Communities, Gateways, and Bridges: Measuring Attention Flow in the Reddit Political Sphere

Cesare Rollo[1], Gianmarco De Francisci Morales[2], Corrado Monti[2], and André Panisson[2]

[1] University of Turin, Turin, Italy, `cesare.rollo@unito.it`
[2] CENTAI, Turin, Italy, `gdfm@acm.org, monti@centai.eu, panisson@centai.eu`

Abstract. Online social media have attracted large and vibrant communities, which shape how people interact online. Platforms such as Reddit provide a safe harbor for groups to discuss a variety of topics, including politics and even conspiracy theories. We propose a framework, dubbed *attention-flow graph*, to investigate the flow of users across Reddit communities from a network perspective. This graph concisely summarizes how users shift their attention from one subreddit to another over time, and allows to capture its community structure. In addition, it enables the operationalization of the concepts of gateways and bridges: particular subreddits that support the transition of users towards specific communities. We apply this framework to identify political and conspiracy communities, thus discovering their bridges and gateways. We find that conspiracy theories help attracting users to the alt-right community from occultist subreddits, but also by diverting users from the radical left.

Keywords: Social media · Network science · Conspiracy · Echo chambers.

1 Introduction

The effect of social media on how the general public forms its opinion is a pressing issue of the last decades, due also to their role in several contemporary political events, such as Donald Trump's election (Enli, 2017). Some researchers argue that their usage has been accompanied by a growth in mass ideological polarization (Kubin and von Sikorski, 2021). At the same time, surveys have reported a large growth in affective polarization (Iyengar et al., 2012)—individuals unwilling to socialize across ideological boundaries. These phenomena have been linked to selective exposure: users prefer information sources that agree with their views (Stroud, 2010), and sort themselves into communities of like-minded individuals, i.e., "echo chambers" (Cinelli et al., 2021). Whether this tendency has been exacerbated by social media is an open debate (Boxell et al., 2017).

Some particularly concerning examples of echo chambers are provided by conspiracy theories such as QAnon, that have led to outbursts of violence;[3] or the growth of far-right extremism in the U.S., which has been nurturing recent

---

[3] https://www.theguardian.com/us-news/2020/oct/15/qanon-violence-crimes-timeline

terrorist events, such as the 2022 Buffalo Attack (Abbas et al., 2022). Models of online radicalization highlight the effect of social interactions in the process, and in particular the "influence of like-minded individuals and the online community on the individual's new worldview" (Neo, 2019). Thus, it is important to study the places where this connection happens, and how *gatekeepers* of information shape the worldview within the echo chamber (Garimella et al., 2018a).

The original work on gatekeeping (Lewin, 1947) referred to the power of gate-keepers to decide which messages may pass through their channels. More modern views of gatekeeping re-imagine it within the context of user-generated content and social media. In this context, users do not act as filters, rather they decide, in an ecosystem of complex processes which may involve algorithms (Thorson and Wells, 2016), which items the collective attention focuses on (Bruns, 2003). Indeed, the paradigm shifts from editors keeping the gates of publishing, to users pointing out which sources (gates of information) to watch, thereby fostering the dissemination of specific worldviews (Welbers and Opgenhaffen, 2018). We adopt a similar view in this work, and ask ourselves which forums can have a role in supporting a shift in the attention of Reddit users towards specific topics and ideas. Given the decentralized, peer-to-peer nature of the modern information gathering process, it is natural to approach it from a network perspective.

It is well-known that new ideas tend to flow into a community via boundary spanners and brokers (Burt, 2004). At the same time, bridges, which span structural holes (Burt, 1992), connect separate communities, while gatekeepers facilitate or hinder information flows (West, 2017). In particular, we are interested in which communities serve as entry points and precursors to other communities— noticeably, conspiracy and hoaxes-spreading ones. In this sense, our work is similar in spirit to the ones by Klein et al. (2019) and Phadke et al. (2022). However, we develop a general framework to analyze the flow of attention within Reddit, which can identify precursors to engagement with specific communities.

Much of the attention has been focused on Facebook (Quattrociocchi et al., 2016), Twitter (Garimella et al., 2018a), and YouTube (Ribeiro et al., 2020b; Fabbri et al., 2022). Our goal is to empirically explore this phenomenon by using Reddit as the object of study. Reddit, which self-defines as "the front page of the Internet", is a social news and discussion website which has consistently ranked among the top ten most visited websites worldwide over the past years.[4] It is organized in topical forums called subreddits, each dedicated to a specific topic of discussion; users typically subscribe to and participate in multiple subreddits of interest. Its role in the development of far-right and conspiratorial attitudes has been extensively scrutinized (Massachs et al., 2020; Phadke et al., 2022). Indeed, these studies have shown that the fora frequented by users are powerful indicators for their future trajectories. For instance, dyadic interactions with members of conspiracy communities are the most important social precursors to joining conspiracy communities (Phadke et al., 2021). Furthermore, engagement with specific communities (e.g., men's rights activists) has been identified as a gateway towards alt-right extremism (Mamié et al., 2021). Therefore, places that

---

[4] https://en.wikipedia.org/wiki/Reddit

act as melting pots for users within and without the conspiracy community are fundamental to understand the pathways taken by users.

We propose a graph-based approach to analyze the flow of attention across communities on Reddit. Fabbri et al. (2022) proposed a similar approach to study how YouTube videos can represent a gateway that introduce users to a radicalization path. This concept is modeled as a path on a what-to-watch-next graph that leads user to harmful content. In our case, we look for specific subreddits that have a comparable role in shifting the attention flow of users towards a community. To do so, we draw inspiration from the work by Davies et al. (2021) which analyses user migration within Reddit. They find an interesting interplay between user migration and controversies, increasing politicization, and the rise of conspiracy theories. We wish to leverage this interplay to uncover the underlying common structure between user attention and (political) topics.

Our approach based on the *attention-flow graph* provides a way to identify the main communities within Reddit based on the flow of users within them (Cai et al., 2019), rather than on the static view used by co-participation networks (Phadke et al., 2022; Waller and Anderson, 2021). Similar approaches have been already used to analyze behavior on Reddit (Massachs et al., 2020; Ribeiro et al., 2020a), however they have been mostly ad-hoc. The attention-flow graph is a way to formalize and systematize this type of analysis. As such, the attention-flow graph allows to identify important landmarks in the pathways across the website such as *gateways* from and *bridges* to other communities. In particular, while subscription information is not public, the production of content on a subreddit can be seen as a proxy for membership (Datta and Adar, 2019). Finally, although our specific insights regard political and conspiratorial communities, our methodology is more general, and can be applied to the study of any topic on Reddit. For instance, while out of the scope of the current study, the attention-flow graph can be used to study the effect of moderation within the platform (e.g., banning of subreddits) (Chandrasekharan et al., 2017; Horta Ribeiro et al., 2021).

2   Data

Reddit is a social news aggregation and discussion platform. It has been recognized as an influential one, especially for the alt-right and conspiratorial communities: Zannettou et al. (2017) have shown that fringe communities on Reddit are comparatively successful in spreading their content to more mainstream media.

Discussions on Reddit are organized in topical communities called *subreddits*. Users (known as *redditors*) can publish *posts* in subreddits, and *comment* on other posts and comments, thus creating a tree structure for the overall discussion. In addition, users can also *upvote* a message to show approval, appreciation, or agreement (and their opposites with a *downvote*). The *score* of a message is the number of upvotes minus the number of downvotes. In this study, we restrict our attention to posts (rather than comments), since they are a better proxy for the active involvement of a user in a specific subreddit. Each subreddit focuses on a specific topic, thus subscribers to the same subreddit can be seen as sharing similar interests. For instance, some subreddits gather politically-aligned

4 C. Rollo et al.

groups of users by limiting their participants to supporters of a political group or figure (e.g., in r/The_Donald, moderators explicitly state that the community is for 'Trump Supporters Only' and that dissenting users will be removed).

We extract all Reddit posts from 2017 to 2020 from the PushShift dataset (Baumgartner et al., 2020). To account only for users positively engaged with a community, we keep only posts with score greater than one. We exclude smaller subreddits by selecting only those that have at least one month with at least 50 distinct users publishing a post on it. Finally, we take care of automated accounts by removing users who posted on more than 50 distinct subreddits for at least one month and users with bot in their username (except common dictionary words). In the end, we obtain a data set with 242 220 979 posts submitted by 3 677 137 users in 25 877 subreddits.

3 Methods

In the following, we first state our definition of *attention flow* (Section 3.1) and how it can be transformed into a graph (Section 3.2). Then, we operationalize of the concepts of communities, gateways, and bridges by using this graph (Section 3.3).

3.1 Attention flow

Our first goal is to formalize the concept of attention flow across subreddits, meant to capture how much each subreddit at time step t' contributes to the user base of another subreddit at a time step $t > t'$. In practice, we always consider $t = t' + 1$, where the time step represents one month.

Let $\mathcal{U}$ be the set of users and $\mathcal{S}$ be the set of subreddits. Define $c_{u,s}^{(t)}$ as the number of interactions of user $u \in \mathcal{U}$ with subreddit $s \in \mathcal{S}$ at time step t, and $\mathbf{C}^{(t)} \in \mathbb{N}^{|\mathcal{U}| \times |\mathcal{S}|}$ as the corresponding matrix for all users and all subreddits at time step t. We define the attention matrix $\mathbf{B}^{(t)}$ as the row-normalized form of $\mathbf{C}^{(t)}$, and each row $\mathbf{b}_u^t$ as the attention vector of user u at time step t. Therefore, for each user u and subreddit s, we quantify how much the user's attention changes between t and t' as follows:

$$\mathbf{\Delta b}_{u,s}^t = \begin{cases} \mathbf{b}_{u,s}^t - \mathbf{b}_{u,s}^{t'} & \text{if } \mathbf{b}_{u,s}^t \cdot \mathbf{b}_{u,s}^{t'} = 0 \\ 0 & \text{otherwise} \end{cases}$$

This representation highlights the subreddits s that are adopted or abandoned by user u between t' and t, represented in a vector $\mathbf{\Delta b}_u^t$ where the adopted subreddits are represented as positive flow $\mathbf{\Delta}^+\mathbf{b}_u^t$, and the abandoned ones as negative flow $\mathbf{\Delta}^-\mathbf{b}_u^t$. By evenly matching the negative flows with the positive ones, we define the attention flow of user u as the outer product $\mathbf{F}^{(t,u)} = \widehat{\mathbf{\Delta}^-\mathbf{b}_u^t} \otimes \mathbf{\Delta}^+\mathbf{b}_u^t$. where $\widehat{\mathbf{\Delta}^-\mathbf{b}_u^t}$ is the l_1-normalized form of the absolute values in $\mathbf{\Delta}^-\mathbf{b}_u^t$.

Each user u active in both t' and t determines a per-user attention flow matrix $\mathbf{F}^{(t,u)} \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|}$ where $f_{i,j}^{(t,u)}$ represents how much of the attention of user u

was transferred from subreddit i to subreddit j between the time steps t' and t. By aggregating over users, $\mathbf{F}^{(t)} = \sum_{u \in \mathcal{U}} \mathbf{F}^{(t,u)}$ represents how much of all users' attention migrated from subreddit i to subreddit j.

3.2 Network construction

Next, we devise a suitable network representation from the attention flow previously defined, by aggregating multiple time steps. $\mathbf{F}^{(t)}$ can be seen as the adjacency matrix of a weighted directed network, which represents the overall attention flow of users among subreddits. Thus, the set of matrices $\{\mathbf{F}^{(t)}\}_{t=1,\dots T}$ represent a temporal flow graph that encodes the user flow among subreddits.

Since network weights represent user flow, they depend on the popularity of both source and target nodes, i.e., the number of users that post on those subreddits in the considered time interval. Node strength s is thus strongly correlated with subreddit popularity. As a consequence, for each node, the main in- and out-neighbors are subreddits with higher popularity, that at all time steps absorb and consume attention from all the others. To avoid this effect, we *rescale* edge weights to dampen their dependency on node popularity: $\widetilde{F_{ij}} := \frac{F_{ij}}{\sqrt{\sum_k F_{ik} \cdot \sum_k F_{kj}}} = \frac{F_{ij}}{\sqrt{s_i^{out} \cdot s_j^{in}}}$. We empirically find that re-scaling by the geometric mean of the in/out-strenghts works best (see Appendix A).

Finally, we define the aggregated network over multiple time steps as $\mathcal{F} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{F}^{(t)}$. Henceforth, we refer to this network as the *attention-flow graph* (AFG). This graph is dense, but most weights have very small values (median 10^{-3}, see Appendix B). Section 4.1 shows that pruning low-weight edges is beneficial to community detection, and the threshold can be determined experimentally.

3.3 Communities, gateways, and bridges

The definition of *attention-flow graph* allows to operationalize the concepts of communities, gateways, and bridges. We use the Stochastic Block Model (SBM) to uncover the underlying community structure of the AFG.[5] A major advantage of using SBM for community detection derives from being a sound statistical model, which lowers the chance to overfit (mistake stochastic fluctuations for actual structure) compared to maximizing modularity (Guimera et al., 2004).

Each community represents a collection of subreddits where users regularly migrate from one to another, while moving to a subreddit external to the community is rare. Indeed, they represent a community that is interested in similar themes. However, they can loosely be seen as *echo chambers*: assuming a user on Reddit reads the same subreddits they participate to, it is unlikely for an individual, during their trajectory on Reddit, to be exposed to content produced outside this bubble. Moreover, if two users belong to distinct communities, their interaction is less likely.

---

[5] Specifically, we use a degree-corrected version of the stochastic block model (Karrer and Newman, 2011), with the addition of the degree sequence $\boldsymbol{k} = \{k_i\}$ of the graph as an additional set of parameters (Peixoto, 2017).

The extraction of such communities enables us to investigate the role of specific nodes in connecting them. In particular, when considering a user that is active inside a community X, we seek to understand which are the likely pathways that drive them towards a different community Y. By modeling the users' movement across subreddits as random walks (where the weighted adjacency matrix $\mathcal{F}$ is the transition matrix of a Markov process), the stationary distribution of walkers that (re-)start their walk from X is given by the Personalized PageRank (PPR) in community X (Garimella et al., 2018b). Thus, we can quantitatively determine which are the specific subreddits $s \in Y$ that are more likely to be crossed by a random walker that belongs to X, and which as bridges from X to Y.

Definition 1 (Bridge node). *A subreddit s belonging to a community Y is a* bridge *from the community X (for $X \neq Y$) if it has a high probability of being reached by a random walk with uniform probability of restart in all $s' \in X$.*

This definition captures the subreddit's property of being a transition point for users that walk away from a community X, heading towards community Y.

Similarly, to find subreddits acting as attractors for X, we can apply a complementary definition, by using the PPR from outside of X.

Definition 2 (Gateway node).
A subreddit $s \in X$ is a gateway *for community X if it has a high probability of being reached by a random walk with uniform probability of restart in all $s' \notin X$.*

Note that the definition of a *bridge* requires to consider a pair of communities (X, Y), starting and arrival, while we define *gateways* only by referring to a unique entry point for community X. By applying these definitions, we obtain for each community a ranking of its top gateway and bridge nodes. Informally, a bridge between X and Y redirects the flow from X to Y, i.e., is an attractor point to Y from X. The subreddits thus identified as bridges can act as transition nodes, through which users typically transit while shifting their attention from one community to another; alternatively, they can represent meeting points for users that have different interests. Conversely, a gateway for X is an ingress point for the whole community X, irrespective of its source. Users ending up in a gateway node have a higher chance of being absorbed into that community. The next section reports some concrete examples within the Reddit political sphere.

4   Results

This section illustrates the results obtained by applying the process described so far to the Reddit data set, with a focus on political and conspiratorial communities. First, we show how we validate our graph construction and community detection by using a human-curated data set (Section 4.1). Then, we restrict our attention to the political communities (Section 4.2), and show how they are connected to each other. Finally, we focus on their gateways and bridges, particular subreddits that support the transition of users from one ideological side to another (Section 4.3).

Table 1: Adjusted mutual information (AMI) scores between ground-truth subreddit categorizations at 3 levels of depth and SBM communities resulting from edge-pruning thresholds at different percentiles.

Pruning	0th	25th	50th	75th
Hierarchy				
Level 1	0.230	0.249	0.280	**0.300**
Level 2	0.283	0.307	0.340	**0.369**
Level 3	0.278	0.293	0.345	**0.409**

4.1 Preprocessing and validation

As mentioned in Section 3.2, the attention-flow graph is very dense; thus, as a preprocessing step, we prune its edges by removing those with smaller weight. To choose the threshold of this pruning process, and to validate the quality of the communities we find, we compare our results to a human-curated data set[6] created by redditors to categorize subreddits into topics. It organizes topics within three levels, e.g., the subreddit `r/chicagobulls` belongs to the *NBA* subtopic, which in turns is part of *Sport*. The data set contains 2017 labeled subreddits, of which 1255 appear in the time frame we consider. Of those, 1211 are also labelled with a second subtopic level, and 712 with a third more specific one.

Table 1 reports the adjusted mutual information (AMI) between the benchmark categorizations and the SBM communities, with edge-pruning thresholds set at different percentiles. First, we observe that with any choice of threshold, the community detection obtains a good fit with respect to the ground truth. In fact, it obtains AMI scores between 0.278 and 0.409 for the finer categorization level—values in line with state-of-the-art results of applications of community detection on real data (George et al., 2020). Secondly, we note that in all levels of depth, we obtain the highest correspondence with the ground truth at the 75th percentile. As such, we consider this threshold in the rest of our analysis.

The giant component of the final pruned network has 24 529 nodes with 581 990 directional edges. A total of 140 communities, with a median size of 122 nodes each, are found by SBM. By inspecting the results, we observe that such communities are based on their topic or partisan affiliation, as expected.

4.2 Political communities

Among all the identified communities, we focus on those related to politics and conspiracies. In particular, we identify four communities of particular interest. **Liberal**. This community covers a wide range of politics and current news topics, mainly from a liberal/centrist point of view. In fact, it includes democratic-wing subreddits such as `r/JoeBiden` or `r/YangForPresidentHQ`, communities opposed to more left-wing politicians (`r/Enough_Sanders_Spam`), and to right-wing ones (`r/EnoughTrumpSpam`). It also contains many generalist, political communities such

---
[6] https://www.reddit.com/r/ListOfSubreddits/wiki/listofsubreddits

Fig. 1: Subgraph of the attention-flow graph with political and conspiratorial communities. Each one is represented by a color: *Alt-right* (green), *Radical Left* (blue), *Liberal* (orange), and *Esoterism* (purple).

as r/politics. This fact is not surprising, since Reddit appears to have a liberal bias in its user base (Shatz, 2017).

Radical left. This community is a meeting point for leftist users. Among its most popular subreddits we find in fact left-wing politicians such as Bernie Sanders (r/SandersForPresident), and more radical communities such as r/ChapoTrapHouse (a popular radical left podcast) and r/LateStageCapitalism (dedicated to anti-capitalistic satire), or intellectuals (r/chomsky).

Alt-right. This community attracts users aligning with the so-called "alt-right" movement. It is mainly composed of right-wing groups (r/TheNewRight) and Donald Trump's supporters (r/The_Donald), but it also hosts subreddits related to right-wing conspiracy theories (Marwick and Partin, 2022) such as r/greatawakening and r/conspiracy (both involved with the QAnon theory). Also some traditional conservative subreddits (such as r/Conservative and r/Republican) are embedded inside this community, albeit their size is smaller.

Community	$|V|$	$|E|$
Alt-Right	128	2079
Radical Left	112	2011
Liberal	113	1517
Esoterism	139	1595

(a) Number of nodes $|V|$ and internal edges $|E|$ of each community.
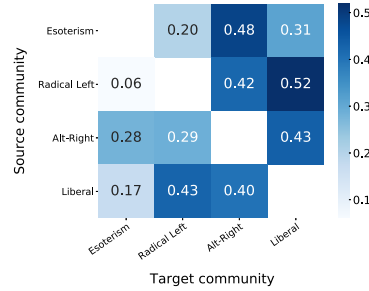
(b) Probability that a random walk on the attention-flow graph starting in source community s reaches a target community t (with $t \neq s$).

Fig. 2: Statistics on the four political communities considered.

Esoterism. This community is not strictly related to politics but its main topic can be viewed as an extension of the classical conspiracy theories to the spiritual sphere. In fact, among its main subreddits we can find groups related to esoterism and occultism such as `r/Paranormal`, `r/Glitch_in_the_Matrix`, and `r/occult`.

The subgraph induced by these four communities in the attention-flow graph includes a total of 492 subreddits and 7202 edges. By using a standard force-directed layout, we depict such a subgraph in Figure 1. Figure 2a shows the size of these communities, which appears to be well-balanced, ranging from 112 to 139 subreddits in each community. Instead, we observe different distances between each pair of clusters. Figure 2b reports the probability of a random walker starting in source community s to reach a target community t (for $s \neq t$). This proximity measure shows that the *Radical Left* and *Liberal* groups are close to each other, while *Esoterism* is more connected to *Alt-Right*.

4.3 Political bridges and gateways

By applying our definitions of *bridge* and *gateway* nodes to the AFG, as given in Section 3.3, we investigate the pathways driving user to different political groups.

Bridges. First, we focus on *bridge nodes* (Definition 1) identified in the four communities of interest. Table 2 shows a ranking of the top subreddits in terms of PPR, for each starting community, considering the attention-flow subgraph induced by these communities. Generally—as shown in Figure 2b—the *Radical Left* and the *Liberal* community are close, as the top bridge for each arrives from the other. The same holds for the *Alt-Right* and the *Esoterism* communities. Nevertheless, while all the top-3 bridges starting in the *Liberal* community arrive in the *Radical Left* one, the opposite is not true, as *Radical Left* has bridges to the *Alt-Right* community, e.g., through `r/WikiLeaks` and `r/ConspiracyII`.

Now, let us focus on the top bridge subreddit for each community, represented in Figure 3. We start with the one with the largest PPR—the subreddit that is more powerful as a bridge between communities—`r/occupywallstreet`, a subreddit

Fig. 3: Same subgraph represented in Figure 1 with emphasis on the top bridge of each community (see Table 2) and their neighbors.

Table 2: Top-3 bridge nodes for each starting community.

Bridge Subreddit	Start	Arrival	PPR ($\times 10^{-4}$)	Bridge Subreddit	Start	Arrival	PPR ($\times 10^{-4}$)
r/numerology	Alt-Right	Esoterism	4.49	r/occupywallstreet	Liberal	Radical Left	7.61
r/jimmydore	Alt-Right	Radical Left	3.44	r/MarchAgainstNazis	Liberal	Radical Left	4.70
r/Lyme	Alt-Right	Esoterism	3.11	r/tulsi	Liberal	Radical Left	4.02
r/YangGang	Radical Left	Liberal	5.54	r/C_S_T	Esoterism	Alt-Right	4.33
r/WikiLeaks	Radical Left	Alt-Right	4.39	r/ConspiracyII	Esoterism	Alt-Right	3.10
r/ConspiracyII	Radical Left	Alt-Right	4.25	r/CriticalTheory	Esoterism	Radical Left	3.00

dedicated to the Occupy Wall Street movement. This movement, born after the financial crisis of 2008, focuses on the problem of growing inequalities in U.S., and argues that politicians—including the liberal administrations—are too influenced by the interests of finance. Our framework reveals that this subreddit acts as a powerful bridge from the *Liberal* community to the *Radical Left* one.

Conversely, the main bridge from the *Radical Left* community is towards the *Liberal* one, and it is r/YangGang. This subreddit is dedicated to supporters of Andrew Yang, a candidate for the U.S. Democratic presidential primaries in 2020. The demographic of Yang's supporters overlaps significantly with Reddit's (Skelley, 2019), which may explain the popularity of the subreddit. While some of Yang's proposals (Universal Basic Income, support for Medicare For All) can be considered left-leaning, Yang does not ascribe himself to the left (e.g., one of his slogans was "not left, not right, forward") and attracts liberal voters (ibid), which explains the role of his supporters' subreddit in the network.

Table 3: Top-3 gateway nodes for each community.

Gateway Subreddit	Community	PPR ($\times 10^{-4}$)	*Gateway* Subreddit	Community	PPR ($\times 10^{-4}$)
r/TheNewRight	Alt-Right	17.99	r/Jewish	Liberal	8.80
r/The_Donald	Alt-Right	16.22	r/AntiTrumpAlliance	Liberal	8.44
r/multiculturalcancer	Alt-Right	14.37	r/Liberal	Liberal	8.31
r/Ani_Communism	Radical Left	13.88	r/numerology	Esoterism	1.83
r/sendinthetanks	Radical Left	13.65	r/occult	Esoterism	1.56
r/GenZedong	Radical Left	13.08	r/cults	Esoterism	1.51

Moving to the right end of the political spectrum, the subreddits r/C_S_T (abbreviation for *Critical Shower Thoughts*) and r/ConspiracyII act as bridges from the *Esoterism* community to the *Alt-Right* one. r/C_S_T collects "politically incorrect" casual thoughts, a type of content known to be connected to the roots of alt-right movements (Massachs et al., 2020). r/ConspiracyII is highly focused on political conspiracy theories, that eventually culminate in supernatural or pseudo-scientific beliefs. Thus, the "bridge" property highlighted by our framework suggests that such subreddits attract users interested in esoterism/occultism towards the alt-right political faction, via involvement in the universe of conspiracy theories. Interestingly, r/ConspiracyII is also one of the top-3 bridges tapping into the *Radical Left* and leading users' attention to the *Alt-Right*. This observation supports the idea that conspiracy theories are "diversionary narratives" (Wu Ming 1, 2021), appealing to a critical view of today's society, but serving the purpose to focus such views within a conservative framework (Jolley et al., 2018). Finally, the last of the four top bridges is r/numerology, an accessible occultist subreddit—interested in mystical and esoteric aspects of numbers—which drives alt-right users towards the *Esoterism* community.

Gateways. To find *gateway nodes* (Definition 2), we compute for all the subreddits in a given community c their PPR outside c (restricted to the 4 communities of interest). However, the random walkers can span the entire network and eventually pass through other communities before reaching the target c. Table 3 reports a summary of the top gateways, which, as expected, differ from bridge nodes. These subreddits are in fact important entry points, in general, for users in a given community. For instance, a user gravitating towards the *Alt-Right* community is more likely to enter it from r/TheNewRight (a subreddit discussing the alt-right movement in general), r/The_Donald (dedicated to Trump's supporters), or r/multiculturalcancer (which attacks the idea of a multicultural society). In particular, r/TheNewRight and r/The_Donald display a much larger PPR than all the other gateways we find: this fact suggests the importance of particular subreddits—and of the Trump presidency—in attracting individuals to the *Alt-Right* community.

These rankings of gateway nodes tell us which are the main entry points of each community for a random walker, but they do not provide us with any information regarding their origin. We investigate this aspect by computing the inflow received by each gateway node from the other (possibly non-politicized) communities in the whole AFG. We compute the inflow of a node s from a

community c as the sum of the weights of the in-links to s from any node in c. We label all neighboring communities by observing its largest subreddits, as done for the four communities of interest; the list of largest subreddits for all the mentioned communities are available in Table S1.

We report some examples of particular interest in Table S2. The subreddit r/Ani_Communism, a powerful gateway node for the *Radical Left* group, self-describes as "*The subreddit of Anime Revolutionaries*". Interestingly, the closest community to r/Ani_Communism is not related to politics, but instead seems to be mainly focused on videogames, including subreddits such as r/LilliaMains (a character of the League of Legends videogame), r/ColdWarZombies (a fantasy war game), and r/PS5restock. As such, its role as an entry point towards *Radical Left* for many Reddit users interested in gaming seems relevant. Similarly, r/The_Donald, gateway for the *Alt-Right* community, displays some of the same communities among its top inflows, such as *Videogames/Fantasy*, but also relevant differences—e.g., a community related to Twitch streamers. Finally, r/numerology, gateway for *Esoterism*, is the most different, as it draws user attention from parts of the network far from politicized communities, such as those related to gaming hardware and crypto-currencies.

5   Conclusion and Future Work

In this work, we analyzed the flow of user attention on Reddit, focusing on the political and conspiratorial communities. We first materialized this flow as a network of subreddits, where links represent user flowing from a subreddit to another. This compact representation allowed us to study four years of user participation to Reddit in a concise format. We did so by relying on the concepts of communities, gateways, and bridges—typically employed in network science—on this attention-flow graph. Thanks to this framework, we identified four main political and conspiratorial communities on Reddit. We then investigated which subreddits have a pivotal role in supporting the flow of users from one community to the other. This way, we discovered, for instance, that the Occupy Wall Street movement supports users flowing from a liberal perspective to a leftist one; and that conspiracy theories helps attracting users to the alt-right community from occultist subreddits, but also from the radical left community.

Our contribution is therefore twofold: on the one hand, we provide a concrete, principled method to study user attention flow on Reddit from a network perspective; on the other hand, our results help to shed light on the topic of opinion formation. Still, our work has limitations: first, our approach is tailored on the community-based Reddit ecosystem, and not immediately applicable to other social media; second, we rely on posts as a proxy for participation, since there is no data for subscriptions. Nevertheless, we are able to identify how Reddit supports the growth of conspiratorial communities, and what is the role of such communities in the wider political ecosystem. Such empirical findings can form the basis for further studies on the political movements of our times, as well as helping community managers of social media to identify the most harmful communities, that could foster the spread of hoaxes and conspiracy theories.

Bibliography

Abbas, T., Somoano, I.B., Cook, J., Frens, I., Klein, G.R., McNeil-Willson, R.: The buffalo attack–an analysis of the manifesto (2022), URL https://icct.nl/publication/the-buffalo-attack-an-analysis-of-the-manifesto/ (Cited on 2)

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., Blackburn, J.: The pushshift reddit dataset. In: Proceedings of the international AAAI conference on web and social media, vol. 14, pp. 830–839 (2020) (Cited on 4)

Boxell, L., Gentzkow, M., Shapiro, J.M.: Is the internet causing political polarization? evidence from demographics. Tech. rep., National Bureau of Economic Research (2017) (Cited on 1)

Bruns, A.: Gatewatching, not gatekeeping: Collaborative online news. Media International Australia **107**(1), 31–44 (2003) (Cited on 2)

Burt, R.S.: Structural holes: The social structure of competition. Harvard University Press (1992) (Cited on 2)

Burt, R.S.: Structural holes and good ideas. American journal of sociology **110**(2), 349–399 (2004) (Cited on 2)

Cai, B., Decker, S., Zheng, C.: The migrants of reddit: An analysis of user migration effects of subreddit bans. preprint (2019) (Cited on 3)

Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., Gilbert, E.: You can't stay here: The efficacy of reddit's 2015 ban examined through hate speech. Proceedings of the ACM on Human-Computer Interaction **1**(CSCW), 1–22 (2017) (Cited on 3)

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., Starnini, M.: The Echo Chamber Effect on Social Media. PNAS: Proceedings of the National Academy of Sciences **118**(9), e2023301118 (2021), https://doi.org/10.1073/pnas.2023301118, URL http://www.pnas.org/lookup/doi/10.1073/pnas.2023301118 (Cited on 1)

Datta, S., Adar, E.: Extracting inter-community conflicts in reddit. In: Proceedings of the international AAAI conference on Web and Social Media, vol. 13, pp. 146–157 (2019) (Cited on 3)

Davies, C., Ashford, J., Espinosa-Anke, L., Preece, A., Turner, L., Whitaker, R., Srivatsa, M., Felmlee, D.: Multi-scale user migration on reddit. AAAI (2021) (Cited on 3)

Enli, G.: Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. European journal of communication **32**(1), 50–61 (2017) (Cited on 1)

Fabbri, F., Wang, Y., Bonchi, F., Castillo, C., Mathioudakis, M.: Rewiring what-to-watch-next recommendations to reduce radicalization pathways. In: Proceedings of the ACM Web Conference 2022, pp. 2719–2728 (2022) (Cited on 2, 3)

Garimella, K., De Francisci Morales, G., Gionis, A., Mathioudakis, M.: Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of

Bipartisanship. In: WWW'18: The Web Conference, pp. 913–922 (2018a) (Cited on 2)

Garimella, K., De Francisci Morales, G., Gionis, A., Mathioudakis, M.: Quantifying Controversy on Social Media. ACM Transactions on Social Computing **1**(1), 3 (2018b) (Cited on 6)

George, R., Shujaee, K., Kerwat, M., Felfli, Z., Gelenbe, D., Ukuwu, K.: A comparative evaluation of community detection algorithms in social networks. Procedia Computer Science **171**, 1157–1165 (2020) (Cited on 7)

Guimera, R., Sales-Pardo, M., Amaral, L.A.N.: Modularity from fluctuations in random graphs and complex networks. Physical Review E **70**(2), 025101 (2004) (Cited on 5)

Horta Ribeiro, M., Jhaver, S., Zannettou, S., Blackburn, J., Stringhini, G., De Cristofaro, E., West, R.: Do platform migrations compromise content moderation? evidence from r/the_donald and r/incels. Proceedings of the ACM on Human-Computer Interaction **5**(CSCW2), 1–24 (2021) (Cited on 3)

Iyengar, S., Sood, G., Lelkes, Y.: Affect, not ideologya social identity perspective on polarization. Public opinion quarterly **76**(3), 405–431 (2012) (Cited on 1)

Jolley, D., Douglas, K.M., Sutton, R.M.: Blaming a few bad apples to save a threatened barrel: The system-justifying function of conspiracy theories. Political Psychology **39**(2), 465–478 (2018) (Cited on 11)

Karrer, B., Newman, M.E.: Stochastic blockmodels and community structure in networks. Physical review E **83**(1), 016107 (2011) (Cited on 5)

Klein, C., Clutton, P., Dunn, A.G.: Pathways to conspiracy: The social and linguistic precursors of involvement in reddit's conspiracy theory forum. PloS one **14**(11), e0225098 (2019) (Cited on 2)

Kubin, E., von Sikorski, C.: The role of (social) media in political polarization: a systematic review. Annals of the International Communication Association **45**(3), 188–206 (2021) (Cited on 1)

Lewin, K.: Frontiers in group dynamics: Ii. channels of group life; social planning and action research. Human relations **1**(2), 143–153 (1947) (Cited on 2)

Mamié, R., Horta Ribeiro, M., West, R.: Are anti-feminist communities gateways to the far right? evidence from reddit and youtube. In: 13th ACM Web Science Conference 2021, pp. 139–147 (2021) (Cited on 2)

Marwick, A.E., Partin, W.C.: Constructing alternative facts: Populist expertise and the qanon conspiracy. New Media & Society p. 14614448221090201 (2022) (Cited on 8)

Massachs, J., Monti, C., De Francisci Morales, G., Bonchi, F.: Roots of Trumpism: Homophily and Social Feedback in Donald Trump Support on Reddit. In: WebSci '20: 12th International ACM Web Science Conference, pp. 49–58, WebSci (2020) (Cited on 2, 3, 11)

Neo, L.S.: An internet-mediated pathway for online radicalisation: Recro. In: Violent Extremism: Breakthroughs in Research and Practice, pp. 62–89, IGI Global (2019) (Cited on 2)

Peixoto, T.P.: Nonparametric bayesian inference of the microcanonical stochastic block model. Physical Review E **95**(1), 012317 (2017) (Cited on 5)

Phadke, S., Samory, M., Mitra, T.: What makes people join conspiracy communities? role of social factors in conspiracy engagement. Proceedings of the ACM on Human-Computer Interaction **4**(CSCW3), 1–30 (2021) (Cited on 2)

Phadke, S., Samory, M., Mitra, T.: Pathways through conspiracy: The evolution of conspiracy radicalization through engagement in online conspiracy discussions. In: Proceedings of the International AAAI Conference on Web and Social Media, vol. 16, pp. 770–781 (2022) (Cited on 2, 3)

Quattrociocchi, W., Scala, A., Sunstein, C.R.: Echo chambers on facebook. Available at SSRN 2795110 (2016) (Cited on 2)

Ribeiro, M.H., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., Long, S., Greenberg, S., Zannettou, S.: The evolution of the manosphere across the web. arXiv preprint arXiv:2001.07600 (2020a) (Cited on 3)

Ribeiro, M.H., Ottoni, R., West, R., Almeida, V.A., Meira Jr, W.: Auditing radicalization pathways on youtube. In: Proceedings of the 2020 conference on fairness, accountability, and transparency, pp. 131–141 (2020b) (Cited on 2)

Shatz, I.: Fast, free, and targeted: Reddit as a source for recruiting participants online. Social Science Computer Review **35**(4), 537–549 (2017) (Cited on 8)

Skelley, G.: What We Know About Andrew Yang's Base. Five Thirty Eight. https://fivethirtyeight.com/features/what-we-know-about-andrew-yangs-base/ (2019), [Online; accessed 18-August-2022] (Cited on 10)

Stroud, N.J.: Polarization and partisan selective exposure. Journal of communication **60**(3), 556–576 (2010) (Cited on 1)

Thorson, K., Wells, C.: Curated flows: A framework for mapping media exposure in the digital age. Communication Theory **26**(3), 309–328 (2016) (Cited on 2)

Waller, I., Anderson, A.: Quantifying social organization and political polarization in online platforms. Nature **600**(7888), 264–268 (2021) (Cited on 3)

Welbers, K., Opgenhaffen, M.: Social media gatekeeping: An analysis of the gatekeeping influence of newspapers' public facebook pages. New Media & Society **20**(12), 4728–4747 (2018) (Cited on 2)

West, S.M.: Raging against the machine: Network gatekeeping and collective action on social media platforms. Media and Communication **5**(3), 28–36 (2017) (Cited on 2)

Wu Ming 1: La Q di Qomplotto: QAnon e dintorni: come le fantasie di complotto difendono il sistema. Edizioni Alegre (2021), URL https://edizionialegre.it/notizie/the-q-in-conspiracy/ (Cited on 11)

Zannettou, S., Caulfield, T., De Cristofaro, E., Kourtelris, N., Leontiadis, I., Sirivianos, M., Stringhini, G., Blackburn, J.: The web centipede: understanding how web communities influence each other through the lens of mainstream and alternative news sources. In: Proceedings of the 2017 internet measurement conference, pp. 405–417 (2017) (Cited on 3)

Appendix

A   Weight rescaling
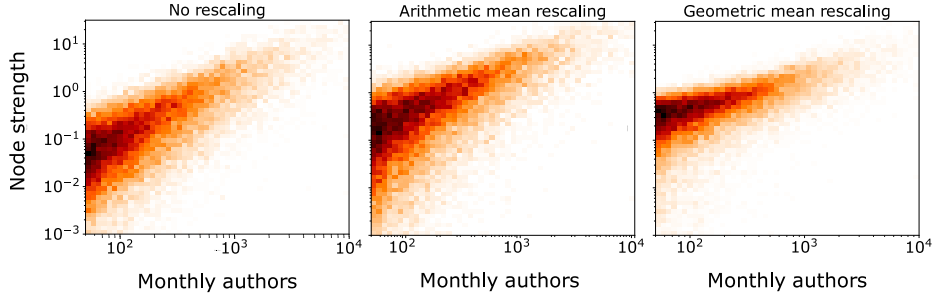
Fig. S1: Distribution of the subreddits in the space defined by their number of monthly users and their (rescaled) node strength in the attention-flow graph, i.e., the sum of the weight of all of their links.

In order to define the weights of the links in the attention-flow graphs, we wish to focus on links that are deviating from the most popular subreddits. E.g., the constant flow of users towards a generalist subreddit such as `r/pictures` is not as interesting for our purposes as the flow between two topical communities. For this reason, we rescale the weight w of a link between two subreddits (s, s') in order to reduce the correlation of w with the number of users of s and s'. Thanks to our framework, we are able to perform such normalization using only information coming from the node weights themselves. In fact, we rescale each weight by the mean of the total weights of the links incident on the head and the tail of each link. We compare two types of means: arithmetic and geometric. Figure S1 reports this comparison, clearly showing that the geometric mean succeeds in this goal, as it dampens the dependency between the resulting subreddit node strength (i.e., the sum of the weights w of all of their links in the attention flow graph) and the number of monthly users in each subreddit.

B   Pruning thresholds

As shown in table Table 1, community detection algorithm benefits from edge pruning. In fact, the weights distribution of the attention-flow graph, plotted in Figure S2, spans over several orders of magnitudes and the considerable amount of lightly weighted links adds a background noise in the network, which is detrimental for the community detection.
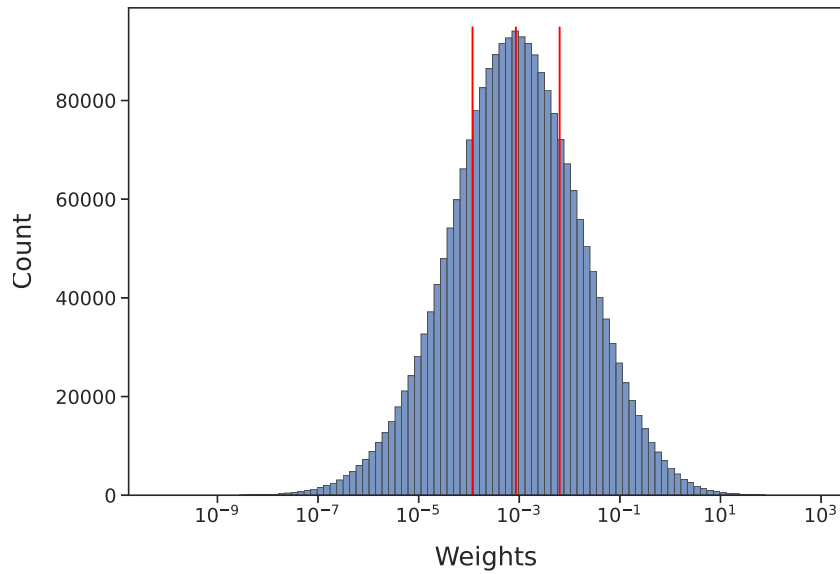
Fig. S2: Weights distribution of the attention-flow graph, with the investigated pruning thresholds at 25th, 50th and 75th percentiles indicated by red vertical lines.

C   Mentioned communities

In Table S1, we report the top-10 largest subreddits by userbase for all the mentioned communities: the four politicized communities of interest, and their neighboring communities mentioned in Table S2.

Table S1: Top-10 subreddits by userbase for mentioned communities.

Label	Top 1 subreddit	Top 2 subreddit	Top 3 subreddit	Top 4 subreddit	Top 5 subreddit	Size
Alt-Right	The_Donald	worldpolitics	conspiracy	greatawakening	CBTS_Stream	128
Esoterism	Dreams	Incels	Paranormal	witchcraft	Glitch_in_the_Matrix	139
Liberal	Coronavirus	PoliticalHumor	politics	CoronavirusMemes	China_Flu	113
Radical Left	ChapoTrapHouse	SandersForPresident	SelfAwarewolves	PresidentialRaceMemes	LateStageCapitalism	112
LGBTQ+	TeenAmIUgly	pointlesslygendered	TransTryouts	latebloomerlesbians	okbuddyhetero	155
Entertainment	KeanuBeingAwesome	frenworld	whatsthatbook	ConsumeProduct	okboomer	293
Videogames/Fantasy	keoXer	MarioJudah	EajParkOfficial	PS5restock	ainbowroad	1291
Gaming Hardware	PS5	XboxSeriesX	buildapc	Amd	nvidia	110
Twitch	LaStalla	EIDL	BPTmeta	Homyatol	Panetty	674
Satire	mildly_interesting	TheCircleTV	covfefe	PolHumor	net_neutrality	198
Crypto-currencies	CryptoCurrency	Bitcoin	btc	ethtrader	ethereum	13
Anime	Genshin_Memepact	codevein	Helltaker	DevilArtemisUniverse	JuJutsuKaisen	285
Self-help	AmItheAsshole	relationship_advice	depression	NoFap	Tinder	104

Label	Top 6 subreddit	Top 7 subreddit	Top 8 subreddit	Top 9 subreddit	Top 10 subreddit
Alt-Right	Conservative	WatchRedditDie	JordanPeterson	NoNewNormal	Libertarian
Esoterism	DecidingToBeBetter	IWantToLearn	step1	spirituality	tarot
Liberal	YangForPresidentHQ	CoronavirusUS	HongKong	EnoughTrumpSpam	JoeBiden
Radical Left	ABoringDystopia	OurPresident	TheRightCantMeme	WayOfTheBern	VaushV
LGBTQ+	femboyhooters	questions	comingout	TheClickOwO	aromantic
Entertainment	Boomerhumour	shower_thoughts	screenshots	opienanthony	disenchantment
Videogames/Fantasy	CanadaPS5	Pxlsspace	PlayGodfall	theblackvoid	usps_complaints
Gaming Hardware	OculusQuest	battlestations	hardwareswap	oculus	GlobalOffensiveTrade
Twitch	USCIS	DarioMocciaTwitch	interestingnewsworld	pga2k21	barexam
Satire	podemos	zuckmemes	AskABrit	EpsteinMemes	NickerNation
Crypto-currencies	CryptoMarkets	Bitcoincash	BlockChain	CryptoCurrencies	dashpay
Anime	TowerofGod	SDSGrandCross	PampamilyangPaoLUL	KanojoOkarishimasu	obeyme
Self-help	ApplyingToCollege	offmychest	CasualConversation	stopdrinking	ADHD

Table S2: Top-5 communities by normalized inflow for three gateway nodes.

(a) r/Ani_Communism

Community	Inflow
Videogames/Fantasy	0.699
Radical Left	0.201
LGBTQ+	0.045
Satire	0.039
Anime	0.004

(b) r/The_Donald

Community	Inflow
Alt-Right	0.411
Videogames/Fantasy	0.133
Satire	0.081
Entertainment	0.059
Twitch	0.049

(c) r/numerology

Community	Inflow
Satire	0.604
Crypto-currencies	0.256
Esoterism	0.075
Self-help	0.039
Gaming Hardware	0.016