

Building Bridges between Users and Content across Multiple Platforms during Natural Disasters

Lynnette Hui Xian Ng
Carnegie Mellon University
Pittsburgh, USA
lynnetteng@cmu.edu

David Farr
University of Washington
Seattle, USA

Iain Cruickshank
Carnegie Mellon University
Pittsburgh, USA

Abstract

Social media is a primary medium for information diffusion during natural disasters. The social media ecosystem has been used to identify destruction, analyze opinions and organize aid. While the overall picture and aggregate trends may be important, a crucial part of the picture is the connections on these social media sites during natural disasters. These connections, or bridges, are essential to facilitate information flow within the network. In this work, we perform a multi-platform analysis (X, Reddit, YouTube) of Hurricanes Helene and Milton, which occurred in close in time to each other in the US in late 2024. We construct network graphs to understand the properties of effective bridging content and users. We find that bridges tend to exist on X, that bridging content is complex, and that bridging users have relatable affiliations related to gender, race and job. Public organizations can use these characteristics to manage their social media personas during natural disasters to more effectively spread helpful information.

CCS Concepts

- Information systems → Social networks; Clustering.

Keywords

social media, network analysis, multi-platform, bot detection

ACM Reference Format:

Lynnette Hui Xian Ng, David Farr, and Iain Cruickshank. 2018. Building Bridges between Users and Content across Multiple Platforms during Natural Disasters. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 6 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Social media has become a critical tool during natural disasters, serving as a platform for authorities and organizations to provide public information, achieve situational awareness, and coordinate grassroots aid [19]. Simultaneously, the distributed and participatory nature of social media enables the collection of valuable, localized information through everyday digital traces and interactions of users [2]. This crowd-sourced data has proven useful

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-XXXX-X/18/06
<https://doi.org/XXXXXXX.XXXXXXX>

for tasks such as damage assessment [2], constructing situational awareness overviews [15], and modeling information diffusion [4].

In addition to aggregate narrative trends and identifying influential users, another critical but often overlooked aspect of social media is the **connections** within the network. These connections, referred to as *bridges*, facilitate the flow of information across different communities and are vital during natural disasters. Bridges ensure that key insights reach diverse social media groups, helping to disseminate life-saving information effectively [6].

Bridges manifest in multiple forms:

- (1) **Intra-platform bridges:** Within a single social media platform, bridges can:
 - Represent *topic-oriented communities* connected through shared narratives or discussions.
 - Be users who hold structural importance in the network, such that their removal disrupts clusters of nodes [14].
 - Include not only human users but also automated bots, which can artificially link highly connected user clusters to influence information flow [10].
- (2) **Cross-platform bridges:** Bridges also operate across different social media platforms. While platforms have distinct ecosystems and rules, narratives about significant events often propagate across them [12]. Examples include:
 - *Linker posts*, where users on one platform (e.g., X) direct audiences to content on another (e.g., YouTube).
 - *Cross-platform bridging users*, who consistently amplify content from specific users across multiple platforms, boosting visibility and coherence in messaging [8].

Bridges, whether within or across platforms, play a critical role in the complex social media ecosystem during natural disasters. Understanding and identifying these bridges is essential to enhancing information dissemination, mitigating misinformation, and improving disaster response effectiveness.

In late 2024, two hurricanes happened in the US within a week of each other, sparking discussions on social media that highlight the use and importance of bridges on these sites. Hurricane Helene was a Category 4 hurricane that affected the Carolinas, Georgia and the surrounding areas. Hurricane Milton was a Category 5 hurricane that mostly affected Florida and Georgia. We collected data from X, YouTube and Reddit related to the discourse on these hurricanes, and analyzed the properties of bridging narratives and users. Using our analysis, we establish general characteristics of social media bridges, and elaborate on how such bridges can be leveraged for disaster management.

	Helene		Milton	
Platform	Posts	Users	Posts	Users
X	5000	3613	5000	4054
YouTube	500	225	500	114
Reddit	247	124	239	272

Table 1: Statistics of Data Collected

Research Questions. In this work, we use a combination of linguistic analysis and network science to distill the features of content and user bridges within the hurricane discourse. As such, we ask one overarching research question: **What makes an effective bridge?** To answer this question, we formulate three sub-research questions:

- What platforms build bridges?
- What types of content build bridges?
- What types of users build bridges?

2 Methodology

2.1 Data

We used the InformationTracer¹ platform to collect data. We used the search terms “hurricane helene” and “hurricane milton”. We retrieved data from three social media platforms: X, YouTube and Reddit. This process retrieved tweets from X, video captions from YouTube, and original posts from Reddit. Table 1 presents the statistics of the data. The data was collected on the week of each of the hurricanes: 24-29 September 2024 for Helene and 5-12 October 2024 for Milton.

2.2 Data Preprocessing

The texts collected are first preprocessed to remove stopwords and social media artifacts like hashtags, @-mention symbols, etc. Retweets are kept because they signify a repetition of a particular message. The texts are then vectorized using a standard BERT-based embedding model². This model is chosen for its ability to vectorize both sentences and paragraphs into the same vector space as we have short texts from X, and varying lengths of texts from YouTube and Reddit.

2.3 Identifying Bridging Nodes

In graph theory, a bridge is a link whose deletion increases the number of connected components in a graph [1, 18]. We adapt this idea to bridging nodes, and define a bridging node as a node whose removal and the removal of its links increases the shortest path between its 1-degree connected neighbors, or increases the number of connected components.

We modeled the spread of narratives through two network graphs: the Content graph and the User’s graph. The Content graph is constructed by comparing the similarity of texts from posts. The nodes represent texts, and links represent similarity of the texts. Text

similarity is calculated in the following manner: The vectors of pre-processed texts are compared with an all-pairs cosine comparison. Two texts are deemed similar if their cosine similarity is greater than 0.8, which is a threshold used in other text network constructions [3, 12]. Bridges in the Content graph show how narratives are interconnected.

The Users graph is constructed from users that have similar texts. This uses the matrix transformation procedure defined in [12] to identify users that are similar through their textual content in a cross-platform setting. In this graph, the nodes are users. Two users have a link if they have similar content. Bridges in the Users graph show which users connect over macro-textual elements like narratives.

After constructing both graphs, we apply the Leiden clustering algorithm on the graphs. We pruned the graph to retain only the top clusters; that is, clusters that have more than 10 nodes. After pruning the graph, the Content network is left with 9 Leiden groups consisting of a total of 10,138 texts. The User network is left with 7 Leiden groups consisting a total of 789 users. From these pruned graphs, we find bridges using the Network Chain Decomposition method [16] from the NetworkX library³. This method decomposes the graph into connected chains of nodes, then calculates betweenness centrality of the nodes. Nodes with high betweenness centrality are potential bridging nodes, because they lie on the shortest paths of multiple nodes. The algorithm then disrupts the chain by removing these suspected bridging nodes and checks if the graph becomes more disconnected (i.e. increase in the number of connected components).

2.4 Analysis of Bridging Nodes

Calculating Multi-platform linkages. We analyzed the proportion of bridging nodes against the number of nodes of each platform. Then, we calculated the multi-platform linkages between the nodes. We analyzed the links for both Content and Network graphs in terms of platform pairs of the nodes that they link between. We plot Content-Platform and Content-User graphs, where nodes are social media platforms, and links are the extent to which the nodes are linked together in the original graph through similarity of Content or User respectively. For the User graph, we excluded self-loops, where the user posted similar text to himself. We also calculated the proportion of cross-platform linkers, which are nodes (Content and User respectively) that bridge between two platforms.

Calculating network centrality metrics. For both the Content and User networks, we calculate four centrality metrics of bridging and non-bridging users. These metrics are: (1) total degree centrality, which measures the number of incident edges to a node. Nodes with many connections have high degree centrality, and can act as connectors between graph components. (2) Eigenvector centrality, which is the measure of how connected the node is to other influential nodes. (3) Hub centrality, which measures the extent the node links to authority or influential nodes. Nodes with high eigenvector or hub centrality can be effective information disseminators because they can spread information to other influential nodes. (4) Betweenness centrality, which is the extent to which a

¹<https://informationtracer.com>

²<https://huggingface.co/sentence-transformers/all-mpnet-base-v2>

³<https://networkx.org>

node lies on the shortest path of other pairs of nodes. Nodes with high betweenness centrality lie on the shortest path of many other sets of nodes.

Calculating linguistic cues. To calculate linguistic cues, we use the NetMapper software⁴. This software identifies psycholinguistic cues from the texts using a lexicon and has been used in many social media studies. We extracted linguistic cues for the texts and the usernames. The salient linguistic cues are picked out and compared between bridging and non-bridging nodes.

Topic analysis of content. For each Leiden cluster in the Content network, we perform Latent Dirichlet Allocation on the vectors of the preprocessed texts. For each cluster, we retrieve the words associated with the top 3 topics. The resultant topics are manually interpreted and annotated in Figure 1(b).

Bot user annotation. To understand whether bridging nodes occur organically or inorganically, we annotate the users as bot or human using BotBuster [9]. BotBuster uses a mixture-of-experts architecture, where each expert is a machine learning model to evaluate each subset of data. This algorithm is ideal for our dataset because the common data pillar for each of the three platforms is the user name. Therefore, we only activate the user name expert when using BotBuster. The algorithm returns $P(\text{bot}) \in [0,1]$. As per the original paper, we used the threshold of 0.5; therefore, $P(\text{bot}) \geq 0.5 = \text{bot}$ and $P(\text{bot}) < 0.5 = \text{human}$.

User identity affiliation annotation. We discern the social identity that users affiliate themselves with by comparing their usernames against a lexicon of identities from a 2015 population survey of US residents that consolidates representations of their occupation and affiliations [17]. The users are annotated with their identities, which are manually analyzed in the context of the User graph, and the manual unification of identity sets are annotated on the network graph itself. These identities are classified into 7 categories (political, family, gender, race/nationality, religion, job, other), and the proportion of the categories are compared between bridging and non-bridging users.

3 Results

Figure 1 presents the results from our analyses. The Content graph has 8,080 links, while the User graph has 35,889 links. Despite the extensive linkages (representing the similarity of nodes to each other) within the networks, the proportion of bridging nodes across all platforms are small (< 1%, Figure 1(a)). Across both Content and Users, X hosts the highest number of bridging nodes, followed by YouTube and Reddit. Reddit is not an ideal platform for bridging nodes as its communities are segregated into distinct subreddits. 17.9% of the links in the User graph and 24.2% of links in the Content are cross-platform links (Figure 1(d)(e)). The most commonly linked Content and User is between X and YouTube.

By definition of a bridge, the betweenness centrality of bridging nodes, whether content or user nodes, are higher than non-bridging nodes, connecting disparate parts of the network (Figure 1(f)(g)). However, the eigenvector centrality metric behaves differently. Eigenvector centrality is higher for bridging users than

non-bridging users, but lower for bridging content than non-bridging content. This means that bridging users cluster around other influential users, while this clustering pattern does not hold for bridging content. These dynamics highlight the nuanced roles of bridging nodes in facilitating information flow.

Specific to each hurricane, the topics are the donations and situational updates (Figure 1(b)). However, bridging content comes in the form of donations, relief efforts and conspiracy theories that are related to both hurricanes. Such examples are: announcements of government funding for both hurricanes, or theories that the hurricanes are man-made because they are in close succession with each other. In particular, bridging content connect groups of content that are talking about one hurricane, and morph the content to talking about the commonalities between both hurricanes. This therefore informs the types of users in the User graph, where the most common user affiliations are news broadcasters, humanitarian aid organizations, region specific news agencies and conspiracy theorists (Figure 1(c)). These users typically want to reach multiple user communities to broadcast their message, be it a situational update, disaster aid information or conspiracy. Therefore, their bridging property serves to enhance their purpose in the graph.

Bridging content is typically complex. This is observed in Figure 1(h)(i), where the mean of the linguistic cues for bridging content is higher than that of non-bridging content. The complexity of bridging content thus allows it to cover a breadth of topics and be similar to multiple sets of other content, as compared to the simpler non-bridging content. Bridging users have usernames that are longer (average sentence length) and have words in capital letters (# all caps), perhaps to attract attention to their content.

From the bot annotation (Figure 1(j)), 67.2% of the bridging users are bots, while 0.5% of the non-bridging users are bots. Bridging users straddle between multiple communities of users, which breaks the principle of homophily, where users congregate together because of common interests [7]. It is easier for bots, who are inorganic users controlled by a computer program, to make connections with multiple groups. Such heterophily is harder for humans to maintain, with studies showing that heterophily can result in poor physical and mental health [5].

In terms of social identity, 34.5% of bridging users and 27.8% of non-bridging users present an identifiable social identity (Figure 1(k)). Most of the identities extracted are job-related, which may be an artifact of the identity survey. Bridging users affiliate more with gender, race, nationality and job identities than non-bridging users. Such identities are common identities that all humans can relate to, as compared to specific interests, which suggests the effectiveness of bridging users to lean towards these identities.

4 What makes an Effective Bridge?

When a bridging node is removed from the graph, the number of connected components in the graph increases. That is, if there were originally 1 connected component in the graph, a bridge removal results in the graph having 2 or more connected components. Since we used that algorithm to identify bridges, the bridges in our study, are by definition, “effective”, and produced the desired result in

⁴<https://netanomics.com/netmapper/>

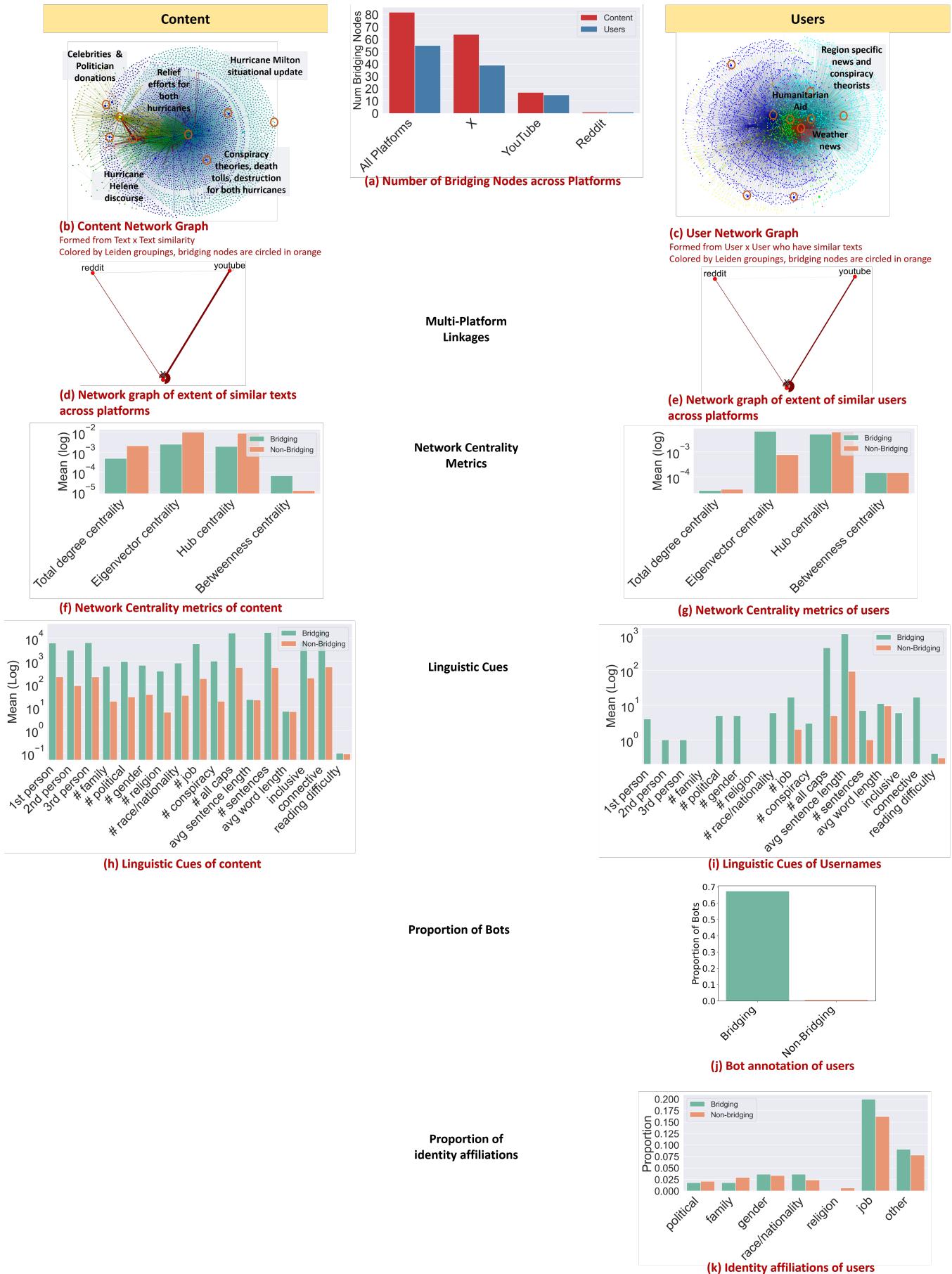


Figure 1: Comparison of bridging and non-bridging nodes

fragmenting the graph. However, fragmenting an extremely interconnected graph such as a social media discourse is a complex operation, therefore, we do not have many bridges.

Our methodology allows establishment of general characteristics of social media bridges:

- Platform: Bridges tend to be nodes on X. X is a platform that connects multiple platforms via summaries or URLs (to the Reddit discussion thread or YouTube video) [8], and are therefore effective bridges. The feed algorithm on X is created to maintain a balance and diverse feed, which allows crossover of content types⁵. In contrast, Reddit gatekeeps users by interest-based subreddits, and YouTube where its recommendation algorithm is based on the user's past watched videos.
- Content: Bridging content are complex, having long sentences, many inclusive and connective words, and many pronouns. Such complex content are effective bridges since they can bridge multiple communities.
- User type: Bots are effective bridging users for they can easily straddle different sets of user communities and take on multiple personas, even changing affiliations. Identity affiliations related to gender, race and job are commonly used among bridging users compared to interest-specific affiliations. Influential users (high eigenvector centrality) are more primed to be bridging users.

Understanding social media bridges can significantly inform disaster management strategies. During natural disasters, governments and non-profit organizations aim to disseminate situational updates, aid-related information, and combat misinformation [19]. To maximize the reach and acceptance of their messages, these organizations' social media personas should resonate with diverse communities. By positioning these personas as bridges, organizations can leverage various strategies, such as adopting characteristics of effective bridges, natural identity affiliations, or influential users (e.g., politicians or local figures). Automation in the form of bots can be used for wider message reach, though malicious bots can also be used to propagate misinformation and conspiracy theories during natural disasters [13]. Messages can also be crafted with complex sentence structures that address multiple topics simultaneously. Furthermore, platforms like X often offer strategic advantages for broad dissemination of disaster-related information.

Dataset Availability. We release our dataset at <https://doi.org/10.5281/zenodo.1479136>.

Limitations and Future Work. Our dataset has some limitations due to restricted access to APIs, which may result in an incomplete view of the discourse. However, data from multiple platforms were collected to provide a more comprehensive perspective. Additionally, as noted in [11], the list of social identities used may not fully align with the descriptions people use on social media, necessitating regular updates to stay current with evolving social media language. Future work could explore the evolution of content through bridges to gain deeper insights into their functionality.

⁵https://blog.x.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm

5 Conclusion

Our study introduces the concept of bridging nodes on social media. Our investigations trace social media discourse on two 2024 US hurricanes on X, YouTube and Reddit. By modeling the similarity of narratives through network graphs, we can identify how these narratives are connected (via content), and who connect narratives (via users). These connections are termed as *bridging nodes*. While only less than 1% of nodes are bridging nodes, these nodes serve as important links for information flow. Removal of these bridges breaks up the connectivity of the graph, and therefore the path of information transversal breaks too. Deeper analysis of these bridging nodes through network cues, linguistic cues, topic analysis and user information show key properties that differentiate these bridging nodes from non-bridging nodes. We hope our work facilitates downstream studies of information diffusion, and can inform public officials and disaster aid organizations on how they should position themselves and their content for announcements to reach diverse communities.

References

- [1] S Axler, FW Gehring, and KA Ribet. 2000. Graduate Texts in Mathematics. *Graduate Texts in Mathematics* (2000).
- [2] Stefano Cresci, Maurizio Tesconi, Andrea Cimino, and Felice Dell'Orletta. 2015. A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages. In *Proceedings of the 24th international conference on world wide web*. 1195–1200.
- [3] Cedric De Boom, Steven Van Canneyt, Steven Bohez, Thomas Demeester, and Bart Dhoedt. 2015. Learning semantic similarity for very short texts. In *2015 ieee international conference on data mining workshop (icdmw)*. IEEE, 1229–1234.
- [4] Rongsheng Dong, Libing Li, Qingpeng Zhang, and Guoyong Cai. 2018. Information diffusion on social media during natural disasters. *IEEE transactions on computational social systems* 5, 1 (2018), 265–276.
- [5] ChangHwan Kim and Harris Hyun-soo Kim. 2023. Network ties, upward status heterophily, and unanticipated health consequences. *Journal of Health and Social Behavior* 64, 2 (2023), 313–332.
- [6] Wei Liu, Matteo Pellegrini, and Aiping Wu. 2019. Identification of bridging centrality in complex networks. *IEEE Access* 7 (2019), 93123–93130.
- [7] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27, 1 (2001), 415–444.
- [8] Isabel Murdock, Kathleen M Carley, and Osman Yağan. 2023. Identifying cross-platform user relationships in 2020 US election fraud and protest discussions. *Online Social Networks and Media* 33 (2023), 100245.
- [9] Lynnette Hui Xian Ng and Kathleen M Carley. 2023. Botbuster: Multi-platform bot detection using a mixture of experts. In *Proceedings of the international AAAI conference on web and social media*, Vol. 17. 686–697.
- [10] Lynnette Hui Xian Ng and Kathleen M Carley. 2023. Deflating the Chinese balloon: types of Twitter bots in US-China balloon incident. *EPJ Data Science* 12, 1 (2023), 63.
- [11] Lynnette Hui Xian Ng and Iain J Cruickshank. 2023. Recruitment promotion via Twitter: a network-centric approach of analyzing community engagement using social identity. *Digital Government: Research and Practice* 4, 4 (2023), 1–17.
- [12] Lynnette Hui Xian Ng, Iain J Cruickshank, and Kathleen M Carley. 2022. Cross-platform information spread during the January 6th capitol riots. *Social Network Analysis and Mining* 12, 1 (2022), 133.
- [13] Samantha C Phillips, Lynnette Hui Xian Ng, and Kathleen M Carley. 2022. Hoaxes and hidden agendas: A Twitter conspiracy theory dataset: Data paper. In *Companion Proceedings of the Web Conference 2022*. 876–880.
- [14] Andrey N Rabchevskiy, Victor S Zayakin, and Evgeny A Rabchevskiy. 2021. A method for identifying bridges in online social networks. In *International Conference on Analysis of Images, Social Networks and Texts*. Springer, 166–175.
- [15] Koustav Rudra, Niloy Ganguly, Pawan Goyal, and Saptarshi Ghosh. 2018. Extracting and summarizing situational information from the Twitter social media during disasters. *ACM Transactions on the Web (TWEB)* 12, 3 (2018), 1–35.
- [16] Jens M Schmidt. 2013. A simple test on 2-vertex-and 2-edge-connectivity. *Inform. Process. Lett.* 113, 7 (2013), 241–244.
- [17] Lynn Smith-Lovin, Dawn T Robinson, Bryan C Cannon, Jesse K Clark, Robert Freeland, Jonathan H Morgan, and Kimberly B Rogers. 2016. Mean affective ratings of 929 identities, 814 behaviors, and 660 modifiers by university of georgia and duke university undergraduates and by community members in durham, nc, in 2012–2014. *University of Georgia: Distributed at UGA Affect Control Theory*

- Website: <http://research.franklin.uga.edu/act> (2016).
- [18] Michael Szell and Stefan Thurner. 2010. Measuring social dynamics in a massive multiplayer online game. *Social networks* 32, 4 (2010), 313–329.
- [19] Cheng Zhang, Chao Fan, Wenlin Yao, Xia Hu, and Ali Mostafavi. 2019. Social media for intelligent public information and warning in disasters: An interdisciplinary review. *International Journal of Information Management* 49 (2019), 190–207.