Subreddit Temporal Network: An Interactive Web Interface

Gabriel Mota
Universidade Federal de Viçosa
Viçosa, Brasil
gabriel.f.mota@ufv.br

Sabrina A. Silveira
Universidade Federal de Viçosa
Viçosa, Brasil
sabrina@ufv.br

Julio C. S. Reis
Universidade Federal de Viçosa
Viçosa, Brasil
jreis@ufv.br

Abstract

Reddit is a social media platform driven by community interactions rather than direct relationships between users; interaction happens through subforums called subreddits. This structure can be modeled as a temporal network of subreddits linked by shared users, evolving as users' interests and participation shift over time. However, temporal analysis of such data as a network is still limited by existing visualization tools' support for exploring change over time. This work models Reddit as a temporal network of subreddits organized into communities and builds on it an interactive web interface for exploring the network and its communities over time. Such a tool supports researchers and community analysts in identifying when a community grows, fragments, or shifts its structural core.

Keywords

Complex Network, Data Visualization, Reddit, Temporal Networks, Web Tool

1 Introduction

Reddit [32] is a social media platform driven by community interactions. Unlike other popular social media platforms, such as Instagram and X, Reddit users do not establish direct relationships with one another; instead, they interact through subforums called subreddits [25]. According to [42], users participate in an average of 28.85 subreddits over their lifetime. This extensive participation across multiple subreddits has motivated researchers to model relationships between subreddits based on shared user participation in order to better understand user interests and community formation [31, 43**?**]. These relationships are dynamic, evolving as users join, leave, and redistribute their attention across subreddits over time [26, 30, 47].

Because these relations evolve over time, understanding how communities and subreddits change is important. In the context, we can explore complex network concepts [41] and model these relationships as graphs to enable visualization and analysis. However, network visualization systems, such as Gephi [7] and NetworkX [14], are limited in their ability to show the temporal evolution of network structure and domain characteristics.

In this work, we model Reddit user interactions as a subreddit network, specialize it into a community graph, and analyze the criteria of user participation, subreddit relations, and community formation to construct a network that represents how subreddits are related and form communities over time.

In addition to the temporal network, this work provides a Web-based interactive interface (Figure 1)https://subreddit-network-interface.netlify.app/ to help users discover the subreddit network and explore its communities, their members, and the structural role of each subreddit in its community over time.

2 Related Work

Due to the large amount of data available, as well as its relevance and popularity [12], Reddit has been used in many studies, including in the field of complex networks. There are also several works dedicated to the visualization of complex networks, as they are characterized by their scale and structural complexity [16].

2.1 Reddit Explored as a Complex Network

Reddit is a valuable resource for complex network research due to its large-scale interaction data and explicit community structure. Depending on the research objective, previous work has proposed different network models to encode users, posts, or community relations.

For instance, several studies model Reddit as a subreddit network with relationships inferred through different similarity measures. [36] used crossposts to link subreddits, creating relations based on posts made in multiple communities. Other work applied natural language processing (NLP) techniques to model subreddit relationships from content. In this direction, [37] used embeddings to identify relationships between subreddits sharing named entities, while [35] used cosine similarity between subreddit representations. [29] proposed a subreddit complex network representation based on user interactions across multiple subreddits and used community detection algorithms to show that the resulting subreddit network could reveal patterns of user interests. However, the analysis was limited to a single temporal snapshot.

Many Reddit complex network studies use a single temporal snapshot of the platform [22, 24, 29]. Although snapshot-based representations are effective for identifying structural properties at a specific point in time, they cannot capture temporal changes in communities or in the structural roles of individual subreddits [17]. Other studies have focused on specific communities or interaction contexts within Reddit. For example, [11] investigated the temporal evolution of Reddit discussion networks, while [40] characterized political communities on Reddit. Their specialized scope limits the exploration of large-scale structural relationships across Reddit.

In contrast to previous work, this study proposes a subreddit network based on shared-user relationships and examines both inter- and intra-community structures across multiple temporal snapshots. This approach enables the analysis of structural evolution, changes in community organization, and variations in subreddit influence over time.

2.2 Visualization Systems for Complex Networks

Visualizing complex networks remains challenging due to the scale, density, and structural complexity of these systems [16]. Consequently, different visualization approaches have been proposed

Figure 1: Subreddit Network Discovery Interface's Global View page.

depending on the analytical objectives and the network characteristics to be represented.

Gephi [7] is a widely used general-purpose network visualization system that provides interactive exploration, filtering, network analysis metrics, community detection algorithms, and multiple graph layout techniques. Despite its analytical capabilities, Gephi primarily supports the exploration of static network snapshots and offers limited support for analyzing structural evolution over time.

NetworkX [14] and IGraph [9] are both widely used software packages for complex network research. They provide functionality for network construction, analysis of graph properties, and visualization. Although both support the creation of network snapshots for different time periods, they require proficiency in programming languages such as Python or R, which poses a barrier to users without prior programming experience.

Previous studies have proposed different approaches for visualizing temporal networks. [28], for instance, introduced the "flip-book" approach, in which network evolution is represented as a sequence of discrete snapshots that users can navigate chronologically. [3] expands this approach with animated transitions between snapshots. Animated transitions are not suitable in our setting because neither nodes nor community assignments are preserved across snapshots.

Consequently, this study adopts a snapshot-based paradigm and extends it into an interactive Web-based visualization system that supports the exploration of subreddit relationships across multiple temporal snapshots, enabling both inter-community and intra-community analysis of network evolution.

3 Method

This study proposes an interactive visualization system for exploring Reddit communities through the lens of complex networks over time. To support it, the subreddit network was built with a six-step pipeline, and each step is explored in depth in the following subsections.

3.1 Data Source and Preprocessing

To construct the network, this study uses historical Reddit data provided by Pushshift [8]. It comprises publicly available submissions and comments posted between June 2005 and December 2024, partitioned by month. The data was preprocessed to remove deleted users, retain only the attributes required for network construction, and anonymize user identifiers using a salted hash function.

The raw dataset contains hundreds of thousands of subreddits with highly heterogeneous activity levels. In December 2024, the average subreddit received 967.4 interactions, while the median received only 7. This skew means that most subreddits contribute negligibly to the overall interaction structure while substantially increasing the size and complexity of the resulting network. Following common practice in large-scale network analysis [1, 20, 27], this study addresses this through node sampling [23], retaining

only the subset of subreddits that contributes most significantly to the network's structure.

To determine an appropriate sampling threshold, we evaluated the cumulative share of interactions and users captured by subreddits ranked by activity (Figure 2). The top 5% of subreddits by interaction count account for approximately 90% of all interactions and 92% of all users. Therefore, only the top 5% most active subreddits were retained, a threshold that preserves the vast majority of user activity and network structure while substantially reducing computational cost and visual clutter in the resulting network.

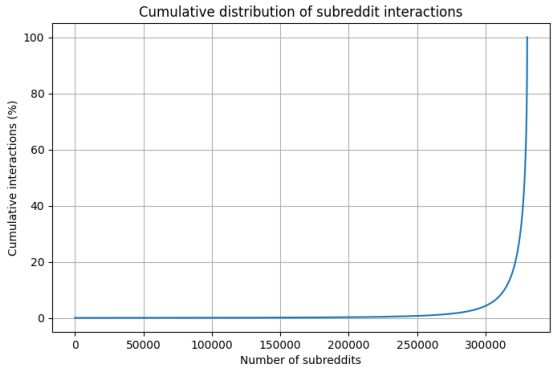

Figure 2: Cumulative distribution of interactions across subreddits in December 2024, with subreddits ordered by interactions in ascending order.

3.2 Network Construction

The subreddit network used in this study is constructed in three steps. First, user participation in subreddits is modeled as a bipartite network in which nodes represent either users or subreddits, and edges are created when a user submits posts to or comments in a subreddit. From this network, meaningful participation is distinguished from one-time interactions. Second, subreddit-to-subreddit relationships are derived from shared user participation, and a backbone extraction procedure is applied to retain only significant connections. Finally, communities are identified, and structural properties are measured within the resulting network.

3.2.1 User–Subreddit Network. To represent user participation in Reddit communities, we constructed a bipartite network connecting users and the sampled subreddits in which they participated. An edge between a user and a subreddit indicates that the user participates in that subreddit based on their interactions. However, defining participation solely by the existence of at least one interaction introduces substantial noise. To better characterize user behavior, we analyzed the distribution of subreddit participation and interaction counts per user within the sampled dataset. As shown in Figure 3, users interacted with an average of 4.2 subreddits; however, only 25% of users participated in four or more subreddits, indicating a highly skewed distribution dominated by a small fraction of highly active users. A similar pattern was observed for interaction counts: users produced an average of 17.1 interactions, while the top 1% of users produced 230 or more interactions in a month.

To reduce the influence of highly active users while preserving meaningful participation patterns, this study combines two filtering strategies. First, a user-centric top-k selection approach was applied, retaining only the three subreddits with the highest interaction counts for each user. According to [15], Reddit users explore new communities, but their participation settles into a small preferred core. Therefore, selecting the top three subreddits captures users' primary community affiliations while limiting the influence of users with highly dispersed activity patterns.

Second, a minimum interaction threshold was applied to distinguish sustained participation from incidental interactions. Prior work suggests that community attachment emerges through repeated participation [33]. Consequently, users were considered members of a subreddit only if they had performed at least two interactions within that community.

3.2.2 Subreddit–Subreddit Network. Given the user-subreddit participation network, subreddits can be related to one another based on shared user participation: two subreddits are connected with a strength proportional to the number of users they have in common. However, a naive relation based on any shared user produces an extremely dense and noisy network, a single shared user is not a statistically meaningful signal of a relationship between two subreddits, yet the large majority of subreddit pairs (73.3%) share only one user (Figure 4). To address this, we first apply a minimum threshold of ten shared users, which substantially reduces network density while retaining subreddits with at least moderate co-participation.

This threshold alone, however, is insufficient: a relation of ten shared users between two large subreddits with hundreds of thousands of users each is not as meaningful as the same shared-user count between two much smaller subreddits. Because the resulting network exhibits a scale-free degree distribution [5], with a small number of high-degree hub subreddits and a long tail of low-degree ones (Figure 5), a single fixed threshold cannot adequately distinguish significant from incidental relationships across such heterogeneous connectivity.

Therefore, we follow [39] and apply the disparity filter, a per-node backbone extraction method that retains edges whose normalized weights are statistically significant relative to the node's total edge weight, under a significance threshold α. This node-adaptive criterion is well suited to scale-free networks, as it applies stricter retention criteria to hub nodes while preserving the comparatively weaker but structurally meaningful connections of peripheral nodes.

To determine the optimal threshold α, we conducted a sensitivity analysis to evaluate the trade-off between retaining network connectivity, measured by the size of the largest component, and enhancing the clarity of the underlying community structure, measured by modularity. As shown in Figure 6, decreasing α yields higher modularity by effectively filtering non-significant "noise" edges; however, extreme values result in severe network fragmentation, where the size of the largest component falls below 60%. We selected $\alpha = 0.05$ as our threshold, maximizing community

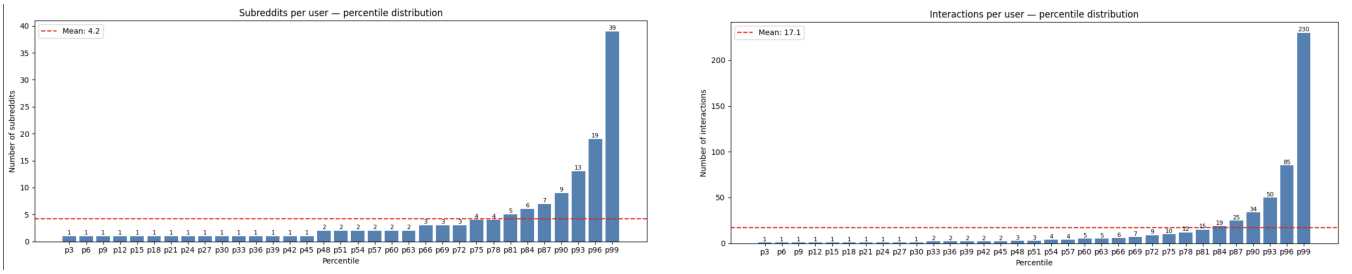

Figure 3: Subreddits and interaction count per user distribution in December 2024.

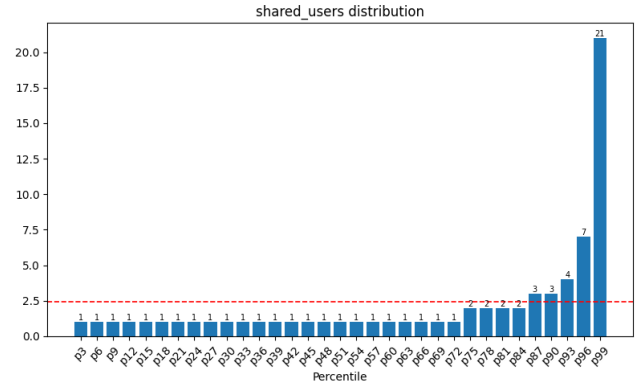

Figure 4: Shared user participation distributions across subreddits in December 2024.

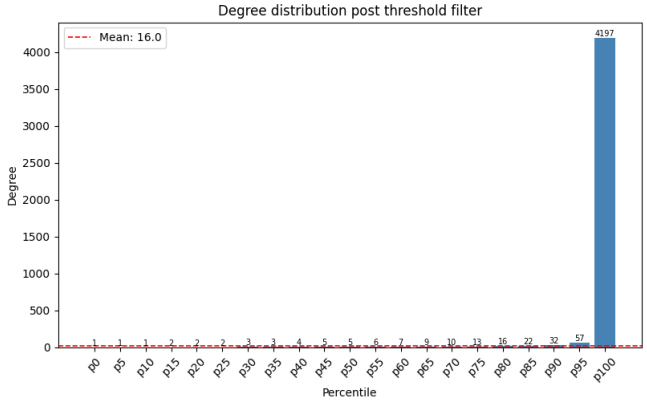

Figure 5: Subreddit network degree distribution after applying the threshold filter.

definition while preserving the global structural integrity of the network.

In this process, some subreddits became irrelevant because all their interactions were dropped. Starting from 16k nodes and 4.4M edges, applying the minimum shared-user threshold reduces the network to 14k nodes and 115k edges; the disparity filter further reduces it to 5k nodes and 7k edges. The resulting backbone retains the statistically significant relationships between subreddits while

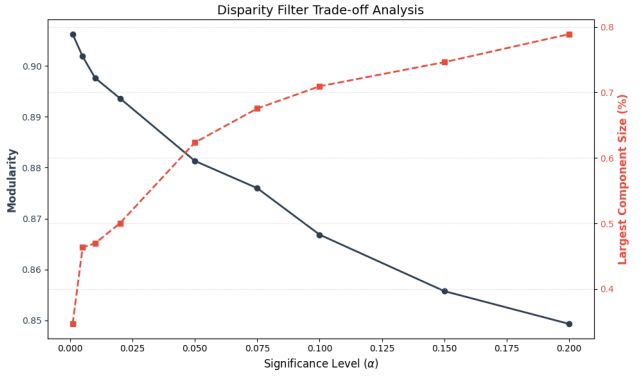

Figure 6: Trade-off analysis of disparity filter thresholds in relation to modularity and largest component size in the December 2024 network.

substantially reducing network density, yielding a graph suitable for community detection.

3.2.3 Community Detection. Before community detection, the backbone network is restricted to its largest connected component. This discards the many small, disconnected fragments and isolated nodes left over from the disparity filter, and focuses the analysis on the network's densely interacted core.

Communities within the subreddit backbone network are identified using the Leiden algorithm [44]. Community detection is performed independently on each monthly network, using a resolution parameter of 1.0 and a fixed random seed to ensure reproducibility. The resulting community assignments form the basis for both the community-level structural metrics and the community-based navigation supported by the visualization system.

3.3 Temporal Analysis

Once monthly subreddit networks are generated, structural properties are computed for every subreddit in each snapshot. Degree and strength quantify the extent and intensity of a node's direct connections [6]. Betweenness centrality [10] indicates its potential role in the flow of information across the network. The clustering coefficient [46] captures local cohesion. Finally, the k-core number [38] provides a measure of a node's structural persistence.

Each metric is computed both on the complete network and on the community subgraph, following the distinction between within-community and between-community node roles proposed by [13]. Global metrics describe a subreddit's structural position within the broader Reddit ecosystem, while community-level metrics describe its relative importance among subreddits in the same community.

Communities are also analyzed over time based on their subreddit members. Using IDF-weighted Jaccard similarity [19] of community members over time, we identify trajectories: chains of communities that, despite membership changes, retain a persistent core, allowing each community's evolution to be tracked individually over time.

In the trajectories, communities are matched to their best counterpart in the next snapshot when similarity exceeds a threshold; for this work, we used 0.3. If a community snapshot does not reach the threshold with any community in the next snapshot, we create its standalone trajectory. Over 5 snapshots, we identified 516 trajectories; of these, only 191 (37%) span more than one snapshot, so most communities are standalone trajectories.

3.4 Interactive Visualization System

To support exploration of the network, the system follows an overview-plus-detail design in which users navigate from a coarse-grained, community-level representation to a fine-grained, subreddit-level one.

3.4.1 Global View. The global view represents each detected community as a single node, sized by its number of member subreddits and colored consistently across views. Edges between community nodes indicate inter-community relationships in the underlying subreddit network. A date slider lets users move between monthly snapshots, reloading the community-level graph for the selected month. This view is intended to support comparison of high-level community structure over time. Selecting a community node opens a detailed view of the community, showing its largest members, and allows the user to transition to the corresponding community view.

3.4.2 Community View. The community view displays the subgraph induced by the subreddits in the selected community, with node size proportional to interaction count. Selecting a subreddit reveals a panel summarizing platform-level attributes, interaction and user counts, and structural metrics, both globally and within the community. This panel plots each metric across every monthly snapshot in which the subreddit appears, showing how the subreddit's structural role changes over time. Together, the two views allow users to move between macro-level analysis of how communities relate to one another and micro-level investigation of individual subreddits' structural roles.

4 Results

4.1 Network Structure

Table 1 summarizes the structural properties of the subreddit backbone network across five annual snapshots from December 2020 to December 2024. All networks exhibit low density (<0.0025), which decreases as network size increases, consistent with observations in large-scale real-world networks [2].

Modularity increased from 0.590 to 0.800 between 2020 and 2024, while the relative importance of high-degree hubs decreased. Together, these observations suggest that subreddit interactions became increasingly organized into more cohesive communities.

4.2 Interface Demonstration

To illustrate how the system supports temporal exploration of community evolution, we present a case study of the Brazilian hub on Reddit, tracked across five snapshots from December 2020 to December 2024.

In December 2020, the Brazilian community was small and strongly centered around r/brasil (Figure 7). By 2024, the community had grown (Figure 8) and become more decentralized, with the influence of r/brasil decreasing (Figure 9) and other subreddits emerging as influential nodes, such as r/MeJulgue and r/opiniaoimpopular (Figure 10).

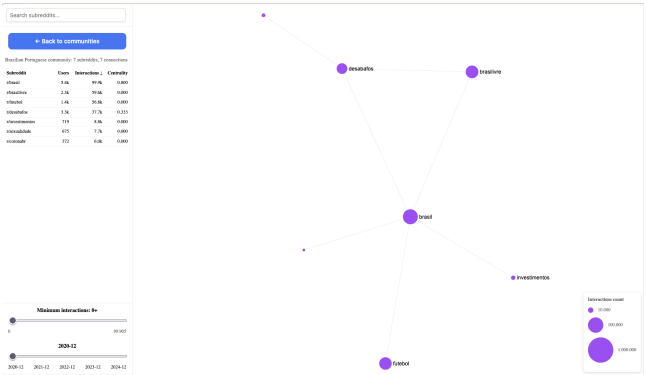

Figure 7: Brazilian community in December 2020

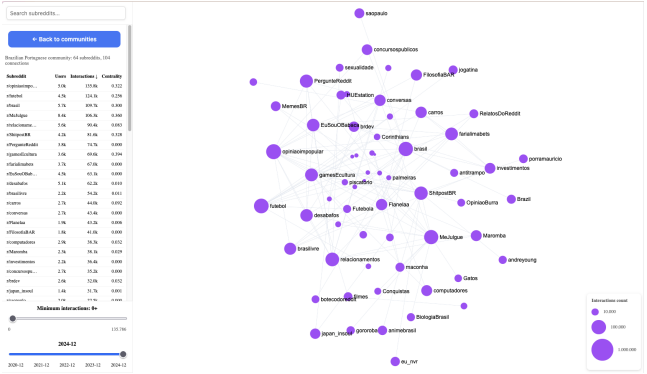

Figure 8: Brazilian community in December 2024.

This case study illustrates how the interface supports identifying the subreddits responsible for community growth and changes in structural influence. The interface is available at https://subreddit-network-interface.netlify.app/.

Table 1: Monthly Snapshot Networks Stats

Date	Number of nodes	Number of edges	Density	Number of communities	Modularity	Max degree	Average degree
2020-12	2224	4511	0.00182	162	0.5899	830	4.0566
2021-12	2573	5282	0.001596	195	0.5964	1027	4.1057
2022-12	2761	5266	0.001382	214	0.6470	932	3.8146
2023-12	2939	5096	0.001180	180	0.7461	713	3.4678
2024-12	3333	5797	0.001044	140	0.7997	464	3.4785

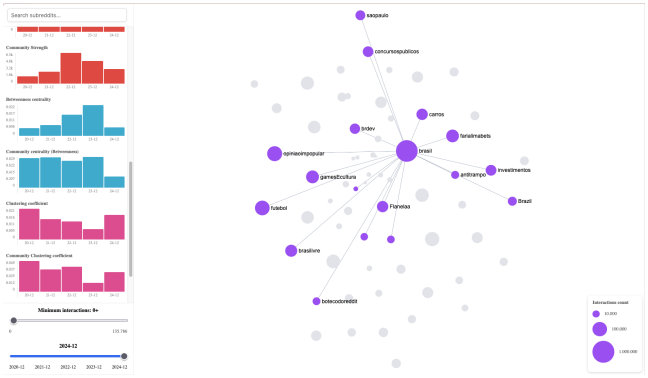

Figure 9: r/brasil in community view highlight, December 2024 snapshot.

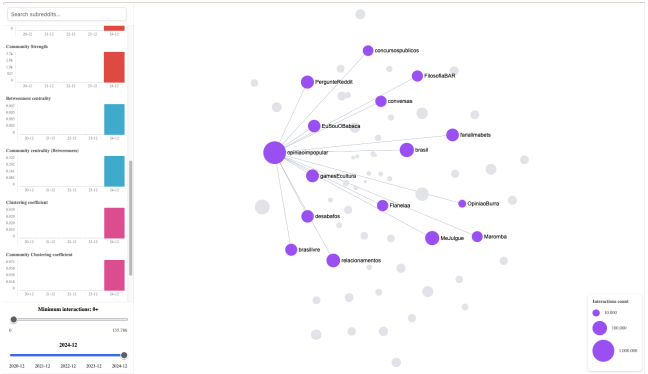

Figure 10: r/opiniaoimpopular in community view highlight, December 2024 snapshot.

5 Discussion

The observed increase in network modularity, together with the reduction in maximum and average degree during the later snapshots, suggests that subreddit interactions are becoming less concentrated around a small number of highly connected hub subreddits and more fragmented. This interpretation is consistent with previous work describing Reddit user behavior, in which users curate their participation across smaller, topic-specific groups rather than relying exclusively on general-purpose subreddits [18].

Thus, this work contributes to the study of user behavior over time with respect to topic and community relations [4] by providing an interactive interface for exploring the temporal evolution of subreddit networks across multiple domains. This approach enables the analysis of how topic and community influence evolve over time in domains such as politics, pop culture, sports, and others.

5.1 Limitations and Future Work

The temporal coverage of the dataset is constrained by computational and data availability limitations. Processing each monthly interaction network requires substantial storage and computational resources because the Reddit dataset exceeds 3 TB. In addition, recent changes to Reddit's API access policy have limited the availability of newer data, further restricting the temporal range that could be incorporated into this study.

There are several opportunities to extend the proposed system. Expanding the temporal dataset with additional network snapshots would enable investigation of longer-term structural trends. Following [21], the visualization can be used to identify trends in community growth and decay, and to predict community behavior. Another opportunity is to use Reddit community trends to understand user behavior and model shifts in interest around real-world events, following the research of [34, 45].

The visualization interface itself can also be enhanced. Future versions could incorporate additional interaction techniques, including filtering, richer edge representations, and community-level analytical summaries.

Finally, although the proposed system was developed for Reddit interaction networks, its visualization paradigm is not inherently tied to this domain. Future work could generalize the preprocessing pipeline and expose a configurable data interface, allowing networks generated from different domains, community detection methods, or node and edge attributes to be visualized within the same analytical framework.

Artifact Availability

The web interface presented in this work is available at https://subreddit-network-interface.netlify.app/. The full source code, including the web application, the network construction pipeline, and the exploratory notebooks that support this study, is available at https://github.com/gfmota/subreddit-community-analysis.

Acknowledgements

Claude, powered by the Claude Sonnet 5 model, was used to improve the readability of this article, to assist in software development, and to generate community labels. The authors take full responsibility for the content of this work. Finally, the authors are grateful to agencies CAPES, FAPEMIG, and CNPq for funding different parts of this work.

References

[1] Yong-Yeol Ahn, Seungyeop Han, Haewoon Kwak, Sue Moon, and Hawoong Jeong. 2007. Analysis of topological characteristics of huge online social networking services. In *Proc. of the International Conference on World Wide Web (WWW)*. 835–844.

[2] Réka Albert and Albert-László Barabási. 2002. Statistical mechanics of complex networks. *Reviews of modern physics* 74, 1 (2002), 47.

[3] Benjamin Bach, Emmanuel Pietriga, and Jean-Daniel Fekete. 2013. Graphdiaries: Animated transitions andtemporal navigation for dynamic networks. *IEEE transactions on visualization and computer graphics* 20, 5 (2013), 740–754.

[4] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 44–54.

[5] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *Science* 286, 5439 (1999), 509–512.

[6] Alain Barrat, Marc Barthelemy, Romualdo Pastor-Satorras, and Alessandro Vespignani. 2004. The architecture of complex weighted networks. *Proc. of the National Academy of Sciences* 101, 11 (2004), 3747–3752.

[7] Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. 2009. Gephi: an open source software for exploring and manipulating networks. *Icwsm* 8, 2009 (2009), 361–362.

[8] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset. In *Proc. of the International AAAI Conference on Web and Social Media (ICWSM)*, Vol. 14. 830–839.

[9] Gabor Csardi and Tamas Nepusz. 2005. The Igraph Software Package for Complex Network Research. *InterJournal* Complex Systems (2005), 1695.

[10] Linton C Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry* (1977), 35–41.

[11] Diletta Goglia and Davide Vega. 2024. Structure and dynamics of growing networks of Reddit threads. *Applied Network Science* 9, 1 (2024), 48.

[12] Jeffrey Gottfried and Eugenie Park. 2025. Americans' Social Media Use 2025. https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-2025/. Accessed: 2026-07-05.

[13] Roger Guimera and Luís A Nunes Amaral. 2005. Cartography of complex networks: modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment* 2005, 02 (2005), P02001.

[14] Aric A Hagberg, Daniel A Schult, and Pieter J Swart. 2008. Exploring network structure, dynamics, and function using NetworkX. In *Proc. of the Python in Science Conference*. 11–15.

[15] William Hamilton, Justine Zhang, Cristian Danescu-Niculescu-Mizil, Dan Jurafsky, and Jure Leskovec. 2017. Loyalty in online communities. In *Proc. of the International AAAI Conference on Web and Social Media*, Vol. 11. 540–543.

[16] Ivan Herman, Guy Melançon, and M Scott Marshall. 2000. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on visualization and computer graphics* 6, 1 (2000), 24–43.

[17] Petter Holme and Jari Saramäki. 2012. Temporal networks. *Physics reports* 519, 3 (2012), 97–125.

[18] Sohyeon Hwang and Jeremy D Foote. 2021. Why do people participate in small online communities? *Proc. of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.

[19] Paul Jaccard. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat* 37 (1901), 547–579.

[20] Jing Jiang, Christo Wilson, Xiao Wang, Wenpeng Sha, Peng Huang, Yafei Dai, and Ben Y Zhao. 2013. Understanding latent interactions in online social networks. *ACM Transactions on the Web (TWEB)* 7, 4 (2013), 1–39.

[21] Sanjay Ram Kairam, Dan J Wang, and Jure Leskovec. 2012. The life and death of online groups: Predicting group growth and longevity. In *Proc. of the ACM International Conference on Web Search and Data Mining*. 673–682.

[22] Pushwitha Krishnappa, Lance Lindner, Eduardo Pasiliao, and Tathagata Mukherjee. 2024. Identifying NSFW groups on Reddit social network by identifying highly interconnected subreddits through analysis of implicit communication patterns. *Applied Sciences* 14, 24 (2024), 11665.

[23] Jure Leskovec and Christos Faloutsos. 2006. Sampling from large graphs. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 631–636.

[24] Lance Lindner. 2023. *Analysis of Interconnected Subreddits on the Reddit Social Network*. The University of Alabama in Huntsville.

[25] Noah Makow, Matt Millican, and Chenyao Yu. 2017. A Network Model for Reddit Post Virality Prediction. (2017).

[26] Humphrey Mensah, Lu Xiao, and Sucheta Soundarajan. 2020. Characterizing the evolution of communities on reddit. In *International Conference on Social Media and Society*. 58–64.

[27] Alan Mislove, Massimiliano Marcon, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. 2007. Measurement and analysis of online social networks. In *Proc. of the ACM SIGCOMM Conference on Internet Measurement*. 29–42.

[28] James Moody, Daniel McFarland, and Skye Bender-deMoll. 2005. Dynamic network visualization. *American journal of sociology* 110, 4 (2005), 1206–1241.

[29] Randal S Olson and Zachary P Neal. 2015. Navigating the massive world of reddit: Using backbone networks to map user interests in social media. *PeerJ Computer Science* 1 (2015), e4.

[30] Elliot Panek, Connor Hollenbach, Jinjie Yang, and Tyler Rhodes. 2018. The effects of group size and time on the formation of online communities: Evidence from Reddit. *Social Media+ Society* 4, 4 (2018), 2056305118815908.

[31] Gale H Prinster, C Estelle Smith, Chenhao Tan, and Brian C Keegan. 2024. Community archetypes: An empirical framework for guiding research methodologies to reflect user experiences of sense of virtual community on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–33.

[32] Reddit, Inc. 2026. Reddit. https://www.reddit.com/. Accessed: 2026-07-12.

[33] Yuqing Ren, F Maxwell Harper, Sara Drenner, Loren Terveen, Sara Kiesler, John Riedl, and Robert E Kraut. 2012. Building member attachment in online communities: Applying theories of group identity and interpersonal bonds1. *MIS quarterly* 36, 3 (2012), 841–863.

[34] Jon Roozenbeek and Adrià Salvador Palau. 2017. I read it on reddit: Exploring the role of online communities in the 2016 us elections news cycle. In *International Conference on Social Informatics*. 192–220.

[35] Jan Sawicki and Maria Ganzha. 2024. Exploring Reddit community structure: Bridges, gateways and highways. *Electronics* 13, 10 (2024), 1935.

[36] Jan Sawicki, Maria Ganzha, Marcin Paprzycki, and Yutaka Watanobe. 2023. Reddit CrosspostNet—studying Reddit communities with large-scale Crosspost graph networks. *Algorithms* 16, 9 (2023), 424.

[37] Jan Sawicki, Maria Ganzha, Marcin Paprzycki, and Yutaka Watanobe. 2024. Applying named entity recognition and graph networks to extract common interests from thematic subfora on Reddit. *Applied Sciences* 14, 5 (2024), 1696.

[38] Stephen B Seidman. 1983. Network structure and minimum degree. *Social Networks* 5, 3 (1983), 269–287.

[39] M Ángeles Serrano, Marián Boguná, and Alessandro Vespignani. 2009. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences* 106, 16 (2009), 6483–6488.

[40] Ahmed Soliman, Jan Hafer, and Florian Lemmerich. 2019. A characterization of political communities on reddit. In *Proceedings of the ACM Conference on Hypertext and Social Media (HT)*. 259–263.

[41] Steven H Strogatz. 2001. Exploring complex networks. *nature* 410, 6825 (2001), 268–276.

[42] Chenhao Tan and Lillian Lee. 2015. All who wander: On the prevalence and characteristics of multi-community engagement. In *Proc. of the International Conference on World Wide Web (WWW)*. 1056–1066.

[43] Nathan TeBlunthuis, Charles Kiene, Isabella Brown, Laura Levi, Nicole McGinnis, and Benjamin Mako Hill. 2022. No community can do everything: why people participate in similar online communities. *Proc. of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–25.

[44] Vincent A Traag, Ludo Waltman, and Nees Jan Van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports* 9, 1 (2019), 5233.

[45] Isaac Waller and Ashton Anderson. 2021. Quantifying social organization and political polarization in online platforms. *Nature* 600, 7888 (2021), 264–268.

[46] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of 'small-world'networks. *nature* 393, 6684 (1998), 440–442.

[47] Jason Shuo Zhang, Brian Keegan, Qin Lv, and Chenhao Tan. 2021. Understanding the diverging user trajectories in highly-related online communities during the COVID-19 pandemic. In *Proc. of the International AAAI Conference on Web and Social Media*, Vol. 15. 888–899.