is characterized by ordinary Americans distributing their attention to elites in a non-random manner. Even when weighted, the communities appear ideologically divided in so far as their mean ideal points are considered, though the ideological distances between them are greatly reduced. Even though four of the five communities share considerable distributional overlap of weighted ideologies with each other, community 4 still remains considerably distinct. This affords substantial evidence regarding the presence of a conservative echo-chamber dominated by political figures and political pundits on American Twitter. This raises potential concerns regarding the health of political discourse on the platform, echoing some of the prior anxieties that scholars have expressed over the shrinking common-ground that the ideal of a deliberative democracy requires [35]. Future studies can seek to characterize the political discourse in these communities, to try and understand whether there is any fragmentation in agenda or topics in what these elites discuss.

References

[1] A. I. Abramowitz and K. Saunders. Why Can't We All Just Get Along? The Reality of a Polarized America. *The Forum*, 3(2), 2005.

[2] A. I. Abramowitz and K. L. Saunders. Is Polarization a Myth? *The Journal of Politics*, 70(2):542–555, 2008.

[3] C. Bail, L. Argyle, T. Brown, J. Bumpus, H. Chen, M. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky. Exposure to opposing views can increase political polarization: evidence from a large-scale field experiment on social media. *Proceedings of the National Academy of Sciences*, 115(37):1–6, 2018.

[4] P. Barbera. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, 23(1):76–91, 2015.

[5] P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, and R. Bonneau. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10):1531–1542, 2015.

[6] P. Barberá and G. Rivero. Understanding the Political Representativeness of Twitter Users. *Social Science Computer Review*, 33(6), 2015.

[7] J. Bohman and W. Rehg, editors. *Deliberative democracy : essays on reason and politics.* MIT Press, 1997.

[8] M. Bossetta, A. D. Segesten, and H.-J. Trenz. Political participation on Facebook during Brexit: Does user engagement on media pages stimulate engagement with campaigns? *Journal of Language & Politics*, 17(2):173–194, 2018.

[9] A. Boutet, H. Kim, and E. Yoneki. What's in Your Tweets? I Know Who You Supported in the UK 2010 General Election. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

[10] L. Boxell, M. Gentzkow, and J. M. Shapiro. Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617, 2017.

[11] M. Conover, J. Ratkiewicz, and M. Francisco. Political polarization on twitter. In *ICWSM 2011 International Conference on Weblogs and Social Media*, 2011.

[12] H. K. Evans, V. Cordova, and S. Sipole. Twitter Style: An Analysis of How House Candidates Used Twitter in Their 2012 Campaigns. *PS: Political Science & Politics*, 47(02):454–462, 4 2014.

[13] M. P. Fiorina, S. A. Abrams, and J. C. Pope. Polarization in the American public: Misconceptions and misreadings. *Journal of Politics*, 70(2):556–560, 2008.

[14] M. P. Fiorina and S. J. Abrams. Political Polarization in the American Public. *Annual Review of Political Science*, 11:563–588, 2008.

[15] M. P. Fiorina and M. S. Levendusky. Disconnected: The Political Class Versus the People. In P. S. Nivola and D. W. Brady, editors, *Red and Blue Nation? Volume I: Characteristics and Causes of America's Polarized Politics*, pages 49–71. Brookings Institution Press, Washington, DC, 2006.

[16] S. Flaxman, S. Goel, and J. M. Rao. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(Specialissue1):298–320, 2016.

[17] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, N. Grinberg, and D. Lazer. Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425):374–378, 2019.

[18] A. Gruzd and J. Roy. Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy & Internet*, 6(1):28–45, 3 2014.

[19] A. Guess, B. Lyons, B. Nyhan, and J. Reifler. Avoiding the Echo Chamber about Echo Chambers: Why selective exposure to like-minded political news is less prevalent than you think. 2018.

[20] J. Habermas. *The Structural Transformation of the Public Sphere*, volume 9. print. 1989.

[21] Y. Halberstam and B. Knight. Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*, 143:73–88, 2016.

[22] M. Lai, C. Bosco, V. Patti, and D. Virone. Debate on political reforms in Twitter: A hashtag-driven analysis of political polarization. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 1–9. IEEE, 10 2015.

[23] Y. Lelkes. Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, 80(Specialissue1):392–410, 2016.

[24] F. Lumban Gaol, T. Matsuo, and A. Maulana. Network Model for Online News Media Landscape in Twitter. *Information*, 10(9):277, 9 2019.

[25] K. E. Matsa and E. Shearer. News Use Across Social Media Platforms 2018 | Pew Research Center. Technical report, Pew Research Center, 2018.

[26] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito. Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(3):033114, 3 2015.

[27] S. Mukerjee, S. Majó-Vázquez, and S. González-Bailón. Networks of Audience Overlap in the Consumption of Digital news. *Journal of Communication*, 68(1):26–50, 2018.

[28] M. Pennacchiotti and A.-M. Popescu. A Machine Learning Approach to Twitter User Classification. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2011.

[29] P. Pons and M. Latapy. Computing Communities in Large Networks Using Random Walks Pascal. *Journal of Graph Algorithms and Applications*, 10(2):191–218, 2006.

[30] S. Ronen, B. Gonçalves, K. Z. Hu, A. Vespignani, S. Pinker, and C. A. Hidalgo. Links that speak: The global language network and its association with global fame. *Proceedings of the National Academy of Sciences*, 111(52):E5616–E5622, 2014.

[31] A. H. Schwartz, J. C. Eichstaedt, M. L. Kern, L. Dziurzynski, M. Agrawal, G. J. Park, L. S. K., S. Jha, M. E. P. Seligman, and L. Ungar. Characterizing Geographic Variation in Well-Being using Tweets. *Seventh International AAAI Conference on Weblogs and Social Media*, (June), 2013.

[32] E. Shearer and E. Grieco. Americans Are Wary of the Role Social Media Sites Play in Delivering the News. Technical report, Pew Research Center, 2019.

[33] A. Siegel, J. A. Tucker, J. Nagler, and R. Bonneau. Tweeting Beyond Tahrir: Ideological Diversity and Political Intolerance in Egyptian Twitter Networks. *Unpublished Working Paper*, 2018.

[34] A. Smith and M. Anderson. Social Media Use 2018. Technical report, Pew Research Center, 2018.

[35] C. R. Sunstein. *#Republic : divided democracy in the age of social media*. Princeton University Press, 2017.

[36] A. Urman. Context matters : political polarization on Twitter from a comparative perspective. *Media, Culture & Society*, 2019.

[37] I. Weber, V. R. K. Garimella, and A. Batayneh. Secular vs. Islamist polarization in Egypt on Twitter. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM '13*, pages 290–297, New York, New York, USA, 2013. ACM Press.

[38] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang. Quantifying Political Leaning from Tweets, Retweets, and Retweeters. *IEEE Transactions on Knowledge and Data Engineering*, 28(8):2158–2172, 8 2016.

[39] S. Yardi and D. Boyd. Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter. *Bulletin of Science, Technology & Society*, 30(5):316–327, 10 2010.

[40] F. A. Zamal, W. Liu, and D. Ruths. Homophily and Latent Attribute Inference : Inferring Latent Attributes of Twitter Users from Neighbors. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2012.