Cultural Fault Lines and Political Polarization

Yongren Shi
Yale University
New Haven, Connecticut
yongren.shi@yale.edu

Kai Mast
Cornell University
Ithaca, New York
kaimast@cs.cornell.edu

Ingmar Weber
Qatar Computing Research Institute, HBKU
Doha, Qatar
iweber@hbku.edu.qa

Agrippa Kellum
Cornell University
Ithaca, New York
ask252@cornell.edu

Michael Macy
Cornell University
Ithaca, New York
mwmacy@cornell.edu

ABSTRACT

Survey research reveals deep partisan divisions in the U.S. that extend beyond politics to include cultural tastes, lifestyle choices, and consumer preferences. We show how co-following on Twitter can be used to measure the extent to which these divisions are also evident in social media. We measure political alignment (location on the red-blue spectrum), relevance (overlap between cultural and political interests), and polarization (internal division) in music, movies, hobbies, sports, vehicles, food and drink, technology, universities, religions, and business. The results provide compelling evidence that "Tesla liberals" and "bird hunting conservatives" are stereotypes grounded in empirical reality.

CCS CONCEPTS

•Information systems → Online advertising; •Networks → Online social networks; •Applied computing → Sociology;

KEYWORDS

Polarization, Culture, Social Media, Networks

1 INTRODUCTION

Extensive research has revealed deep cultural fault lines in America, not only aligned with politically salient issues, but also geographic regions, class, gender, race, ethnicity, and education [1, 3, 11, 12, 16]. Liberal and conservative ideology has also been shown to align with moral hot buttons, lifestyle choices, and consumer preferences. In his popular book *The Big Sort*, Bishop [5] shows a polarized nation that is divided into red and blue regions not only by voting behavior but also by lifestyle profiles of "Books, Beers, Bikes, and Birkenstocks" (p. 196). A sweeping analysis of opinion items in the General Social Survey cumulative file (1968-2010) shows a consistent pattern of association between political opinions and cultural tastes [10].

WebSci'17, June 25-28, 2017, Troy, NY, USA.

978-1-4503-4896-6/17/06…$15.00
DOI: http://dx.doi.org/10.1145/3091478.3091520

Much of the previous research on cultural polarization has relied on opinion surveys, which measure retrospective responses to items that interest researchers though not necessarily the respondents. More importantly, recent studies reveal a rising rate of nonresponse in landline-administered surveys, which raises concerns about the external validity of survey methodology. Growing availability of data from social media opens up unprecedented opportunities to use digital traces of human communications to record real-time behavioral indicators of salient preferences [15]. Recent examples include the study of partisan preferences for popular science [21], linguistic usage [17], and TV shows (http://nyti.ms/2ize1lN). People follow, like, comment, reply, purchase, listen to or watch celebrities, organizations, topics, issues, news items, events, books, bands, and videos on myriad social media platforms and in real time. These data capture preferences that may not be detected in survey responses. For example, respondents are less likely to reveal to researchers their "politically incorrect" attitudes that are often criticized by mainstream media. The failure of polls to predict the Brexit vote in Britain or the 2016 Presidential election in the U.S. have been attributed to this "shy voter" problem (http://wapo.st/2ik1EJO), but this explanation has also been challenged (http://53eig.ht/2jbciXC).

This study examines data from one of the largest social media sites, Twitter, to identity lifestyle preferences that reinforce or bridge political divisions in the U.S. Our study proposes a novel method based on co-following graphs to impute political interest in lifestyle accounts using the followers of the U.S. Congress. In contrast to most studies [2, 4, 6, 8] that measured polarization as bi-modality in the distribution of opinion, this study measures polarization as the clustering of the population into opposing camps, each of whose members agree internally across multiple dimensions of the cultural landscape, such that one can predict how someone will vote by knowing their favorite TV show or hot beverage – a pattern also known in political science as "constraint" [3, 9] that creates partisan echo chambers [10]. Co-following graph, measuring alignment between political and cultural issues based on a common set of followers, extends greatly the current research on political polarization in that it allows to measure the political leaning of Twitter accounts that have no direct linkage with politicians. For instance, Pepsi's Twitter account would avoid to be associated with any Congress member, while it doesn't prevent ordinary users from following both Pepsi and members in Congress. Note that a conservative political alignment, e.g. country music, does not mean that the musicians are conservative (although some of them no doubt are). We mean that their social media followers are conservative. We also measure political relevance (overlapping interest in culture and politics) and political alignment (relative interest in the topic by followers of liberal versus conservative Congress members).

2 DATA AND METHODS

The Twitter dataset was created by identifying 553 members of the 111th, 112th, and 113th Congresses who have at least 1000 followers on Twitter. (Below 1000 followers, observations are too sparse for reliable measurements of co-following association between two accounts, as the expected size of co-followers would approach zero). We then collected all the followers of 553 Congress members as well as the Twitter accounts associated with popular culture, personalities and organizations, including pundits, musicians, Fortune 100 companies, restaurants, sports teams, Billboard 100, and religions. The co-follower graph [13, 22] is based on the assumption that each user who follows a popular lifestyle account (e.g. a country music star) and a conservative Congress member is voting with their keyboard for the conservative alignment of a musical genre.

2.1 Political Alignment Index

Political alignment is measured in two steps, which can be illustrated using a Twitter example. First, we calculate the *lift* in the probability that a random member c of the entire set of co-followers in our dataset will follow both a particular politician (a Congress member) j and a particular musician i, compared to the probability that c will follow either j or i. Second, having calculated lift for all the politicians that are co-followed with i, we use lift as a weight to compute the weighted mean of the political scores (S_j) across all the politicians. The mean-centered result is i's political alignment score. Lift can also be interpreted as the ratio between the observed and expected overlap of i and j. The expected value controls for the differences in popularity. For instance, Taylor Swift has many more followers on Twitter than Rodriquez, so she would have a larger overlap with politicians. By adjusting the observed overlap with the expected, we eliminate the bias that is caused by popularity.

$$L_{ij} = \frac{P(i \cap j)}{P(i)P(j)} \tag{1}$$

where $P(i \cap j)$ indicates how likely a randomly chosen user shares interests in both the cultural (i) and political (j) categories. $P(i)P(j)$ indicates the expected probability of intersection between these two categories, given the size of each. In step 2, we use L as a weight when averaging the political scores of all the political categories (Congress members in Twitter data):

$$S_i = \frac{\sum_j L_{ij} \times S_j}{\sum_i L_{ij}} - \bar{S}_J \tag{2}$$

S_i is the mean-centered weighted mean of political scores for all j political categories. We subtract the simple mean ($\bar{S}_J$) to adjust for uneven partisan distribution (which is especially problematic on Twitter where Republican Congress members outnumber Democrats). S_j is derived from Poole and Rosenthal's [18, 19] DW-Nominate (DWN) scores, a standard and reliable measurement of political ideology imputed from the roll call voting records in the U.S. Congress. The DWN scores range from -1 to 1, indicating most liberal and most conservative leaning in the voting behavior in the Congress. For Twitter, we simply assign the DWN score for the Congress member to the political category j for which that Congress member is the unique member.

2.2 Political Relevance Index

Political relevance is measured by the fraction of followers of all accounts in that cultural dimension who also follow members of Congress (whether liberal or conservative). The higher the political relevance, the greater the likelihood that followers of Congress will follow accounts in the cultural topic. Two factors contribute to the political relevance of a Twitter account, the size of politically interested followers and the popularity of the account. Extremely popular cultures, e.g. Billboard 100 musician accounts, are least politically relevant.

2.3 Political Polarization Index

Political polarization of a category, e.g. religion, is the extent to which partisans follow different entities. Polarization measure draws an important distinction from the alignment measure, in that political alignment reports the average ideological leaning of all the users in a category, while political polarization reveals a possible internal divide even when the average alignment is in the middle of the political spectrum. We measure the polarization of a category using kurtosis, the fourth standardized moment of a distribution.

$$k = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2} - 3 \tag{3}$$

where μ is the mean of the distribution. If the distribution of political alignment scores in one category is peaked (indicating a high level of integration and consensus of ordinary users' choices), kurtosis is positive. If it is flatter than the normal distribution, kurtosis is negative. If the distribution reaches bimodality (roughly half of the accounts are co-followed with Republican Congress members, and the other half co-followed with Democratic Congress Members), kurtosis approaches -2. The kurtosis was used to measure attitude polarization in [11].

3 POLITICAL ALIGNMENT

3.1 Congress

We first impute various culture domains' political alignment and validated our methodology by comparing imputed political alignment with the ground truth obtained from external sources. These included Congress members, pundits, musicians who endorsed candidates, restaurants whose customers were surveyed, and Fortune 100 companies who made political contributions.

For Congress members, we compared the observed DWN score of each member with the imputed score using data from the other 552 members. The results in Figure 1 show a near-perfect correlation (r=0.955). The correlation increases with recency of the Congress, indicating that the following behavior on Twitter may have become more politically salient over time (and slightly less influenced by celebrity), in line with observations in a long-term study [14].

We repeated the procedure for celebrities and companies with known political affiliations. Table 1 shows near-perfect correlations for Congress members and political pundits and highly significant

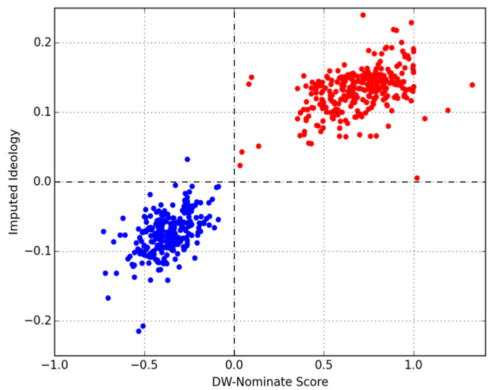

Figure 1: Scatterplot of the imputed ideology and DW-Nominate scores for 553 Congress Members.

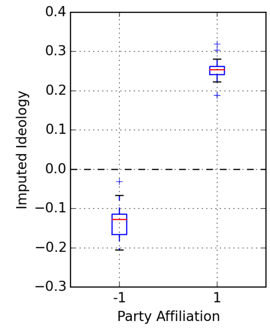

Figure 2: Comparison of the imputed ideology with self-identified party affiliations of 35 political pundits.

but somewhat weaker correlations for the other categories, likely reflecting noise in the ground truth measures. For example, the political alignment of corporations is based on employee contributions, which are often given to both parties in order to hedge the election outcome.

Category	Correlation	p	N
Congress (111th-113th)	0.955	$p < 0.001$	552
111th	0.951	$p < 0.001$	362
112th	0.954	$p < 0.001$	462
113th	0.960	$p < 0.001$	499
Musicians	0.746	$p < 0.001$	66
Fortune 100	0.597	$p < 0.001$	100
Pundits	0.981	$p < 0.001$	35
Chain Restaurants	0.657	$p < 0.01$	24

Table 1: Correlation between imputed ideology and observed political alignments for politicians, pundits, musicians, corporations, and restaurants.

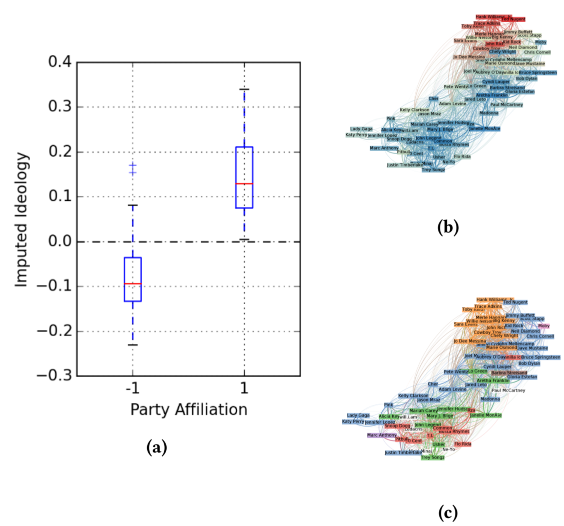

Figure 3: a. Comparison of imputed ideology with party affiliation (based on candidate endorsements) for 66 musicians. b and c are bipartite network projections of musicians, colored by imputed ideology (b) and music genre (c) and located by pairwise affiliations.

Figure 2 reports imputed ideology for 35 popular liberal and conservative political pundits in the U.S. corresponding to the correlation reported in Table 1 (r=.981, $p < 0.001$).

3.2 Musicians

Similarly, we collected all the musicians who publicly endorsed presidential candidates in the 2008 and 2012 elections. Figure 3 compares imputed alignments of musicians with their political endorsements obtained from Wikipedia pages. Figure 3a plots imputed political alignment by party affiliation (based on candidate endorsements). The results correspond to the correlation reported in Table 1 (r= 0.746, $p < 0.001$). Figure 3 show the bipartite network projections, with node color indicating imputed political alignment (3b) and genre (3c), and node location based on pairwise affiliation. In Figure 3b and all the network projections that follow, nodes are located based on pairwise bipartite affiliations in the projected network [22]. In the network projection, the strength of the affiliation between two pundits increases with the relative overlap of their co-followers. These affiliations correspond to pairwise proximities in the visualization using the "backbone" procedure introduced in [20] for complex weighted networks. As shown in the figure, country music is the genre most likely to be co-followed with Congressional conservatives, while all the other genres are aligned with moderates or liberals.

Figure 4 extends the analysis to musicians of unknown political affiliation, using data from Musicbrainz.com and allmusics.com. Results are consistent with those in Figure 3b, with country musicians clustered and distinct from other genre whose political alignments are mostly moderate to liberal.

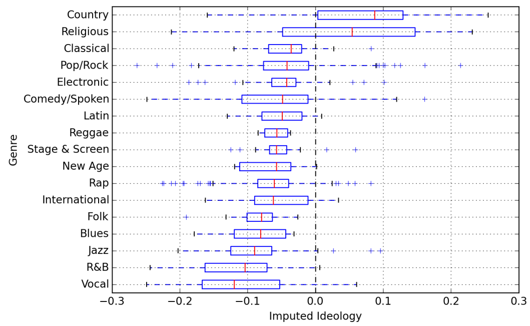

Figure 4: Music genre and political alignment. Ordered political alignment scores for music genres.

3.3 Fortune 100 Companies

We identified political alignments of Fortune 100 companies using employee donations to federal and state political campaigns listed in the Dime database built from the Federal Election Commission data [7]. We constructed two measures of political alignment based on these employee donations, the proportion of the total amount donated that went to the Republican Party (which is heavily biased by the large donations from top executives) and the proportion of the number of unique donations that went to the Republican Party (which indicates the political preference of the majority of employees). We limited the analysis to the top 100 most followed companies that have at least 100 unique donations in the 2008 to 2012 election cycles. Figure 5 visualizes the correlation reported in Table 1 (r = 0.597, $p < 0.001$) between employee campaign donations and the ideology imputed from co-followers of the company and members of Congress, indicating that the larger the proportional contribution made to the Republican Party, the more likely the company is co-followed on Twitter by followers of Congressional conservatives.

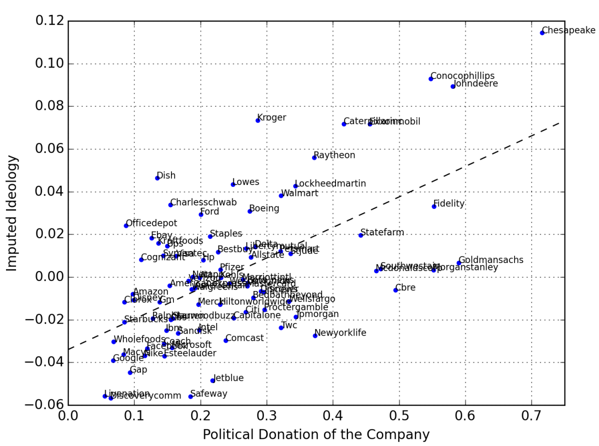

Figure 5: Imputed political alignment of the Fortune 100 and the political donations of employees.

3.4 Restaurants

We measured the political preferences of customers of major restaurants and supermarkets using data from marketing surveys conducted by Experian Marketing Services (http://on.wsj.com/1lGJKzb). Figure 6 visualizes the correlation reported in Table 1 (r=.657, $p < 0.01$) between observed scores and the political alignment imputed from co-followers with Congress. The results indicate general agreement between the restaurants' customers and followers, with the curious exception of the Big Boy restaurant.

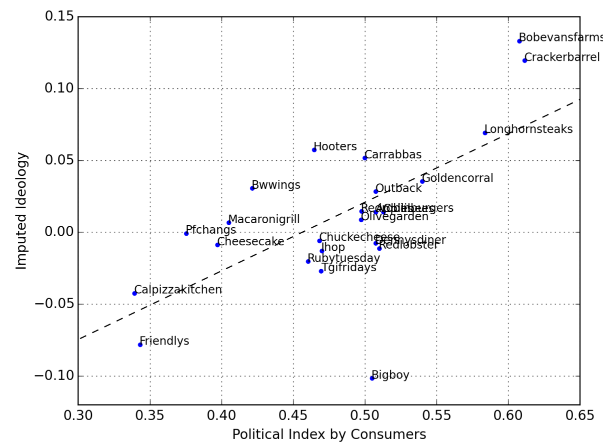

Figure 6: Imputed political alignment of restaurants and the political preferences of customers.

4 POLITICAL RELEVANCE AND POLITICAL POLARIZATION

Figure 7 reports political relevance and polarization of several cultural dimensions, including music and sports, with political alignment indicated by color (ranging from red for conservative followers to blue for liberal). Figure 7 includes Congress members and political pundits to provide a benchmark for assessing the politicization of presumably non-political cultural interests. Results for political relevance show political interest in all the widely followed cultural dimensions is substantially less than for pundits, as might be expected. Nevertheless, these cultural dimensions are not apolitical. Religion attracts the most political interest, followed closely by restaurants. Religious accounts also have the strongest conservative alignment. Religion is also among the most highly polarized, indicating that liberals and conservatives follow different religious accounts. In contrast, music tends to attract more liberals than conservatives, and the Billboard 100 artists are the most polarized, due to the concentration of conservative interest in country western and religious music. Figure 8 graphically shows the polarized following pattern of Billboard 100 musicians. Country Western musicians are highly clustered and separated from other musicians by having much more conservatives followers, in contrast to other musicians whose followers tend to be slightly to the left of center. Figure 7 also shows cultural dimensions that are healing the political divide, including Automobile, News and Media, Fortune 500, and TV programs.

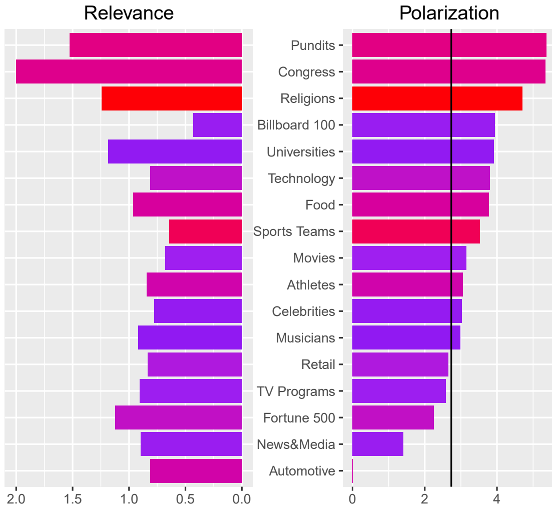

Figure 7: Comparisons of political relevance, polarization and alignment of widely followed cultural dimensions. Color indicates alignment from liberal (blue) to conservative (red), corresponding to the relative interest in the topic by liberals versus conservatives. The widths of the bars (x axis) in the left panel indicate political relevance of cultural topics, and those in the right panel indicate political polarization. Congressmen and pundits are included as a political benchmark for assessing the politicization of cultural dimensions.

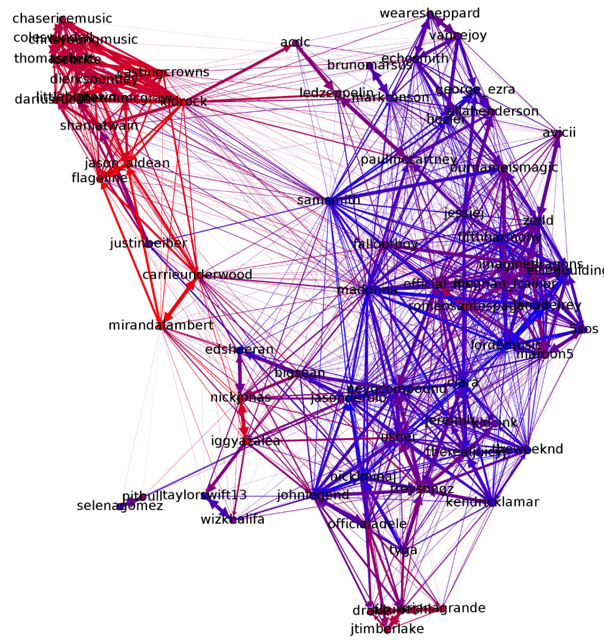

Figure 8: Network Visualization of the Co-following Pattern of the Billboard 100 Music Artists. Nodes are colored according to their imputed political alignment.

5 DISCUSSION AND CONCLUSION

Our analysis of Twitter data suggests that the widely discussed partisan divisions on public policy extend to cultural tastes and lifestyle preferences. Placed in historical context, the cultural fault lines evident on social media have a troubling implication: A long history of assimilating immigrants into a cultural "melting pot" appears to have given way to the formation of politically aligned lifestyle enclaves. Our research also shows how social media data can provide an important window into the polarization of culture that complements traditional survey methods. Opinion surveys are limited to retrospective responses to questions that interest investigators, while social media data provide real time behavioral indicators, significantly expanding the scope and depth of the potential research questions that can be addressed.

REFERENCES

[1] Alan I. Abramowitz. 2011. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy.* Yale University Press.
[2] Lada A. Adamic and Natalie Glance. 2005. The Political Blogosphere and the 2004 U.S. Election: Divided They Blog. In *LinkKDD*. 36–43.
[3] Delia Baldassarri and Andrew Gelman. 2008. Partisans Without Constraint: Political Polarization and Trends in American Public Opinion. *Amer. J. Sociology* 114, 2 (2008), 408–446.
[4] Pablo Barberá. 2015. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis* 23, 1 (2015), 76.
[5] Bill Bishop. 2008. *The Big Sort: Why the Clustering of Like-minded America Is Tearing Us Apart.* Houghton Mifflin.
[6] Robert Bond and Solomon Messing. 2015. Quantifying Social Media's Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook. *American Political Science Review* 109, 01 (2015), 62–78.
[7] Adam Bonica. 2013. Database on Ideology, Money in Politics, and Elections: Public version 1.0. (2013). http://data.stanford.edu/dime.
[8] Michael Conover, Jacob Ratkiewicz, Matthew R. Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. 2011. Political Polarization on Twitter. In *ICWSM*.
[9] Philip E. Converse. 2006. The nature of belief systems in mass publics (1964). *Critical Review* 18, 1–3 (2006), 1–74.
[10] Daniel DellaPosta, Yongren Shi, and Michael Macy. 2015. Why Do Liberals Drink Lattes? *Amer. J. Sociology* 120, 5 (2015), 1473–1511.
[11] Paul DiMaggio, John Evans, and Bethany Bryson. 1996. Have American's Social Attitudes Become More Polarized? *Amer. J. Sociology* 102, 3 (1996), 690–755.
[12] Morris P. Fiorina, Samuel J. Abrams, and Jeremy Pope. 2005. *Culture War? The Myth of a Polarized America.* Pearson Longman.
[13] Venkata Rama Kiran Garimella and Ingmar Weber. 2014. Co-following on twitter. In *HyperText*. 249–254.
[14] Venkata Rama Kiran Garimella and Ingmar Weber. 2017. A Long-Term Analysis of Polarisation on Twitter. In *ICWSM*. to appear.
[15] Scott A. Golder and Michael W. Macy. 2014. Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology* 40, 1 (2014), 129–152.
[16] James Davison Hunter. 1991. *Culture Wars: The Struggle to Define America.* Basic Books.
[17] William Labov. 2010. *Principles of Linguistic Change, Volume III, Cognitive and Cultural Factors.* Wiley-Blackwell.
[18] Keith T. Poole and Howard Rosenthal. 1991. Patterns of congressional voting. *American Journal of Political Science* 35 (1991), 228–278. Issue 1.
[19] Keith T. Poole and Howard Rosenthal. 2001. *Congress: a Political-economic History of Roll Call Voting.* Oxford University Press.
[20] M. Ángeles Serrano, Marián Boguñá, and Alessandro Vespignani. 2009. Extracting the multiscale backbone of complex weighted networks. 106, 16 (2009), 6483–6488.
[21] Feng Shi, Yongren Shi, Fedor Dokshin, Michael W. Macy, and James Evans. 2017. Millions of Online Book Co-purchases Reveal Partisan Differences in the Consumption of Science. *Nature Human Behavior* 3, 0079 (2017). DOI: http://dx.doi.org/10.1145/1219092.1219093
[22] Yongren Shi and Michael W. Macy. 2016. Measuring Structural Similarity in Large Online Networks. *Social Science Research* 59 (2016), 97–106.