Online Social Networks and Media 43–44 (2024) 100292

Contents lists available at ScienceDirect

Online Social Networks and Media

journal homepage: www.elsevier.com/locate/osnem

Measuring centralization of online platforms through size and interconnection of communities

Milo Z. Trujillo *, Laurent Hébert-Dufresne, James Bagrow

Complex Systems Center, University of Vermont, 05401 VT, Burlington, United States

ARTICLE INFO

Keywords:
Social centralization
Structural influence
Insularity

ABSTRACT

Decentralization of online social platforms offers a variety of potential benefits, including divesting of moderator and administrator authority among a wider population, allowing a variety of communities with differing social standards to coexist, and making the platform more resilient to technical or social attack. However, a platform offering a decentralized architecture does not guarantee that users will use it in a decentralized way, and measuring the centralization of socio-technical networks is not an easy task. In this paper we introduce a method of characterizing inter-community influence, to measure the impact that removing a community would have on the remainder of a platform. Our approach provides a careful definition of "centralization" appropriate in bipartite user-community socio-technical networks, and demonstrates the inadequacy of more trivial methods for interrogating centralization such as examining the distribution of community sizes. We use this method to compare the structure of five socio-technical platforms, and find that even decentralized platforms like Mastodon are far more centralized than any synthetic networks used for comparison. We discuss how this method can be used to identify when a platform is more centralized than it initially appears, either through inherent social pressure like assortative preferential attachment, or through astroturfing by platform administrators, and how this knowledge can inform platform governance and user trust.

Online social spaces are vulnerable to centralized authorities making decisions that negatively affect the community. In 2022, the Software Freedom Conservancy recommended that all developers migrate their projects away from GitHub [1], after Microsoft bought the software development collaboration platform and used open source projects as training data for their commercial CoPilot software, in violation of open source licenses and community standards. The same year, users and advertisers departed Twitter after its purchase by Elon Musk and subsequent changes in community policy and staffing, including firing content moderators [2] and reinstating a number of accounts banned for violating the platform's hateful content and harassment policies [3]. Reddit moderators have historically engaged in blackouts to protest administrative policies [4], and these trends are ongoing; in June, 2023, Reddit announced plans to begin charging for API access, sparking warnings from scientists [5], outrage among users, and a protest across nearly 9000 subreddits, the long-term effects of which remain to be seen. As users express dissatisfaction with platform administrators, they have sought alternative platforms without centralized control, leading to the rapid growth of "federated" platforms like Mastodon [6] and Bluesky.[1] Alternatively, other users have promoted self-hosted platforms, such as independently operated git servers, or peer-to-peer hosting solutions such as the Interplanetary File System (IPFS) or web-torrent video hosting software PeerTube. Some deplatformed users have also responded by creating close facsimiles of existing centralized platforms with extremely permissive content-policies, frequently called "alt-tech" platforms [7].

What exactly is "centralization" in an online social network? Does it describe ownership of the platform? Its technical infrastructure? The creation and enforcement of community norms? The distribution of activity and reach of content producers? Centralization has long been ill-defined by academics [8], and "decentralization" joins as a widely-used but contextually redefined term today [9]. Of particular interest to us is a notion of group social influence: How much does one community impact others across a platform? For example, how independent are subreddits on Reddit, and how closely interlinked are Mastodon instances, the nascent "decentralized Twitter alternative?" Our goal is to measure the influence of a socio-technical platform's subcommunities on their peers, providing a mesoscale metric to quantify centralization at an inter-group level.

* Corresponding author.
E-mail addresses: milo.trujillo@uvm.edu (M.Z. Trujillo), laurent.hebert-dufresne@uvm.edu (L. Hébert-Dufresne), james.bagrow@uvm.edu (J. Bagrow).

[1] Bluesky is still in beta, and while the protocol is federated, only one instance exists at the time of writing.

https://doi.org/10.1016/j.osnem.2024.100292
Received 24 April 2024; Received in revised form 29 July 2024; Accepted 4 October 2024
Available online 25 October 2024

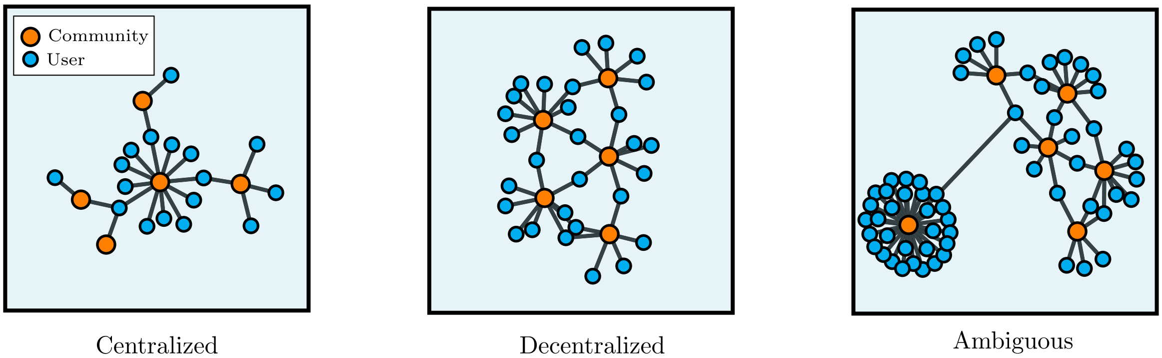


Fig. 1. The influence of a community is tied to both its size and topological role in a network. In the centralized network, the orange community at the center both has the largest population of blue users, and serves as a bridge between four other communities. In the decentralized example, communities are of variable size, but none have a pivotal position to influence their peers. In the ambiguous case, one community is much larger, but the remaining network matches the "decentralized" example. Neither a distribution of community sizes nor purely structural measurements like betweenness centrality or graph conductance adequately capture this notion of community-level influence.

Measuring group-level influence has applications in content moderation, platform governance, and public awareness of administrative behavior. First, it allows administrators, moderators, and community organizers to identify and proactively avoid risks to community welfare. For example, if Mastodon's goal is to create a decentralized Fediverse then measuring the influence of a large instance can inform decisions on when to close registration on that instance, or stop recommending it on new-user onboarding websites like , to direct new users to more diverse instances. If administrators of smaller instances want to mitigate the viral spread of information from influential instances, they can de-emphasize posts from the larger instance, for example by hiding them from the federated feed. This aligns with recent proposals to design social media for abusability, by making design choices that limit usability for a minority of users to protect usability for the majority [10].

Next, measuring community influence allows platform users to identify when administrators are engaging in "decentralization astroturfing". Some administrators misrepresent the level of decentralization or community self-governance on their platforms, allowing them to abdicate responsibility for community moderation and social policy. For example, Bluesky has no dedicated moderation or trust and safety team, because they publicly aspire to provide tools and protocols for communities to self-govern [11]. However, after two years and almost three million users, Bluesky's federated protocol only has one server, administered by Bluesky employees, who willingly or not have immense influence over acceptable speech on their platform. By contrast, Mastodon has thousands of federated instances, each with their own moderators and content policies. However, if instance administrators wish to federate with the largest three instances, containing more than half the Mastodon population, they must have a compatible content policy, enforcing an implicit monoculture. These patterns can be identified by measuring the number of communities on a platform, and the influence that the largest communities have over their peers.

One common approach to measuring community-level centralization is through community size-distribution. If a small oligarchy of Mastodon instances dwarf the population sizes of their peers, then one could presume that the platform is centralized around these instances. Indeed, several prior studies on Mastodon use community size disparity as a starting point, or presuppose that the largest instances are the most significant and focus their study on the largest communities [12–15]. While the community size distribution is related to centralization, assuming they are the same precludes the possibility that a collection of many smaller instances may be more influential than the few largest, or that the influence of the largest instances may not be directly related to their size.

We reject the assertion that the largest communities must be the most significant, or that their size alone implies centralization, on the grounds that community size does not correlate with the number of cross-community links in observed real-world networks. In fact, our results show multiple platforms where the largest communities are *not* well integrated with the platform as a whole (discussed in results, especially Fig. 6), allowing a more decentralized network of communities to exist outside of the largest groups. Under this view, the largest communities would be the most significant only when they also act as important information bottlenecks for the entire system.

To illustrate this discrepancy, consider Fig. 1. In the centralized panel the largest community serves as a central hub, connecting several smaller communities together through shared membership. In the decentralized panel community size is normally distributed, and no community has a pivotal role as a bridge between its peers. Community size-distribution and graph-centric metrics like betweenness-centrality would agree that the former network is centralized, while the latter is decentralized. However, the third ambiguous panel presents a more complex scenario: the community size distribution is highly unbalanced, but the largest community has almost no impact on the remainder of the network. The largest community has a high betweenness-centrality because of its pivotal role in connecting so many users to the rest of the graph, but it has a long path distance from users in other communities and does not serve as a bridge between communities, and so betweenness-centrality does not match our intuition that the largest community has a small role in the rest of the network.

We propose a definition of centralization meant to capture the alignment between rankings of community size and information bottlenecks. To do so, we combine theoretical ideas from graph theory on bottlenecks and applied concepts from network science about network robustness. Our metric then measures how removing a community would impact users within remaining communities, based on the number of "bridges" between communities. We study a variety of real and simulated networks with this method to examine platform behavior under a range of conditions, and we compare our metric to existing measurements of centralization and network "bottlenecks". Finally, we discuss how this work contributes to broader discussions of centralization online, and how techniques like ours can be extended with richer interaction data.

1. Prior work

Centralization of online platforms is sometimes defined in terms of decision-making power, or who has the authority to make what kinds of decisions about the use of the platform. This definition can be traced to Elinor Ostrom's work on Institutional Analysis and Development [16], which describes "layers" of decisions, from operational rules (elementary actions any user can perform), to collective rules (the context in which users operate and interact, such as the Twitter feed or the Amazon marketplace), to constitutional rules (the "meta" rules through which the system changes itself). Modern research on platform design often assesses who has decision-making power, and what levers of change are available to different categories of participants [17,18].

While qualitative studies examine power structures through analyzing governance and rule sets [19,20], network science infers structure through the observed interactions between humans [8,21]. We quantify centralization using attributes that fall into three categories: vertex-level attributes, cluster-level attributes, and graph-level attributes. Vertex-level attributes like betweenness centrality [8] or eigenvector centrality [22] measure the prominence of a particular node in terms of how well it is connected to its peers, or how many paths flow through the node. Cluster-level attributes describe groups of vertices, such as the size of the population that contains a particular attribute, or the assortativity describing how likely vertices with a particular attribute are to be connected to one another. Graph-level attributes describe aspects that span the entire network, including diameter, density, and graph conductance [23]. Quantifiability should not be conflated with objectivity; the modeling choice of what entities are included as vertices and what relationships are represented as edges or attributes presupposes what can be considered influential or centralized [24].

Another thread of research tries to join the social theory of centralization and graph theoretical metrics. [20] distinguish between the technical underpinnings of a network and its social layers, focusing on community-run moderation in infrastructure-centralized (Slack, Discord) and self-hosted (Minecraft) services. Prior Mastodon research also bridges this gap, including both geographic and data-center distribution of instances [12], important for understanding resiliency to disruption or power-outage. This approach aligns with notions of network robustness where centralization can be measured by how a network breaks down under targeted pruning of central nodes [25]. Other studies on Mastodon also integrate its social interaction graph [15], important for understanding the influence of sub-communities and their administrators on discourse. Studies on the social structure of Mastodon primarily focus on individual-centralization, such as a “border-index” of what fraction of a user’s neighbors are on a foreign instance [14] and whether some users serve as critical bridges for information flow between instances [26], or community-centralization, such as how clustering coefficients differ between communities (instances) [13]. Our work intends to add to these options, by considering both a community-level centralization metric of how much influence one community has on the broader platform, and a graph-level centralization score of how quickly a network deteriorates as its largest communities are removed, indicating how much it tends towards monopoly or oligopoly.

Recent social media studies highlight the difference between size and importance, demonstrating the need for a better understanding of smaller-yet-influential subgroup dynamics. For example, [10] identifies a single low-follower Twitter user that has a disproportionate influence on national COVID-19 discussion by starting arguments in the replies to the tweets of public officials. Despite not fitting the typical high-follower and high-engagement profile of an “influencer” or “Internet celebrity”, this account’s behavior and structural role adjacent to prominent accounts leads to outsized impact. At a regional scale, [27] focuses on sentiment-spreading dynamics between Japanese prefectures, proposing a causal measure of social influence based on correlated sentiment between geographic regions in a forecasting model. Other researchers have focused on cross-platform misinformation campaigns, including [28], showing how bad actors can coordinate across YouTube, Facebook, and Twitter to thwart content moderation. We believe that measuring community-level influence through observed social interlinking will contribute to this conversation on disproportionate influence at multiple scales.

2. Methods and materials

In the following sections we introduce our metric and two data sets: five real world networks that encompass a breadth of configurations, and a set of common synthetic networks.

2.1. Measuring centralization: Disruption curves

Prior studies on centralization of social networks often focus on graph-level attributes such as detecting components, the size of the giant component, modularity, density, degree distribution [29]. Others may use “bottleneck” metrics like graph conductance [23] to identify bridges and key clusters. These metrics are most appealing in unipartite settings where the structure of the network is not prescribed. However, we focus on bipartite graphs where communities are well defined, such as subreddits, Mastodon instances, or newsgroups. In these contexts, we are not attempting to infer the number or boundaries of communities, but to measure how influential the known communities are on their neighbors. The size distribution of communities tells us how large a subgroup is, but does not capture the overlap between communities. A graph-wide modularity score describes how well-partitioned the graph is into clusters, and so approximates how insular communities are, but cannot provide more nuance as to whether the largest communities are more integrated than smaller ones, whether small communities are well connected to larger peers but not to each other, or other topological features.

We propose that the influence of a community should be measured in terms of how users outside the community would be impacted by its absence. In other words, a community’s influence should be proportional not to its size, but to the number of bridges between it and other communities. Or, in graph theoretic terms, what fraction of edges would be cut by removing a community, not counting users that do not participate outside the community. More succinctly, “what percentage of edges from surviving vertices would be cut by removing a community?”

We measure disruption cumulatively, rather than discretely per-community. This allows us to answer questions like “how influential are the largest three communities on the rest of the platform?” Since “oligarchies” of large and densely interconnected communities may be common, a cumulative metric is more useful than measuring the influence of a single community on the rest of the oligarchy.

Formally, we define a set of communities that are being cut, A, with associated edges $|A|$. Each user has a set of edges to one or more communities. If users *only* have edges to communities in A, then the user is removed along with A. Surviving users with an edge to at least one remaining community are denoted S, with total edges $|S|$, and edges to cut communities in A denoted ∂S. The disruption curve is calculated as $\partial S/|S|$. This notation was chosen for its similarity to the Cheeger number [23], stressing how our metric measures the alignment of community size and information bottleneck. Familiarity with the Cheeger number is unnecessary to understand our metric; we make a detailed comparison in Section 3.3, but in summary the Cheeger number is a single-valued metric combining community detection and graph conductance, while disruption is a cumulative metric that utilizes available bipartite knowledge.

```
disruption = []
for c in communities:
  remaining = 0
  original = 0
  removeCommunity(c)
  for user in users:
    if degree(user) > 0:
      remaining += degree(user)
      original += originalDegree(user)
  disruption += [1-(remaining/original)]
```

Listing 1: Pseudocode for disruption algorithm

We additionally outline the algorithm as pseudocode in Listing 1:. We recommend caching the size of the smallest community that each user participates in, and pre-sorting users by the order in which they will be removed, to avoid computationally expensive references to a graph or adjacency matrix during each removal-step.

Table 1
Definitions of communities and edges for each platform examined.

Platform	Community definition	Edge definition	Edge weight
Mastodon	Mastodon Instances	Between each user and every instance on which they follow users	The number of users followed on an instance
Penumbra	A git server	Between a user (identified by email) and each server on which they have contributed to a repository	The number of repositories committed to on each server
BitChute	BitChute channels	Between each user and every channel they have commented on videos from	The number of comments made
Voat	A Voat "subverse"	Between each user and subverses they have participated in	Number of comments made in a subverse
Usenet	A Usenet newsgroup	Between each user and every newsgroup they have posted in	The number of posts made

Table 2
Population size of each network in terms of community count, user count, and relationship edge count, before compressing duplicate edges into weighted edges.

Platform	Comms.	Users	Edges
Mastodon	3825	479,425	5,649,762
Penumbra	841	41,619	108,038
BitChute	29,686	299,735	11,549,058
Voat	7515	3,624,486	16,263,309
Usenet	333	2,080,335	58,133,610

Our disruption curve metric is intended for bipartite networks, where communities are clearly distinguishable with ground-truth definition and users can participate in multiple communities. However, some consideration is also given to applying our metric to unipartite settings in the synthetic network section.

We plot disruption similarly to a cumulative disruption function (CDF), where the x-axis represents the number of communities removed, cumulatively ordered by degree, and the y-axis represents the fraction of edges from surviving users that have been cut. In other words, the x-axis is the size of A as a fraction of all communities in the graph, and the y-axis is $\partial S/|S|$, where both the numerator and denominator are dependent on $|A|$.

While disruption curves offer insight into the role of the largest communities on a platform, some readers may desire a scalar summary statistic to describe how "centralized" a platform is under our metric. For these scenarios we recommend calculating the area under the curve, as shown in Figs. 4(b) and 5(b). We calculate the Distruption AUC (DAUC) using a trapezoidal approximation in logarithmic space, by re-normalizing the x-axis as:

$$\frac{\log_{10}(\text{number of removed communities})}{\log_{10}(\text{total number of communities})}$$

We measure in logarithmic space because most disruption curves have a long tail: there is a significant impact as the largest communities are removed first, which we are primarily interested in, and there is typically less change to disruption as smaller communities are removed later, especially once the giant component is fractured. Measuring in linear space allows the long-tail to heavily influence the DAUC, while measuring in logarithmic space emphasizes the role of the largest communities.

For some synthetic networks it is possible to write a closed-form integral for the disruption curve, but because this is not possible for real-world data, we use a trapezoidal approximation for all real- and synthetic-networks for consistency.

2.2. Mathematical analysis of disruption

We can analyze the expected behavior of disruption curves using random bipartite networks parameterized by their joint-degree distribution. This approach fixes the distribution $\{g_m\}$ of users part of m communities, the distribution $\{p_n\}$ of community size n, and the joint-distribution $P_{n,m}$. Beyond these constraints, we assume the networks to be very large and fully random.

We can calculate the *expected* disruption $D(n)$ involved when removing communities of size $n' < n$. Disruption is given by the number of edges that belong to communities of size n minus the fraction u_n of those that are the sole edge of the corresponding users (since these users are removed in the pruning) divided by the number of edges belonging to communities of size equal or smaller than n minus the $u_n n p_n$ users removed. We write:

$$D(n) = \frac{\overbrace{np_n}^{\text{Edges to comms. of size n}} - \overbrace{u_n n p_n}^{\text{Edges to removed users}}}{\underbrace{\sum_{n' \leq n} n' p_{n'}}_{\text{Edges to comms. n or smaller}} - u_n n p_n} . \tag{1}$$

The quantity u_n is defined as the probability that a random user of a community of size n has no community smaller than n:

$$u_n = \sum_m \overbrace{\frac{P_{n,m}}{\sum_{m'} P_{n,m'}}}^{\text{Fraction of users in comm. size n that have m edges}} \Bigg(\underbrace{\frac{\sum_{n' \geq n} P_{n',m}}{\sum_{n'} P_{n',m}}}_{\text{Fraction of users with m edges in comms. larger than size n}} \Bigg)^{m-1} . \tag{2}$$

In a simple experiment, we create a random Erdős-Rényi-like bipartite network and correlated equivalent networks with the same degree distributions and variable community-user degree matrices $P_{n,m}$. The random network has a simple $P^{\text{rand}}_{n,m} \propto np_n m g_m$ (normalized). We also calculate the maximally assortative $P^{\max}_{n,m}$ by assigning users with highest degrees $m_{\max}$ to the largest communities, and maximally disassortative $P^{\min}_{n,m}$ by assigning users with the lowest degree to the largest communities.

Using Eq. 1 on networks linearly interpolating between $P^{\max}_{n,m}$, $P^{\text{rand}}_{n,m}$ and $P^{\min}_{n,m}$, we find that positive user-community degree correlations increase disruption and therefore *centralizes* the resulting socio-technical network. Conversely, negative correlations decreases disruption and *decentralizes* the network. We thus know that dispersion curves will be affected by network structure beyond its distribution of community sizes.

2.3. Real-world network data

We analyze five real-world datasets, each describing online social interactions in bipartite configurations where vertices represent either "users" or "communities". We utilize a 2021 scrape of the Mastodon follow graph [13]. Mastodon is a Twitter alternative where users are located on one of thousands of "instances", which are Twitter-like servers with their own administrators and content policies. However, Mastodon users can follow users on other instances, exchanging content between the two communities, so long as the servers are "federated" (willing to exchange content). For a second example of a platform with distributed servers, we include the Penumbra of open-source [30], a data set of independent git servers (not GitHub or GitLab), and users that contribute to repositories on each server. We also include an interaction graph from BitChute [31], an alt-tech YouTube alternative, consisting of users and the channels (video uploaders) whose videos they commented on. We utilize a similar scrape of Voat [32], an alt-tech Reddit alternative active until late 2020, consisting of users and the "subverses" (subreddits) they commented in. We additionally include an archive of Polish Usenet groups [33], providing a much older but similarly structured platform for comparison. Details on the vertex and edge definitions for each network are included in Table 1, and the size of each network is listed in Table 2.

We selected these platforms because they have clear bipartite user and community representations, data is readily available, and each platform is small enough to obtain a nearly-complete sample. Sub-sampling a larger platform like Reddit may miss lower-population or lower-activity sub-communities, and we are particularly interested in the interactions between smaller communities. The resulting dataset encompasses a variety of approaches to hosting and community governance, providing a spectrum of "centralization".

2.4. Ethical considerations

Any method for community detection, or the measuring of community "importance" (in our case, disruptive potential) in a social context implies risk. A bad actor could use such a metric to identify communities with the most reach in an effort to spread misinformation more efficiently, or to identify the smallest set of content moderators that must be influenced to enforce a desired social policy change across a platform. However, the same methods can be used to preemptively identify risk, allowing a platform like Mastodon to proactively take steps to limit their dependence on their largest server instances, for example by closing user registration on their largest instances, or by de-prioritizing those servers on instance-recommendation websites like . We believe the benefits of studying the structure of social media outweigh the risks in this regard.

Concerning the datasets used in this paper, we present only aggregate group behavior to provide insight into social welfare. We do not publish any usernames or study individuals' behavior in-depth, and present only the names of some of the largest communities to help contextualize our findings. We believe this presents a minimal risk to the privacy of users included in our five real-world datasets. Since our real-world data is sourced from prior publications, we are not republishing it with this study, but also do not have the opportunity to further anonymize user data.

2.5. Synthetic network data

To understand disruption curves and contextualize our real-world results, we examine a variety of well understood synthetic network topologies. Small visualizations of these types of networks are shown in Fig. 2, and the methodology and parameters used in our experiments are described in detail below.

First we construct a bipartite star network, as a default example of a network centralized around a single hub. Bipartite Star networks are analogous to a unipartite star network with duplicate edges. Starting with a unipartite star, we replace each edge from the hub to a leaf with a two-path from the hub community to a new "user" vertex, to the leaf community. Duplicate edges from the unipartite hub to leaves are converted into multiple users that share a community, and serve to break ties when pruning communities for disruption curves. In our example plots, we construct a graph with 150 communities and 3000 users, such that every user has an edge to two communities: the central hub, and one other, assigned uniformly. Removing the hub eliminates 50% of all edges, and removing any subsequent communities incurs no additional disruption, because all impacted users will have a degree of zero and be pruned from the graph (see Fig. 5(a)). This graph type is therefore highly centralized but has a decentralized periphery after the removal of the central community, illustrating how different topologies can co-exist in the same network, muddying the definition of "centralization".

We then test disruption on a variety of bipartite networks with power-law degree distributions. We first adapt the Barabási-Albert preferential attachment model to a bipartite setting, initializing a network with 300 empty communities and introducing users that connect to a given community with probability proportional to their size plus one. We also introduce a range of bipartite configuration models: in each, we assign a degree to each community drawn from a power law with a specified γ exponent. For each community, we create edges according to degree, connecting the community to users uniformly randomly without replacement. Therefore, we control for the size of communities, which follow a power-law distribution, but we do not control for the degree of users, which follow a normal distribution, nor do we control for assortativity. Each of these networks produces a curve that slowly decays towards a diagonal, implying that removing the largest communities has some disproportionate impact, after which removing additional communities has a less pronounced result.

We also adapt the Erdős-Rényi model to a bipartite setting by creating vertices for communities and users, then creating all possible edges with a probability p (in our tests, $p = 0.05$), while preserving the bipartite constraint. These networks produce a disruption curve with a second derivative near zero, indicating that most communities have near-equal influence on the population, and so removing the largest communities does not have a much larger impact than removing subsequent communities.

Lastly, we create a bipartite Watts–Strogatz small-world model. We begin by producing a *unipartite* network with desired neighborhood size ($n = 5$) and edge density ($p = 0.05$) parameters. We apply a clustering algorithm (in our examples we used weighted community label propagation) to place each user in one community, we create a vertex for each detected community, and we replace all user-user edges with user-community edges. These networks have the most uniform community size distribution of any we tested, and their disruption curves are similar to those of Erdős-Rényi networks, with slightly more variability. By applying community detection, as discussed here and illustrated in Fig. 3, it is possible to measure disruption in unipartite networks as well as bipartite.

3. Results

We plot the cumulative population size, disruption curve, and disruption AUC for real-world networks in Fig. 4, and plot the same results for synthetic network data in Fig. 5. We first focus on discrepancies between the size distribution and disruption curves for real networks, then return attention to synthetic network data when we examine the role of assortativity.

3.1. Comparison to size distribution

Upon comparing the size distribution and disruption curve in Fig. 4(a), it is apparent that the community size distribution is insufficient to describe the structure of a network. Voat has the most skewed population distribution: almost all users participate in the largest community, yet removing the largest community impacts less than 0.03% of the remaining graph, and only after removing the largest three communities is more than 10% of the graph impacted. Mastodon and BitChute have the next most skewed size distributions, but there is a large distance between the proportional sizes of their largest communities, and almost identical disruption curves as those communities are removed. By population distribution, the Penumbra appears to be more skewed towards its largest git servers than Usenet is towards its largest newsgroups. This is not mirrored in disruption curves, where Usenet has a consistently higher disruption than the Penumbra.

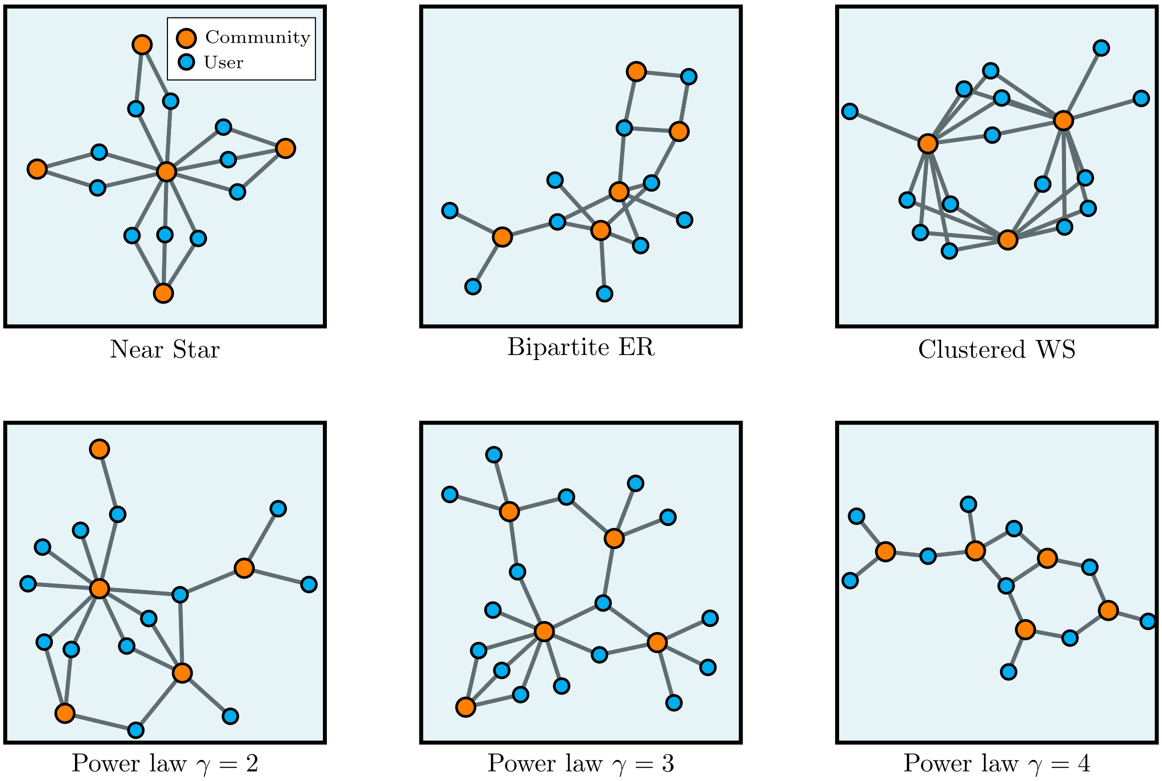


Fig. 2. Example visualizations of synthetic bipartite networks. The near-star approximates a star-graph, but in a bipartite configuration where each node participates in the "hub" and one other community. Bipartite ER is an Erdős-Rényi graph adapted to a bipartite setting. A bipartite Watts–Strogatz model produces a "small-world" graph where each community is close by path length to each other (see also Fig. 3 for more detail). Power-law networks vary from high heterogeneity at low γ where one community is much larger than others, to low-heterogeneity at high γ resembling an Erdős-Rényi graph. For brevity, a visualization of the Barabási-Albert model, an alternative preferential-attachment scheme, is excluded.

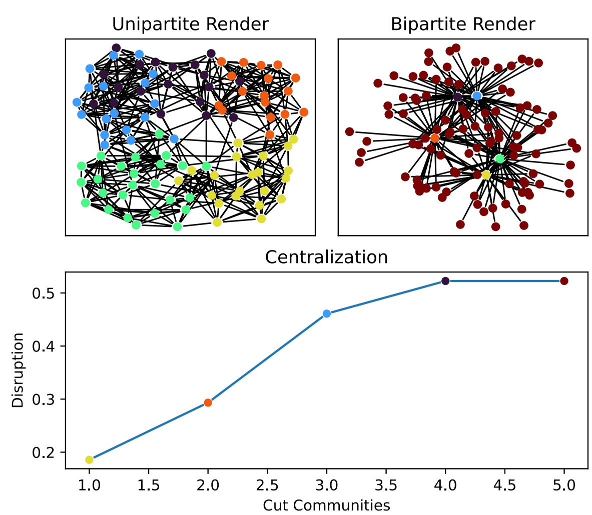
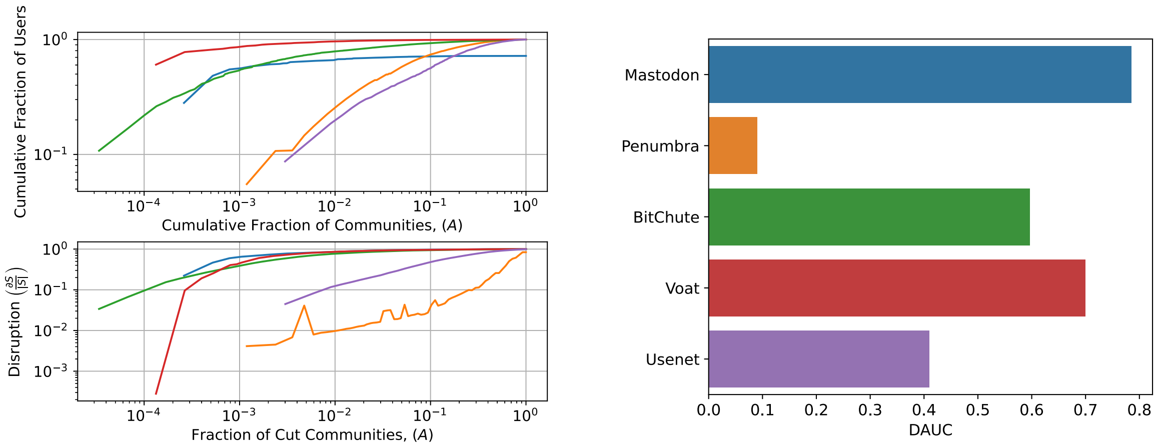


Fig. 3. Example of applying our disruption metric to unipartite graphs by detecting communities on a unipartite small-world network (top-left), converting labeled communities into a bipartite representation (top-right), and running our influence metric on the bipartite graph (bottom).

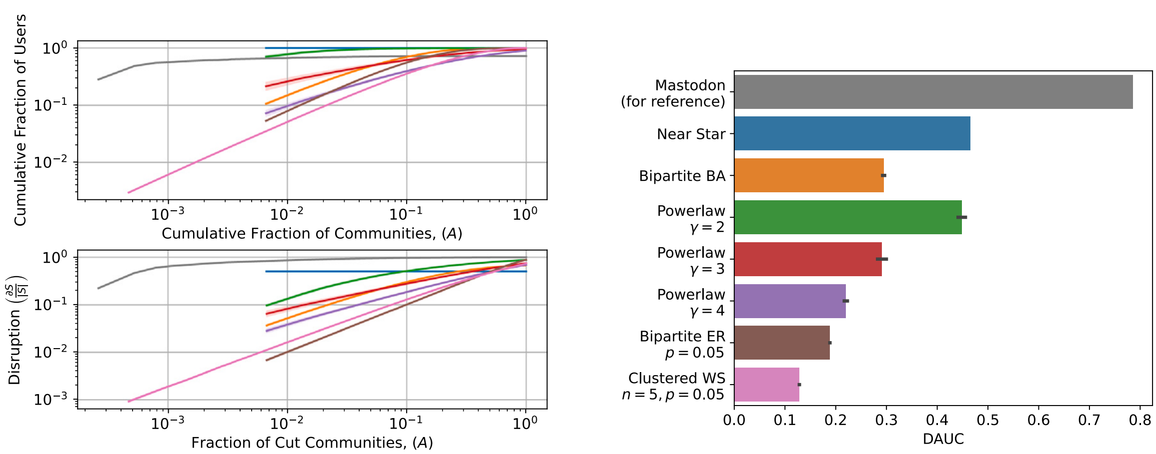
To explain these discrepancies, we examine each network in greater detail. Voat was a Reddit-like platform where users commented and posted in one or more "subverses". While users chose to subscribe from among 7515 public subverses, new accounts were automatically subscribed to a set of 27 subverses by default. This "default subscription" has no parallel on other platforms we examined. Since these default subverses have an automatic population, they are more likely to receive engagement than subverses that must be discovered according to a user's area of interest, and we may expect them to be densely connected with most users on the platform. However, the largest two subverses on Voat by number of unique users were *not* default subverses; `v/QRV` was a QAnon conspiracy group, and `v/8chan` was a right-wing news and discussion forum whose name references the white supremacist image-board 8chan (now "8kun"). Both subverses were highly insular, with little population overlap with the rest of the platform, as illustrated in Fig. 6. Therefore, it is only when we remove the *third-largest* subverse, `v/news`, that we see a large impact on remaining users on the site.



(a) Community size distribution and disruption curves

(b) Area under the disruption curve (DAUC)

Fig. 4. Summary measures of centralization. (a) The population distribution of communities (top) does not correlate with our measure of community disruption (bottom). (b) The area under the disruption curve (DAUC) provides a summary statistic of the disruption curve that reinforces how network structure combined with community size provide greater insight into centralization. Note that while panel (b) provides some insight in to how "centralized" a network is, it fails to capture the nuance of panel (a), such as the minimal effect of removing Voat's largest community. Panel (a) consists of cumulative distribution plots of population and disruption, where the top subplot is a CDF of the platform population as smaller communities are included, and the bottom subplot shows how networks are damaged as more of the largest communities are removed. Each line represents a different network, using the color key from panel b.



(a) Community size distribution and disruption curves

(b) Area under the disruption curve (DAUC)

Fig. 5. In simulated networks with a variety of degree distributions, the disruption curves for each network much more closely match the population distribution (Fig. 5(a)), suggesting that non-degree network attributes such as assortativity play a crucial role in determining centralization. As in Fig. 4, the left figure represents cumulative population and disruption as more communities are considered. Each line represents a network sharing the color-key in the right figure. Simulated networks were generated 100 times, and the mean and a 95% confidence interval are shown in both figures. The number of communities is fixed at 150 during network generation for most simulated networks, except for clustered Watts–Strogatz, where such control is impossible. Therefore, all lines start at the same point on the x-axis except for WS and the reference Mastodon line.

The Penumbra of open-source represents software development on git servers outside of GitHub and the primary GitLab instance. Each community represents a git server with one or more public repositories, and edges indicate that a user (identified by email address) contributed to a repository on a server. Servers are often created per-organization; for example, the Debian Linux distribution hosts their own GitLab server at . Users often contribute to multiple repositories on a single server, but connections *between* servers are extremely sparse. This sparsity is responsible for the "spikes" in the Penumbra's disruption curve; removing a git server may sever an edge to some users, and removing a second, related server may prune all remaining edges to those same users. When the cross-server collaborative users are removed, the impact on the remaining less-collaborative community decreases. In all other networks enough users have a sufficiently high cross-community degree that disruption only increases as communities are removed.

3.2. Comparison to giant component size

Rather than examining the cumulative community size distribution, one could instead examine the size of the giant component of each network. The giant component will shrink as communities are cumulatively removed, providing another means of examining the influence of large communities.

We illustrate this cumulative shrinking in Fig. 7. Most curves are smooth until the tail of the distribution, with two notable exceptions: Voat's giant component changes once the largest insular communities are removed (see Fig. 6), and the Penumbra's curve is much "spikier" as a result of its highly sparse structure.

Measuring the change in giant component size captures some of the same features as our disruption metric. In particular, removing large insular communities may not change the giant component size if the community is completely isolated from the giant component. However, the impact of a community is boolean: if it touches the giant component, then removing the community will shrink the giant

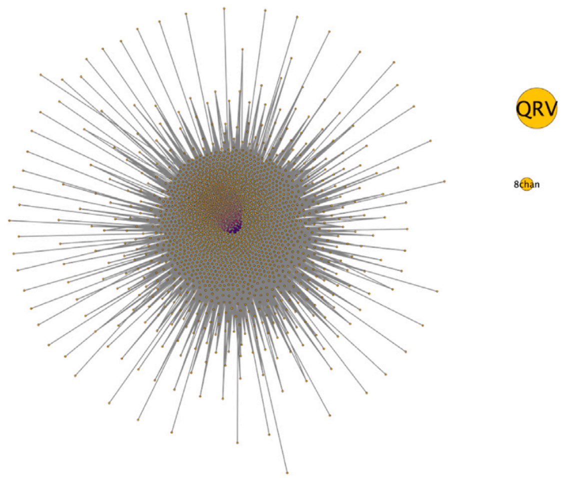


Fig. 6. In this projection of inter-community links on Voat, most of the platform is an amorphous hair-ball, but the overall conclusion is clear: The two largest Voat communities ('QRV' and '8chan') are dramatically larger than their peers, but have almost no overlap in population, making community size a poor proxy for platform-wide influence or centralization. In this network visualization, nodes represent Voat "subverses", and edges represent at least thirty shared users active in two communities. Node size correlates with user count, and color correlates with strength; i.e. the level of overlap with neighboring communities. The purple communities at the center are default subverses all new users are subscribed to ("news", "whatever", etc.), surrounding pink and orange communities are popular with lots of user overlap. The largest two communities, "QRV" and "8chan", have almost no user overlap with other communities and are rendered to the right.

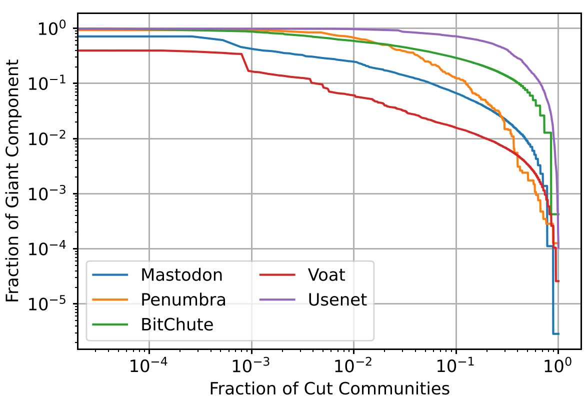


Fig. 7. The giant component shrinks as communities are pruned from largest to smallest, indicating both the size of a community and whether it was part of the giant component before pruning. However, this boolean inclusion does not account for how well-integrated the community was among its peers. The y-axis is normalized as a fraction of the un-pruned giant component size, such that "0.5" indicates the giant component is half the size of the original.

component by the size of that community. There is no distinction between a minimally integrated and tightly integrated community. Measuring the impact of a community in terms of fraction of edges severed, rather than component vertex size, offers finer insight into the interplay between size distribution and network structure.

3.3. Comparison to network bottlenecking

The Cheeger number [23] is a single-valued metric representing how large of a "bottleneck" inhibits conductance across a graph. It is a minimization problem that seeks to divide vertices into two large clusters with a small number of links between them, which is similar to maximizing two-partition modularity. It is typically written as:

$$\min \left\{ \frac{\overset{\text{Edges crossing the boundary of A}}{|\partial A|}}{\underset{\text{All edges in+across A}}{|A|}} : \overset{\text{A is a subset of vertices of G}}{A \subseteq V(G)},\ \underset{\text{A contains at most half of all vertices}}{0 < |A| \leq \frac{1}{2}|V(G)|} \right\} \tag{3}$$

Graph conductance is a global search that measures how similar a graph is to a "barbell", where a small score indicates large communities with few edges across the bottleneck. We are also interested in the size of the bottleneck created between large communities and the rest of a platform, where a large bottleneck and low number of edges among "surviving" users indicates high disruption. However, while the Cheeger number is a community-search algorithm, the communities in our bipartite social-network setting are predefined, and we are interested specifically in the size of the bottleneck for surviving users when the largest communities are removed. Our partition selection is bipartite-aware, such that A includes all the largest communities we are pruning, and all users that only have edges to those communities. Additionally, while the Cheeger number returns a single value for the most "barbell-like" partitioning the graph can achieve, we are interested in the cumulative effect of pruning more and more communities as a means of identifying oligarchic patterns in a network.

Unfortunately, evaluating the graph conductance of all possible subsets of vertices is an NP-hard problem [34] such that it is impractical to directly measure the Cheeger constant on most large graphs. The Cheeger inequality offers upper and lower bounds on the Cheeger number based on the second eigenvalue of the normalized Laplacian of the adjacency matrix, but in our tests these bounds were too wide to offer insightful comparison.

3.4. Assortativity and centralization

High degree disparity is not enough to create a network as centralized as Mastodon. When we control for degree distribution using a variety of "centralized" models including star networks and power-law distributions we cannot achieve more than 50% disruption (Fig. 5(b)). To achieve higher disruption you must have duplicate edges, representing for example a Mastodon user following many accounts on the same server. Therefore, we expect that degree assortativity (or degree-degree correlations) plays a significant role in the differences between observed community disruption (Fig. 4(a)) and network behavior under controlled degree distributions (Fig. 5(a)). In a purely random setting, users are likely to have edges to multiple large communities, because most edge stubs in a configuration model come from high-degree communities. In real social settings, the content of communities may inhibit assortativity, as in Voat, where the largest two communities

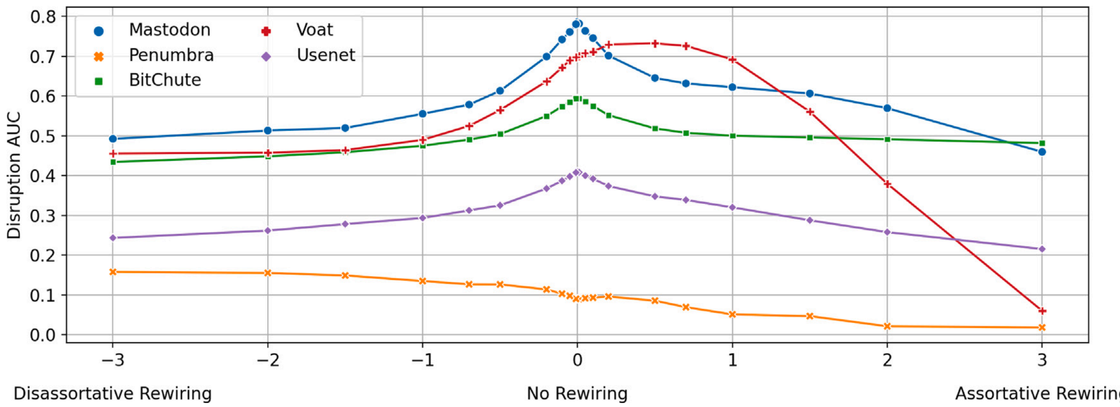


Fig. 8. Rewiring edges to introduce randomness and tune degree-degree correlations. The x-axis corresponds to the expected number of times an edge has been swapped to increase assortativity (positive x-axis) or decrease assortativity (negative x-axis) In most social graphs, any such perturbation leads to a decrease in centralization, regardless of the sign of the change in correlations. In other words, introducing randomness has more impact than changing user-degree to community-degree correlations. There are two exceptions to this rule: in Voat, the two largest communities are highly insular, and so increasing degree assortativity connects these communities and initially *increases* disruption. Conversely, the Penumbra is extremely sparse, and increasing degree assortativity cuts the few inter-community edges, lowering disruption further. Decreasing assortativity makes well-connected users less likely to connect to the largest communities, reducing their role and minimizing the "rich-get-richer" dynamics common to social media. However, in the sparse edge-case of the Penumbra, such rewiring better connects the more populous servers with their smaller peers, increasing disruption.

are highly insular (see Fig. 6), creating a large disparity between the community size distribution and disruption metric.

To explore this hypothesis, we randomly rewired each social network and used this as an opportunity to tune their assortativity. We select pairs of edges uniformly without replacement, and can swap the communities of the edges if doing so would increase or decrease user-community degree assortativity as desired. We continue this process until we have rewired a target percentage of edges; if we exhaust the edge supply before finding sufficient valid swaps, we re-shuffle the edge list and continue drawing. For each rewired network we calculate its disruption and the area under the disruption curve, as in Fig. 4(b), and plot the change in AUC during rewiring in Fig. 8.

The results of Fig. 8 highlights two important results. First, most networks are at their most centralized without any rewiring whatsoever (Mastodon, BitChute, Usenet) or when very little assortative rewiring is applied (Voat). The Penumbra dataset is the only network whose centrality can actually increase under random rewiring. This leads us to our second point which is that the strength of the decentralizing effect of randomness depends on the sign of the correlations introduced. This effect is subtle however, with negative correlations centralizing the Penumbra dataset and positive correlations centralizing the Voat dataset. Unlike the previous mathematical experiment from Eq. 2 which assumes a random infinite network, the finite size of real networks means that correlations can help a user focus their activity on a single community. This mechanism can help us understand the complex role of correlations and some interesting features of individual DAUC curves.

Further exploring correlations in our experiment is useful in distinguishing the idea of network centralization from classic ideas of monopoly. These are two different, but related, problems that are easy to confuse when focusing solely on summary statistics like community size distributions. When a network consists of disconnected communities, it is decentralized under the disruption metric regardless of the size distribution of these communities. This conclusion follows from our definition of centralization since removing a community in a sparse (or disconnected) network, has little (or no) impact on other communities. This rewiring experiment highlights this logic: As networks get rewired to increase correlations, we increase the likelihood of having all the activity of a user focused on a single community and therefore progressively disconnect the community and decentralize the network. The only exception is Voat, whose initial state contains large disconnected communities that can get coupled to the rest of the network by rewiring, before being re-disconnected as we rewire more and more. Small correlations in large networks can therefore increase centralization, since large communities can broker more bridges when they contain well-connected users; while strong correlations in smaller networks can decrease centralization by focusing user activity on single communities.

There are multiple interpretations of degree assortativity in a bipartite setting. The linear correlation between user degrees and community degrees measures whether high-degree users are likely to be connected to high-degree communities. In our network definitions edges represent activity, like follow relationships or participation in conversations, so this measures whether active users are likely to be connected to communities with lots of activity. A second metric of interest is whether large communities are likely to be connected to other large communities, or the assortativity of a unipartite-projected community-community graph. This can be broken into two sub-cases: assortativity of community size (do communities with many users share users with other high-population communities), and assortativity of degree (do communities with lots of activity share users with other high-activity communities). These three notions of assortativity may correlate if high community population correlates with high activity, but this is not guaranteed, so the three metrics should be measured separately.

While rewiring to promote user-community degree assortativity we also plotted the changes in community-community degree assortativity, shown in Fig. 9. Strikingly, the community assortativity *decreases* as we rewire to promote user assortativity. This is because as we rewire edges to focus user connections on the largest communities we implicitly decrease the number of edges between communities. This also matches the changes in disruption in Fig. 8: increasing assortativity may reconnect large and insular communities with the rest of the network, briefly increasing their influence, but continued assortativity rewiring also cuts bridges to and between smaller communities, yielding a sparse network that is far less centralized.

4. Conclusion and future work

We have added to the wealth of centralization metrics by proposing a mesoscale measurement that indicates how much influence one subcommunity has over a broader network, by accounting for how many edges to remaining users would be severed if a community were removed. This metric allows us to differentiate between networks with a substantial community size-imbalance, and networks where the largest communities play a core structural role in their smaller peers. We extend our metric to create a graph-level measurement that indicates how "oligopic" a network is, or how well-integrated its largest communities are with the population at large.

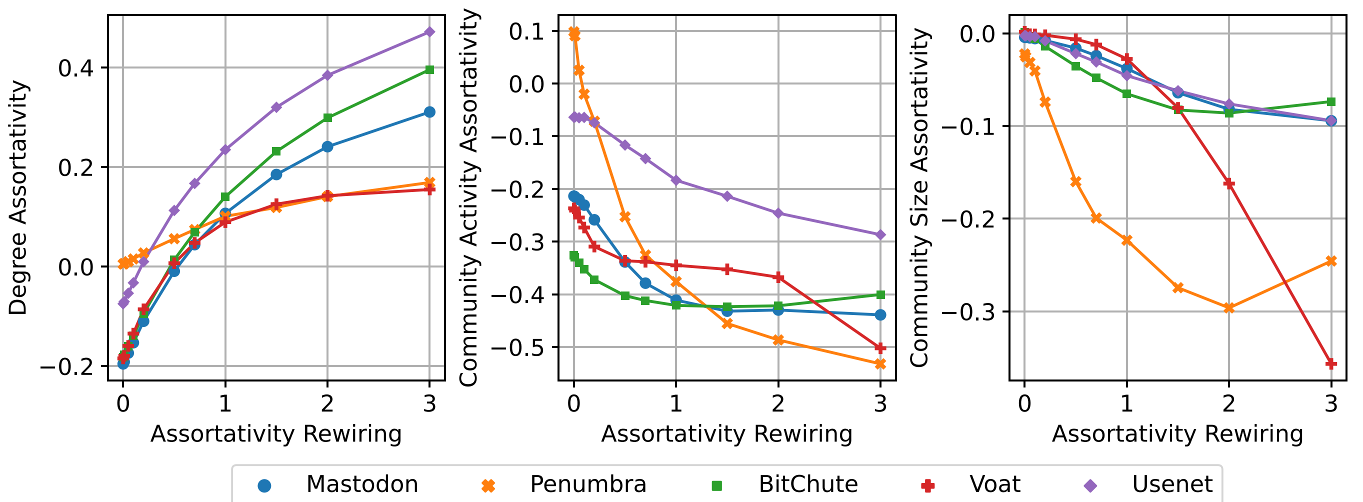


Fig. 9. Rewiring to increase user-community degree assortativity (left) decreases the projected community-community degree assortativity (middle) and community-community population assortativity (right).

We assert that a more nuanced measurement of community influence, accounting for both size distribution and structural role, has utility for content moderation and administrative transparency. Identifying communities on a platform with disproportionate influence can help moderators and administrators limit that influence through changes in content recommendation and integration. Conversely, identifying large communities with lower influence than expected can aid in detecting a large influx of external users, as in the QAnon communities on Voat. Furthermore, third party analysis of community influence can reveal when platform administrators overstate claims of decentralization and community self-governance, understating their own control over and responsibility for a platform.

We have utilized our disruption metric to examine a range of real-world social networks, comparing their network topology, distribution of community sizes, and the influence of those communities. We find that some platforms, like Voat, are much less centralized than their skewed community-size distribution would suggest, while others, like Usenet and the Penumbra of Open-Source, have similar size distributions and widely divergent disruption curves. Mastodon, while vocally supportive of decentralization, has a disruption curve mostly characterized by the skewed population distribution of its sub-communities and is in fact more centralized than any other real or synthetic network considered in this study.

Using simulated networks with a range of degree distributions, and rewiring techniques to adjust assortativity, we have begun to explore the interplay between community size, structure, and community-level centralization. However, we limited ourselves to traditional network generative models like Erdős-Rényi and power-law configuration model networks. Future research could directly simulate networks with chimeric centralization which combine decentralized and centralized components to more realistically represent the diversity observed in social networks.

Our network representations are oversimplified in that we assume that each edge on a network represents a path of information flow. However, one user following another represents *potential* information flow; a bridge between two communities is only realized if the following user is online and chooses to propagate information from the edge to their own followers and instance.

More thorough research should examine how many potential bridges are utilized by, for example, monitoring the number of “boosts” (Mastodon’s equivalent to “retweets”) across instance boundaries on Mastodon. Observed information spread, and examining the reception of cross-pollinated ideas in non-originating communities, would provide much greater insight into how multi-community platforms function in practice.

CRediT authorship contribution statement

Milo Z. Trujillo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laurent Hébert-Dufresne:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **James Bagrow:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Milo Trujillo reports financial support was provided by Google Inc. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

All authors were supported by Google Open Source under the Open-Source Complex Ecosystems And Networks (OCEAN) project. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Google Open Source.

Data availability

Data will be made available on request. Real-world networks used in this study are available from [13,31–33]. Code used in this research, including for synthetic network generation, will be submitted to a public archive before publication or available upon request.

References

[1] D. Gingerich, B.M. Kuhn, Give up GitHub: The time has come!, 2022, URL https://sfconservancy.org/blog/2022/jun/30/give-up-github-launch/.

[2] K. Kupferschmidt, As musk reshapes Twitter, academics ponder taking flight, Sci. (New York, NY) 378 (6620) (2022) 583–584.

[3] D. Hickey, M. Schmitz, D. Fessler, P.E. Smaldino, G. Muric, K. Burghardt, Auditing elon musk's impact on hate speech and bots, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 17, 2023, pp. 1133–1137.

[4] J.A. Jiang, P. Nie, J.R. Brubaker, C. Fiesler, A trade-off-centered framework of content moderation, ACM Trans. Comput.-Hum. Interact. 30 (1) (2023) http://dx.doi.org/10.1145/3534929.

[5] B.I. Davidson, D. Wischerath, D. Racek, D.A. Parry, E. Godwin, J. Hinds, D. van der Linden, J.F. Roscoe, L. Ayravainen, Social media APIs: A quiet threat to the advancement of science, 2023.

[6] H.B. Zia, J. He, A. Raman, I. Castro, N. Sastry, G. Tyson, Flocking to mastodon: Tracking the great twitter migration, 2023, arXiv preprint arXiv:2302.14294.

[7] J. Donovan, B. Lewis, B. Friedberg, Parallel ports: Sociotechnical change from the alt-right to alt-tech, in: Post-Digital Cultures of the Far Right: Online Actions and Offline Consequences in Europe and the US, transcript, 2019, pp. 49–65, http://dx.doi.org/10.25969/MEDIAREP/12374, URL https://mediarep.org/handle/doc/13283.

[8] L.C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (3) (1978) 215–239, http://dx.doi.org/10.1016/0378-8733(78)90021-7, URL https://linkinghub.elsevier.com/retrieve/pii/0378873378900217.

[9] G. Di Bona, A. Bracci, N. Perra, V. Latora, A. Baronchelli, The decentralized evolution of decentralization across fields: from governance to blockchain, 2022, arXiv:2207.14260[physics]. URL http://arxiv.org/abs/2207.14260. Number: arXiv:2207.14260.

[10] A. Beers, S. Nguyen, M. Sioson, M. Mayanja, M. Ionescu, E.S. SPira, K. Starbird, The firestarting troll, and designing for abusability, 2021.

[11] The Bluesky Team, Moderation in a public commons, 2023, URL https://blueskyweb.xyz/blog/6-23-2023-moderation-proposals.

[12] A. Raman, S. Joglekar, E.D. Cristofaro, N. Sastry, G. Tyson, Challenges in the decentralised web: The mastodon case, in: Proceedings of the Internet Measurement Conference, ACM, 2019, pp. 217–229, http://dx.doi.org/10.1145/3355369.3355572.

[13] M. Zignani, S. Gaito, G.P. Rossi, Follow the "mastodon": Structure and evolution of a decentralized online social network, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 12, 2018, pp. 541–550.

[14] M. Zignani, C. Quadri, S. Gaito, H. Cherifi, G.P. Rossi, The footprints of a "Mastodon": How a decentralized architecture influences online social relationships, in: IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2019, pp. 472–477, http://dx.doi.org/10.1109/INFCOMW.2019.8845221.

[15] L. La Cava, S. Greco, A. Tagarelli, Understanding the growth of the fediverse through the lens of mastodon, 2021, arXiv:2106.15473. URL http://arxiv.org/abs/2106.15473.

[16] E. Ostrom, Background on the institutional analysis and development framework: Ostrom: Institutional analysis and development framework, Policy Stud. J. 39 (1) (2011) 7–27, http://dx.doi.org/10.1111/j.1541-0072.2010.00394.x, URL https://onlinelibrary.wiley.com/doi/10.1111/j.1541-0072.2010.00394.x.

[17] R.E. Kraut, P. Resnick, Building Successful Online Communities: Evidence-Based Social Design, Mit Press, 2012.

[18] S. Frey, P.M. Krafft, B.C. Keegan, "This place does what it was built for": Designing digital institutions for participatory change, Proc. ACM Hum.-Comput. Interact. 3 (2019) 1–31, http://dx.doi.org/10.1145/3359134, URL https://dl.acm.org/doi/10.1145/3359134.

[19] N. Schneider, P. De Filippi, S. Frey, J.Z. Tan, A.X. Zhang, Modular politics: Toward a governance layer for online communities, Proc. ACM Hum.-Comput. Interact. 5 (2021) 1–26, http://dx.doi.org/10.1145/3449090, URL https://dl.acm.org/doi/10.1145/3449090.

[20] J. Fan, A.X. Zhang, Digital juries: A civics-oriented approach to platform governance, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ACM, 2020, pp. 1–14, http://dx.doi.org/10.1145/3313831.3376293, URL https://dl.acm.org/doi/10.1145/3313831.3376293.

[21] M. Nadini, L. Alessandretti, F. Di Giacinto, M. Martino, L.M. Aiello, A. Baronchelli, Mapping the NFT revolution: market trends, trade networks, and visual features, Sci. Rep. 11 (1) (2021) 20902, http://dx.doi.org/10.1038/s41598-021-00053-8.

[22] P. Bonacich, Power and centrality: A family of measures, Am. J. Sociol. 92 (5) (1987) 1170–1182.

[23] B. Mohar, Isoperimetric numbers of graphs, J. Combin. Theory Ser. B 47 (3) (1989) 274–291, http://dx.doi.org/10.1016/0095-8956(89)90029-4, URL https://www.sciencedirect.com/science/article/pii/0095895689900294.

[24] C.T. Butts, Revisiting the foundations of network analysis, Science 325 (5939) (2009) 414–416.

[25] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, Nature 406 (6794) (2000) 378–382.

[26] L. La Cava, A. Tagarelli, Information consumption and boundary spanning in decentralized online social networks: the case of mastodon users, 2022, arXiv:2203.15752. URL http://arxiv.org/abs/2203.15752.

[27] M. Fukuda, K. Shudo, H. Sayama, Detecting and forecasting local collective sentiment using emojis, 2022.

[28] N. Kin Wai, S. Horawalavithana, A. Iamnitchi, Multi-platform information operations: Twitter, facebook and youtube against the white helmets, 2021.

[29] Z. Ao, G. Horvath, L. Zhang, Are decentralized finance really decentralized? A social network analysis of the aave protocol on the ethereum blockchain, 2022, arXiv preprint arXiv:2206.08401.

[30] M.Z. Trujillo, L. Hébert-Dufresne, J. Bagrow, The penumbra of open source: projects outside of centralized platforms are longer maintained, more academic and more collaborative, EPJ Data Sci. 11 (1) (2022) 31.

[31] M.Z. Trujillo, M. Gruppi, C. Buntain, B.D. Horne, The MeLa BitChute dataset, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 16, 2022, pp. 1342–1351.

[32] A. Mekacher, A. Papasavva, "I can't keep it up." a dataset from the defunct voat. co news aggregator, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 16, 2022, pp. 1302–1311.

[33] B. Taudul, Archiwum polskiego usenetu, 2021, https://usenet.nereid.pl/. (Accessed 27 May 2022).

[34] V. Kaibel, On the expansion of graphs of 0/1-polytopes, in: The Sharpest Cut: The Impact of Manfred Padberg and His Work, SIAM, 2004, pp. 199–216.