(Anti-)Echo Chamber Participation

Examining Contributor Activity Beyond the Chamber

Ella Guest
University of Manchester
Department of Social Statistics
Manchester, United Kingdom
ella.guest@postgrad.manchester.ac.uk

ABSTRACT

The existence of echo chambers as a phenomenon in social media has been widely debated in recent years. This research examines the social news website *Reddit*, where users engage with each other indirectly as part of topic-based communities. These discrete communities are free to define their own rules for contributor behaviour within the community. In this paper we examine two communities : *r/The_Donald*, which demands echo chamber behaviour, and *r/changemyview* which discourages it. We seek to determine whether contributors to the opposing types of chambers behave differently beyond their respective chambers. We find evidence that participants of both the echo chamber and the anti-echo chamber are more active across the platform than the average *Reddit* user.

CCS CONCEPTS

• **Human-centered computing** → **Social networks**; **Computer supported cooperative work**; **Social media**; **Empirical studies in collaborative and social computing**; **Social content sharing**; • **Applied computing** → **Sociology**;

KEYWORDS

Social media, social news, echo chamber, Reddit

1 INTRODUCTION

The term "echo chamber" has become a serious focus of public and academic attention in recent years [8]. Much of this discussion has been theoretical and often taken for granted the existence of echo chambers on social media, without clearly defining what constitutes an echo chamber [2]. To date, little empirical evidence has been published to support the belief that online users tend to engage exclusively with people of similar views. Recent work has provided evidence of echo chambers on Twitter [1, 5]. This research seeks to find empirical evidence for the existence, or lack thereof, of echo chambers on the social news site Reddit.

Reddit is organized into separate topic-based communities, called "subreddits". Users of the platform, "redditors", can contribute relevant content to a subreddit in the form of a post. Other redditors then discuss posts in comment threads. Unlike social networking sites such as Twitter where users can directly connect with others via following other users, connections between redditors are less explicit and can be defined in many ways, such as browsing or subscribing to the same subreddits.

A common difficulty in researching echo chambers is identifying the boundaries of a chamber and determining who is or is not a participant. However, subreddits are unlike many other online communities in that users congregate around a topic and within a defined space. Each subreddit has a unique nature and can develop its own standards of content and behaviour. Subreddits thus have the ability to define themselves as encouraging (indeed, even demanding) echo chamber behaviours.

This research focuses on two subreddits with clearly defined content policies that allow them to be readily classified as an echo chamber and an "anti-echo chamber", respectively. As we will see in the next section, *r/The_Donald* is an echo chamber for supporters of President Donald Trump, while *r/changemyview* is an anti-echo chamber where redditors actively invite discussion between people of different opinions.

We look at the redditors who participate in each of these subreddits and, in turn, the range of other subreddit communities in which they participate. We seek to determine whether redditors who choose to participate in the echo chamber show echo chamber-like behaviours across the platform, and the same for anti-echo chamber-like behaviours for participants of *r/changemyview*.

A direct comparison alone would not allow us to determine whether chamber participants behave differently than the average redditor. Therefore we include a random selection of other subreddits in our analysis to determine a baseline of expected behaviours for the platform.

2 RESEARCH OUTLINE

2.1 Research Context

2.1.1 Selecting a Definition. Brun' recent work on Twitter provides a useful definition: "an echo chamber comes into being where a group of participants choose to preferentially connect with each other, to the exclusion of outsiders" [2].

The emphasis on preference to connect highlights the process of choice homophily, whereby redditors choose to participate in a community, knowing it is a space for sharing common views [5]. In Bruns' context, Twitter users connect by following each other

[3], while on Reddit users are connected by the common subreddits in which they participate. Thus, for the purposes of this research a subreddit-community is considered to be an echo chamber if it fosters participation among a group of homophilous users, and excludes heterophilous outsiders.

Most subreddits have written content policies outlining the purpose of the community. Some subreddits also have clear policies on types of content that are actively excluded. We apply these explicit, community-provided descriptions to determine whether a subreddit can be considered an echo chamber, or an anti-echo chamber. For the purposes of this research we focus on the content of contributions, i.e. their topic and perspective. Future work will consider the form oof contributions, i.e.

2.1.2 The Echo Chamber. *r/The_Donald*'s content policy clearly excludes outsiders: "Trump Supporters ONLY - This sub is for supporters of Donald J. Trump ONLY. This is not a place for you to debate with us about Donald Trump, or to ask us to convince you to like Donald Trump. This is not a neutral place - we are 100% in support of Donald J. Trump. Moderators reserve the right to ban non-supporters as we see fit" [7].

Contributors must be Trump supporters, and anyone with a non pro-Trump opinion is discouraged (or possibly banned) from participating. Thus, *r/The_Donald* fits our working definition of an echo chamber because it explicitly bans participation that does not support the homophilous opinion.

2.1.3 The Anti-Echo Chamber. *r/changemyview*'s content policy clearly welcomes a variety of opinions: "[*r/changemyview*] is dedicated to the civil discourse of opinions, and is built around the idea that in order to resolve our differences, we must first understand them... anyone can post so long as they have an open-mind and are looking to consider other perspectives" [6].

In opposition to the echo chamber policy of exclusion, *r/changemyview* is intended to be a meeting place for people for heterophilous opinions. As such we define it as an "anti-echo chamber" in the context of content diversity.

2.2 Research Questions & Methodology

2.2.1 Range of Participation. By excluding outside views, *r/The_Donald*'s participants choose to limit the range of content with which they engage within the community. This willingness to engage with a only limited range of content may be specific to participants' behaviour in *r/The_Donald*, or it might also reflect a limitation in the range of content they choose to engage with outside of the echo chamber.

Conversely, *r/changemyview* participants are likely to engage with a wider range of content. Their openness to diverse content may only occur within the anti-echo chamber, or it may be reflected in the range of content they engage with beyond the anti-echo chamber. To examine the range of content a redditor engages with we look at the total number of communities they participate in on *Reddit.*

- *Hypothesis 1a:* r/The_Donald *participants engage with fewer communities than the average redditor.*
- *Hypothesis 1b:* r/changemyview *participants engage with more communities than the average redditor.*

2.2.2 Concentration of Participation. We are also interested in the level of engagement participants have with their community. The literature suggests that echo chambers form in part because people are biased towards having their beliefs reinforced, and that this choice homophily leads to a bonding social capital [1, 4]. If echo chamber participation produces these positive experiences of validation and bonding we would expect participants to be drawn to engage often with the community.

Following on from Hypothesis 1a, if *r/The_Donald* participants interact with a limited number of communities, they may contribute relatively more to the echo-chamber than to other communities. In addition, given the narrow focus of content in *r/The_Donald*, participants are expected to be at least partially interested in most content, and may thus be drawn to participate more.

Conversely, following Hypothesis 1b, if *r/changemyview* participants are active in a wide range of communities they are likely to contribute relatively less to the anti-echo chamber than other communities. In addition, given that a wide range of content is discussed in *r/changemyview*, participants are expected to only engage with a subset of content that is of interest to them.

- *Hypothesis 2a:* r/The_Donald *participants' Reddit contributions will be more concentrated within the echo chamber than the contributions of participants in the average community.*
- *Hypothesis 2b:* r/changemyview *participants' Reddit contributions will be less concentrated within the anti-echo chamber than the contributions of participants in the average community.*

2.2.3 Summary of Research Questions. The first research question examines, broadly, whether the echo chamber behaviour of choosing to engage with a limited range of content extends beyond the echo chamber. The second research question examines the level at which redditors' platform-wide participation is concentrated within the echo chamber. Together these two research questions are the first stage in informing future work into whether redditors' echo chamber behaviours in one community do or do not reflect their engagement with content across the platform.

2.3 Data Collection & Sampling

2.3.1 Data Collection. Redditor *u/Stuck_In_the_Matrix* uses the Reddit API to collect all comments made across the platform and stores them in monthly subsets on Google BigQuery [9, 10]. As this research used an existing, publicly available dataset of pseudonymous participants it did not require approval by a research ethics board. For this initial analysis we used comments from the month of June 2017. Comments are used instead of posts because redditors comment much more frequently than they post and because comments better reflect engagement between redditors. First the usernames of all redditors to comment in *r/The_Donald* or *r/changemyview* in June 2017 were collected. For each of those redditors we then collected a list of all subreddits they commented in that month, and the number of comments they contributed to each subreddit.

2.3.2 Selection of Baseline Subreddits. Data was also collected for a selection of other subreddits for baseline comparison. First a random sample of 1000 subreddits was selected. Then up to 1000
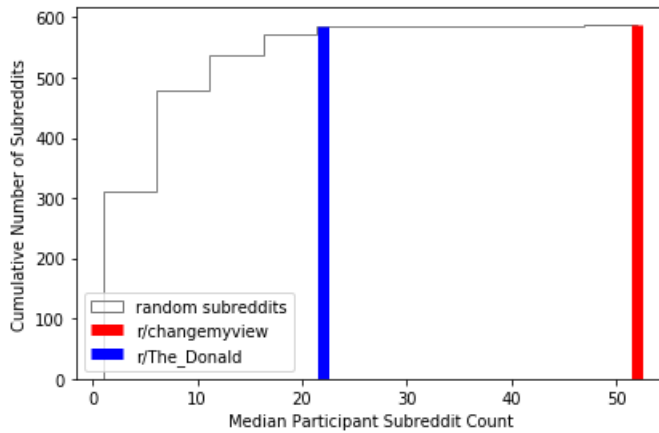
Figure 1: Cumulative Distribution Function of Median Contributor Subreddit Count

Figure 2: Cumulative Distribution Function of Median Contributor In-subreddit Ratios

random participant names were selected for each subreddit. Subreddits with fewer than 1000 participants were dropped from analysis, leaving 578 baseline subreddits. Participant samples were taken independently so one participant could be sampled for multiple subreddits. Again, for each author we collected a list of all subreddits they contributed to and the number of comments they made per subreddit.

2.3.3 Removal of Known Bots. Bots are frequently used on Reddit to perform automated tasks. Bot accounts cannot be automatically identified, thus it is impossible to ensure all bots are removed from the dataset. For the purposes of this preliminary work only accounts that we could conclusively identify as bots were removed. As bots can post much more frequently than human redditors, the majority of most frequently commenting authors in the dataset were manually identified as bots. The 20 authors with the most comments were manually checked and 17 were identified as bots and excluded from analysis. Future work will test more developed techniques for bot identification.

3 RESULTS AND DISCUSSION

3.1 Range of Participation

We analysed 580 subreddits in total: *r/The_Donald*, *r/changemyview*, and the baseline 578 randomly selected. Each participant has a "subreddit count", the total number of subreddits they commented in during June 2017. Subreddit counts ranged from a minimum of 1 to a maximum of 2862. The redditor with the subreddit count of 2862 is likely a bot but this could not be confirmed, thus they were included in this conservative analysis (see 2.3.3). As the range of subreddit counts was so large (μ=17, σ=91) we chose to use median as our measure of average when aggregating contributor subreddit counts for each subreddit.

The median subreddit count for all contributors included in the analysis was 6, with an interquartile range (IQR) of 4-9. Figure 1 shows the cumulative distribution function of median contributor subreddit count for each of our 580 subreddits. Participants in *r/The_Donald* have a median subreddit count of 22 (IQR 6-72). Only
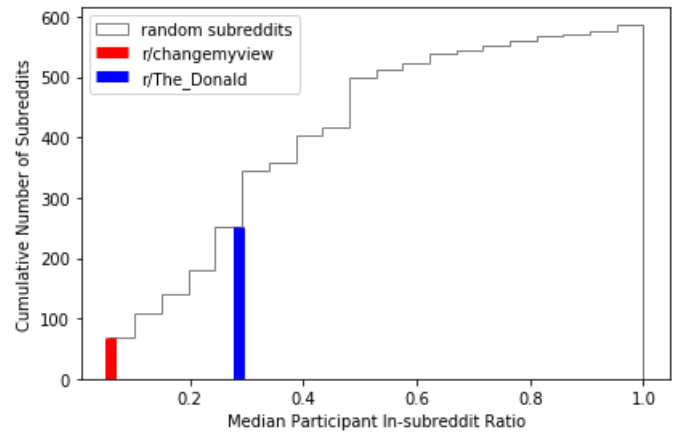
5 of the 580 subreddits have a median contributor subreddit count greater than *r/The_Donald*'s. Thus we must reject Hypothesis 1a:

- *Finding 1a: Contrary to our hypothesis,* r/The_Donald *contributors engage with more communities than the average redditor.*

Contributors to r/changemyview have a median subreddit count of 52 (IQR 16-142), the largest of the 580 subreddits. The second highest median contributor subreddit count was 26 (IQR 14-46) for the community *r/AgainstHateSubreddits.* Thus r/changemyview has the highest median contributor subreddit count by a factor of 2 (52/26 = 2) and we find strong support for Hypothesis 1b:

- *Finding 1b: Following our hypothesis,* r/changemyview *contributors engage with more communities than the average redditor.*

3.2 Concentration of Participation

For each contributor we measured how much their participation across *Reddit* was concentrated within one subreddit by calculating their "in-subreddit ratio". The in-subreddit ratio is the number of comments they made within the subreddit divided by the total number of comments they made across all communities [11]. Figure 2 shows the cumulative distribution function of contributor in-subreddit participation ratios for the 580 subreddits.

The overall median is 0.33 (IQR 0.2-0.5), meaning for half of the sampled subreddits contributors tend to make $1/3^{rd}$ or less of their comments on the platform within the sampled subreddit. Participants of r/The_Donald had a median in-subreddit ratio of 0.28 (IQR 0.07-0.84). On average, 28% of all comments *r/The_Donald* contributors made on *Reddit* in June 2017 were made in *r/The_Donald*. This puts *r/The_Donald* in the 42nd percentile for median in-subreddit ratio. Thus we must reject Hypothesis 2a:

- *Finding 2a: Contrary to our hypothesis:* r/The_Donald *contributors' Reddit participation is less concentrated within the echo chamber than that of contributors in the average community.*

Contributors of *r/changemyview* had a median in-subreddit ratio of 0.06 (IQR 0.02-0.2). On average, 6% of all comments *r/changemyview* participants made on Reddit were made in r/changemyview. Only

5 subreddits in the analysis had a smaller median in-subreddit ratio than r/changemyview's, putting it in the 1st percentile. Thus we find strong support for Hypothesis 2b:

- *Finding 2b: Following our hypothesis:* r/changemyview *contributors' Reddit particiation is more concentrated within the anti-echo chamber than that of contributors in the average community.*

4 DISCUSSION AND NEXT STEPS

4.1 Range of Participation

We have found evidence to support our hypothesis that participants of the anti-echo chamber *r/changemyview* participate in a larger range of communities than the average redditor. However, we have also found that participants of the echo chamber *r/The_Donald* also participate in a greater range of communities than the average redditor. This might suggest that echo chamber participants either a) engage with the same range of content as the average redditor, or b) the range of communities they engage with do not correspond with a range of content. The latter would be possible if there is overlap in the content of the subreddits *r/The_Donald* contributors participate in.

As such, further analysis is needed to determine whether engagement with a range of communities actually corresponds to engagement with diverse content. Our next steps will be to use text analysis to assess the topical content of subreddits, and network analysis to determine whether echo chamber participants regularly contribute to the same communities beyond the echo chamber. These methods of analysis will allow us to determine whether the limits of the echo chamber are contained within *r/The_Donald* or if a cluster of subreddits form a meta-community of like-minded individuals.

Similarly, for *r/changemyview* we will also examine whether the greater number of communities participants engage with actually corresponds to engagement with a greater range of content. We must also test the validity of our categorization of *r/changemyview* as an anti-echo chamber. The community invites people with a range of views to debate their opinions in a rational, evidence-based argument style. While at face value this is a sign of inclusivity, by extension it excludes redditors who do not want to have their opinions questioned and/or who do not have the knowledge or willingness to participate in formal, evidence-based debates. This may be a mechanism of choice homophily of shared forms of behaviours instead of shared opinions, which has been the focus of this research so far.

4.2 Concentration of Participation

We found that for participants of both the echo chamber and the anti-echo chamber, their platform wide participation was less concentrated in the chamber than on average. In other words, they contribute relatively more of their time to other communities. This suggests that those in the anti-echo chamber of *r/changemyview* are not especially loyal, or regular contributors to the chamber and/or that they engage with more compelling communities elsewhere. However, we find this result surprising for participants

of *r/The_Donald*. Existing research has shown that people enjoy having their views validated, thus we expected participants to be regularly drawn to *r/The_Donald* for this reason.

We suggest two possible causes, which we will address in further research. First, as stated in 4.1, *r/The_Donald* participants may be engaging with other communities that form part of a larger echo chamber. Thus they may be finding validation of their opinions in other places, in addition to *r/The_Donald*. Second, our use of aggregate measures overlooks the range of behaviours participants of the same community may exhibit.

Each subreddit has many participants who contribute only a few comments, and conversely a few participants who provide a disproportionate number of comments. As such, the next step in our analysis will be to examine the range of contributor activity within subreddits, and to consider a taxonomy of engagement levels. This will allow us to differentiate between casual participants of the echo chamber, and the more engaged participants who we suggest may most exhibit echo chamber behaviours.

A ONGOING WORK

A more detailed overview of the data, methods, and results can be seen at: http://ellaguest.com/antiechochambers/

REFERENCES

[1] Andre Boutyline and Robb Willer. 2017. The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology* 38, 3 (2017), 551–569.

[2] Axel Bruns. 2017. Echo chamber? What echo chamber? Reviewing the evidence. In *In 6th Biennial Future of Journalism Conference (FOJ17)*. Cardiff, UK.

[3] Axel Bruns, Brenda Moon, Felix Munch, and Troy Sadkowsky. 2017. The Australian Twittersphere in 2016: Mapping the Follower/Followee Network. *Social Media + Society* 3, 4 (2017), 370–379.

[4] Lorien Jasny, Joseph Waggle, and Dana R. Fisher. 2015. An empirical examination of echo chambers in US climate policy networks. *Nature Climate Change* 5, 8 (2015), 782.

[5] J . Miller McPherson and Lynn Smith-Lovin. 1977. Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review* 52 (1977), 370–379.

[6] r/changemyview. 2018. r/changemyview Wiki. https://www.reddit.com/r/changemyview/wiki/index

[7] r/The_Donald. 2018. r/The_Donald Wiki. https://www.reddit.com/r/The_Donald/wiki/index

[8] Cass Sunstein. 2007. *Republic.com* (2nd. ed.). Princeton University Press, Princeton, N.J.

[9] u/Stuck_in_the_Matrix. 2015. I have every publicly available Reddit comment for research. https://www.reddit.com/r/datasets/comments/3bxlg7/i_have_every_publicly_available_reddit_comment/

[10] u/Stuck_In_the_Matrix. 2017. Reddit Comments June 2017. https://bigquery.cloud.google.com/table/fh-bigquery:reddit_comments.2017_06