INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Fan C, Jiang Y, Mostafavi A. 2020 Emergent social cohesion for coping with community disruptions in disasters. *J. R. Soc. Interface* **17**: 20190778.
http://dx.doi.org/10.1098/rsif.2019.0778

Received: 15 November 2019
Accepted: 7 February 2020

Subject Category:
Life Sciences–Physics interface

Subject Areas:
environmental science

Keywords:
emergent cohesion, disaster, social media, network dynamics

Authors for correspondence:
Chao Fan
e-mail: chfan@tamu.edu
Ali Mostafavi
e-mail: mostafavi@tamu.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.4860615.

THE ROYAL SOCIETY PUBLISHING

Emergent social cohesion for coping with community disruptions in disasters

Chao Fan[1], Yucheng Jiang[2] and Ali Mostafavi[1]

[1]Zachry Department of Civil and Environmental Engineering, Texas A&M University, College Station, TX 77840, USA
[2]Department of Computing and Information Science, Cornell University, Ithaca, NY 14850, USA

CF, 0000-0002-5670-7860; AM, 0000-0002-9076-9408

Social cohesion is an important determinant of community well-being, especially in times of distress such as disasters. This study investigates the phenomena of emergent social cohesion, which is characterized by abrupt, temporary and extensive social ties with the goal of sharing and receiving information regarding a particular event influencing a community. In the context of disasters, emergent social cohesion, enabled by social media usage, could play a significant role in improving the ability of communities to cope with disruptions in recent disasters. In this study, we employed a network reticulation framework to examine the underlying mechanisms influencing emergent social cohesion on social media while communities cope with disaster-induced disruptions. We analysed neighbourhood-tagged social media data (social media data whose users are tagged by neighbourhoods) in Houston, TX, USA, during Hurricane Harvey to characterize four modalities of network reticulation (i.e. enactment, activation, reticulation and performance) giving rise to emergent social cohesion. Our results show that, unlike regular social cohesion, communication history and physical proximity do not significantly affect emergent social cohesion. The results also indicate that weak social ties play an important role in bridging different social network communities, and hence reinforce emergent social cohesion. The findings can inform public officials, emergency managers and decision-makers regarding the important role of neighbourhood-tagged social media, as a new form of community infrastructure, for improving the ability of communities to cope with disaster disruptions through enhanced emergent social cohesion.

1. Introduction

The objective of this study is to characterize the dynamics of emergent social cohesion in neighbourhood-tagged social networks in coping with disaster disruptions. Information seeking and sharing are among the main components of human protective actions when facing disasters. With the increasing use of social media, online communication is increasingly prominent in information gathering and sharing among community members, especially during times of crisis [1,2]. This is because, when disasters cause inevitable disruptions in physical infrastructure and subsequent distress for humans, social networks enable information sharing and adjustment behaviours that play a key role in helping communities cope with disaster impacts [3,4]. Social cohesion has been shown to be an important element for human well-being [5,6], especially in times of distress and crisis [7]. In particular, the increasing use of social media during disasters could give rise to emergent social cohesion, which is characterized by abrupt, temporary and extensive social ties with the goal of sharing and receiving information regarding a particular event influencing a community [8]. Existing literature has already examined online social networks (OSNs) from multiple important aspects, such as fundamental structure of social networks [9,10]; information diffusion [11–14]; and social segregation [15,16]. Some methodologies and techniques have been employed to examine user activities, network structures and information propagation on social media for different purposes. For example, community network analysis has been used for examining roles of different actors/organizations, as well as information diffusion [17,18]. Topic analysis has also been used to evaluate important topics discussed by communities. While each of these studies (and the adopted analytics) provide insights about a particular modality in OSNs, they do not inform us about the relationships among various modalities influencing important phenomena such as emergent social cohesion. To this end, this study aims to examine various modalities and mechanisms in OSNs influencing emergent social cohesion in disaster-impacted communities using an integrated approach. Understanding emergent social cohesion in OSNs and its underlying dynamics may hold the key to promoting formal policies for social media selection and usage during disasters by city officials, emergency managers and community leaders.

In this study, we specifically examine the dynamics of emergent social cohesion in neighbourhood-tagged social media during disaster disruptions. Among different social media platforms, most of them, such as Twitter, tag their users with a city, a state or a country in their profiles to enable large-scale communication among masses of users regarding different topics and events (such as disaster-induced events) [19]. Community disruptions, however, tend to be localized and vary from one neighbourhood to another. For example, one neighbourhood may be affected by flooded roads while another neighbourhood might be dealing with sewage back-up. Without tags of neighbourhood information, it is hard for social media users to share and receive information regarding the same disruption event with users from the nearby neighbourhood and cope with disruptions. The examination of the effect of geographical proximity on emergent social cohesion is not feasible to do using Facebook and Twitter data since the distance and neighbourhood locations are not known. Hence, analysing neighbourhood-tagged OSNs makes it possible to uncover the relationships among neighbourhood-specific disruptive events, human activities and cohesion in OSNs. Such characterization would be unattainable through the use of other communication technology tools such as Twitter [20,21]. In this study, we examine the underlying dynamics of emergent social cohesion in OSNs using a theoretical network reticulation framework based on four modalities: enactment, activation, reticulation and network performance in neighbourhood-tagged OSNs. The study uses neighbourhood-level social media data (from Nextdoor) related to Hurricane Harvey in Houston, TX, USA, in analysing each modality of network reticulation and its relationship to the emergence and stability of social cohesion.

Hurricane Harvey was a category 4 tropical storm which affected Houston, the fourth largest city in the USA, from 26 August to 29 August in 2017 [22]. The extremely high intensity of the rainfall brought by Hurricane Harvey led to the water levels of Addicks and Barker reservoirs in the west of Houston rising to their maximum [23,24]. To prevent the reservoirs from abrupt disruptions, a large amount of water was released from the reservoirs; this water flooded the nearby 28 neighbourhoods. We collaborated with the volunteers living in these neighbourhoods and collected unique and publicly available data from their Nextdoor accounts in which the user profile's location is verified by the user's physical address and tagged by a neighbourhood name. The dataset spans 19 days, from 20 August to 7 September, including 7 days before Hurricane Harvey and 12 days during Hurricane Harvey and the subsequent flooding. Surveying the affected area, the identified 28 neighbourhoods have various numbers of users (electronic supplementary material, appendix, table S1). In total, our dataset includes 2690 active users who posted or commented at least one message, 1939 posts and 32 776 comments. Daily OSNs are mapped based on this dataset, in which active users are abstracted as nodes and the post-comment relationships are abstracted as edges (i.e. social ties) (electronic supplementary material, appendix, fig. S1).

2. Network reticulation theory

Communication theories such as Corman's network reticulation theory (NRT) [25], Giddens' structuration theory [26] and Homans' theory of human groups [27] have defined basic concepts and processes related to communication networks, such as the concept of triggering events, activity foci and communication relationships. These theoretical elements can enrich the existing analytics used for evaluating distinct modalities in OSNs and provide an integrative framework for examining the relationships among various modalities influencing important phenomena in OSNs. Hence, building upon these theoretical concepts from the communication field, we proposed an extended NRT as an integrative framework for examining four modalities and their relationships that affect emergent social cohesion in OSNs in disasters (figure 1*a*). The NRT framework used in this study characterizes the dynamics of social networks based on four modalities: enactment, activation, reticulation and performance. Each modality is a concept that relates the systematic phenomena to structural features in OSNs [26]. Enactment modality represents disruptive events (e.g. community disruptions and infrastructure failures) which trigger human activities in social networks. Events tend to evolve, trigger activities in individuals and cause transformation in the structure of OSNs. For example, the information-sharing behaviour when the environmental events occur and the situation evolves is more active than the behaviour before the events. Activation modality specifies user activities—posting and commenting about disruption events. Specifically, activities such as reporting and discussing situations in neighbourhoods (e.g. damage to housing, floodwater levels and road closures) stimulate communication instances in OSNs. Because of the dynamic nature of the unfolding of disaster-induced events, the activity themes (see the definition in Material and methods, §4.5) vary from one activity focus to another and over time. Accordingly, activity foci are identified to represent a cohesive cluster of online users corresponding to a particular activity theme in OSNs during a specific time period. Reticulation modality characterizes the structural properties of communication instances among clusters of OSNs corresponding to different activity foci. Specifically, reticulation modality signifies the creation and reinforcement of social ties among users regarding specific activity foci. Finally, network performance modality determines the influence of user activities on the structure of OSNs. In particular, we determine the measure of network assortativity at both the network and activity focus levels and their changes over time to signify the structural anatomy and stability of emergent social cohesion in OSNs. We employed the NRT framework and its four modalities in examining emergent social cohesion in

royalsocietypublishing.org/journal/rsif J. R. Soc. Interface 17: 20190778

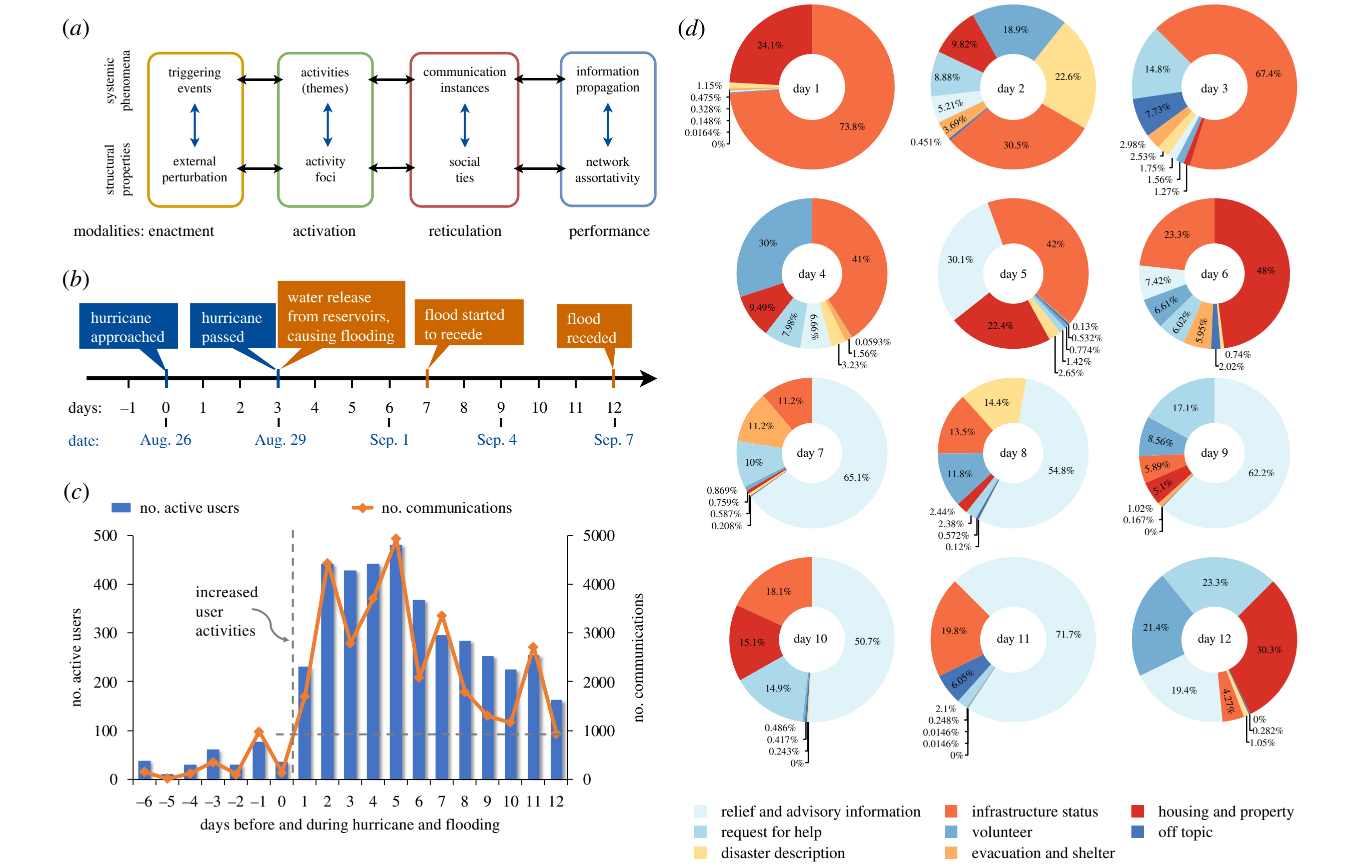

Figure 1. (*a*) Modalities of the NRT framework. Each modality (i.e. enactment, activation, reticulation and performance) is composed of two components: systemic phenomena and structural properties. (*b*) Disaster events happened in the neighbourhoods near reservoirs, including Hurricane Harvey and water release from reservoirs. The first day when the hurricane approached the neighbourhoods is tagged as 1. (*c*) Number of active users and communications on social media before and during the hurricane and flooding. (*d*) The proportion of eight themes in online communications among users from 28 neighbourhoods for 12 days, in which the proportions were weighted by the number of communications under the posts regarding a certain theme.

neighbourhood-tagged OSNs during Hurricane Harvey in 2017 and the subsequent flooding event in Houston.

2.1. Enactment modality

The enactment modality in the NRT framework focuses on examining the nature and timing of the triggering events (e.g. Hurricane Harvey and water release from reservoirs) to better understand their impacts on the affected neighbourhoods. As shown in figure 1*b*, Hurricane Harvey started having impacts on the neighbourhoods in the west of Houston on day 0 (26 August 2017). It dropped torrential and unprecedented amounts of rainfall over the Houston metropolitan area on days 1 and 2, and then weakened and moved towards Louisiana after day 3 (29 August 2017). Because of the heavy rainfall, the two major flood control infrastructures, Addicks and Barker reservoirs, reached their maximum capacity. To protect the reservoirs from breaching, the US Army Corps of Engineers decided to release the water from these two reservoirs without issuing any statement, inundating thousands of houses in nearby neighbourhoods [28]. The entire duration of the flooding lasted for 10 days, spanning from the end of the hurricane until the flood water receded. The houses in west Houston had never flooded before; and 80 per cent of the residents did not have any flood insurance [29]. Hurricane Harvey and the subsequent flooding event resulted in extreme panic among the affected neighbourhoods. People in these neighbourhoods had to seek situational information, look for relief resources and evacuate. As such, the communications of people on Nextdoor increased significantly during that period. As shown in figure 1*c*, more than 200 active users from these neighbourhoods communicated on Nextdoor each day during Hurricane Harvey and the flooding, generating more than 1000 posts and comments (see details in electronic supplementary material, appendix, table S6). The posts and comments are analysed in the same way, and both are considered as communications among online users. The peak of the communication and active users occurred at the end of Hurricane Harvey and the beginning of the flooding. When the flood started to recede, the activities of the users decreased as well.

2.2. Activation modality

In our NRT framework, the triggering events cause user activities related to seeking and sharing information to cope with disruptions. User activities on the neighbourhood-tagged social media include posting queries to seek help or ask questions, as well as commenting posts to respond or share information. The activation modality helps to link the enactment modality (triggering events) to the reticulation modality (communication instances). In the activation modality, online users organize and create the conditions that necessitate communication instances and social ties with each other. In this study, the activation modality identifies activity themes and foci, around which the reticulation of communications unfolds. Hence, collectively, four modalities enable the understanding of emergent social cohesion in OSNs during disasters.

From day 1 to day 2, when the hurricane landed, the most concerned activity theme was the status of infrastructure (accounting for 73.8% in all communications on day 1), including road, power and water (figure 1*d*). Beginning on day 3, water release from the flood control reservoirs started and the released water flooded the nearby neighbourhoods. Hence, again, infrastructure-related communication was the main activity theme discussed by users (accounting for 67.4% in all communications on day 3). Meanwhile, more and more requests or coordination of help from the users in these affected neighbourhoods were posted on the social media (7.98% of all communication on day 4). The damage to housing and properties became increasingly severe, and this theme increased from 9.49% of all communications on day 4, to 22.4% of all communications on day 5, and finally reached 48% of all communications on day 6. When the flood water started to recede on day 7, residents focused more on relief and advisory information (65.1% in all communications on day 7). The relief theme included information related to insurance, recovery tips as well as federal aid (accounting for more than 50% of all communications between days 7 and 11). When the flood water receded on day 12, multiple activities including volunteering, requests for help, relief information seeking, and housing and properties had equal shares of communications. This indicates that, as the triggering events dissipated, the prominence of the main activity themes dissipated as well. This analysis shows the direct relationship of activity themes with the timing of triggering events.

The primary activities and themes are the results of activity foci, which can be examined as structural clusters in OSNs communicating activity themes [25]. We employed the Louvain algorithm to detect the activity foci in the OSNs (see §4.1) [30]. In the activity foci, users collectively drew on information to co-act with each other and cope with the disruptions. The more severe the triggering events in the physical environment, the more individuals tend to find and develop new activity foci [7]. Consistent with the structural approach underlying the focus theory [31], the number of activity foci in the days at the beginning of flooding (especially from day 2 to day 6) is greater than that in other days (electronic supplementary material, appendix, table S3). This is because information seeking is an important component of human protective action and people had to look for a variety of information to cope with the adverse impacts of flooding (figure 1*d*).

Evaluation of the activity foci based on the structural network properties enables the evaluation of which users and activity themes give rise to emergent social cohesion. For example, figure 2*a* depicts the activity foci and their focused themes in OSN on day 6. The activity themes can be identified from the content generated by the users among the activity foci. Each activity focus has its own focused themes such as volunteer and infrastructure status, which brings users together to share information and resources. The presence of activity themes and foci further facilitates collective protective action, and hence enhances cohesiveness among users in activity foci. As shown in figure 2*b*, compared with the density of the entire network, the densities of the activity foci are significantly higher. This implies that users in each activity focus are densely connected to users within the same activity focus and have fewer edges connected to users outside the activity focus. The results are consistent with the results of modularity, which measures internal (and not external) connectivity (see electronic supplementary material, appendix, table S3). The higher the modularity, the more connected the activity focus (i.e. nodes in the activity foci are closely connected, and are less connected to nodes outside the activity foci) [32]. The information related to nodes and edges for each activity focus can be found in electronic supplementary material, appendix, table S4.

2.3. Reticulation modality

Activities by users on social media lead to the formation of communication instances (a.k.a. social ties). To determine the extent to which the increase in communication instances and social ties was due to the triggering events, we examined the effects of two latent factors: communication history and physical proximity of user locations.

2.3.1. Effects of communication history and physical proximity

Communication history was analysed to determine whether the emergent social cohesion was more influenced by triggering disruption events or by the past communication of users. We calculated the proportion of existing social ties which were established during 7 days before the hurricane landed (figure 2*c*). The results show that the proportions of prior social ties to post-event social ties are very low across the entire 12 days. All of the proportions do not exceed 4% and the proportions are even less than 0.5% in 7 days (i.e. days 2, 3, 4, 5, 8 and 11). These results demonstrate that, unlike the case of regular social cohesion, the communication history did not have a significant influence on the formation of social ties in emergent social cohesion.

Another latent factor is the physical proximity of users [33,34]. As all users are tagged by their neighbourhoods, we defined the physical proximity between users based on the distance between their neighbourhoods (electronic supplementary material, appendix, figure S3A). The measurement of the distances between neighbourhoods is defined in the Material and methods section. Electronic supplementary material, appendix, figure S3B displays the communication frequencies between neighbourhoods across the entire period of analysis (see the heatmaps for daily communication frequency in electronic supplementary material, appendix, figure S2). However, Pearson's correlation coefficient indicates that physical proximity has a weak correlation with the frequency of online communications between different neighbourhoods. Thus, physical distance does not hinder the process of emergent social cohesion in the context of community disruptions either. The findings imply that the main reticulation mechanisms (creation or reinforcement of social ties) were derived from activity themes triggered by disruptive events, and not by the communication history and physical proximity.

2.3.2. Spatial and temporal changes in social ties

We mapped the weight distribution of social ties for each day in the OSNs (figure 2*d*). The distribution of weights for social ties varies across different days. Generally, the weight for most edges is 1 regardless of whether the users connected by an edge are from the same neighbourhood or not. This is also shown by the median weights of the weight distribution in each day (see electronic supplementary material,

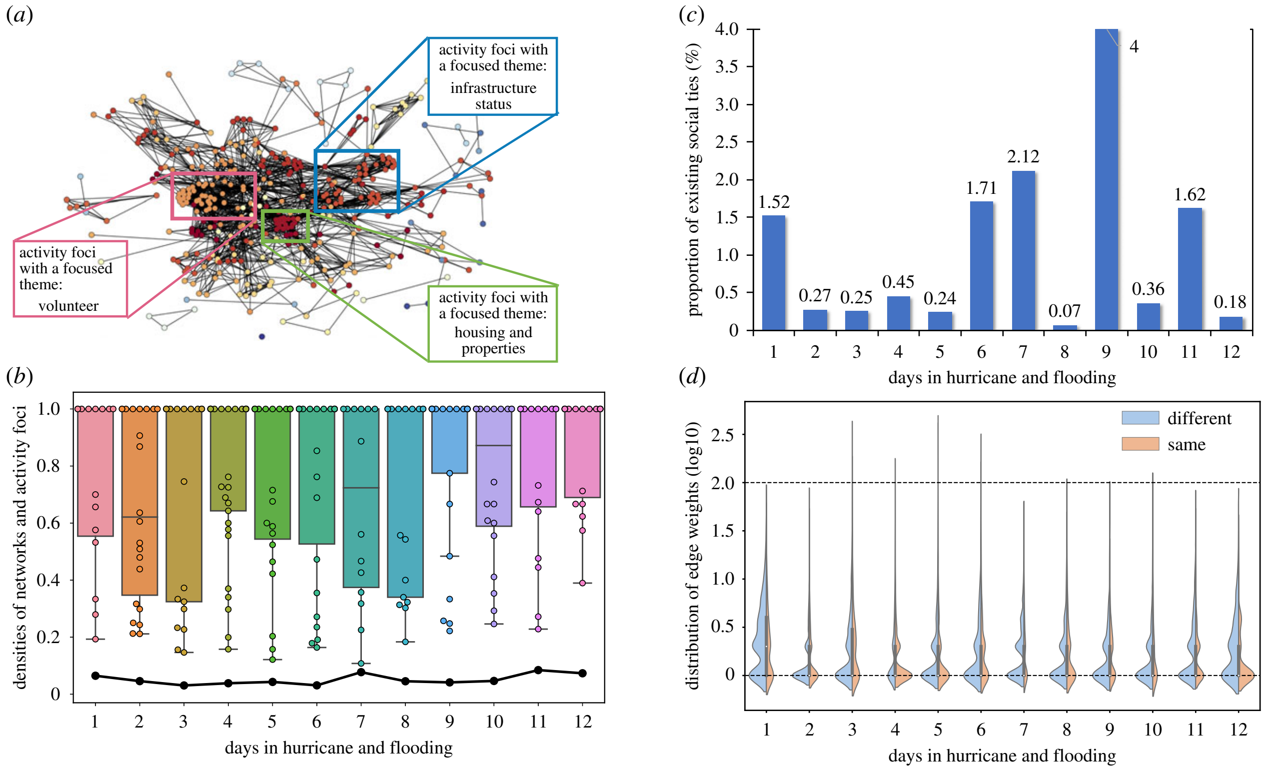

Figure 2. (*a*) An illustration of the focused themes of activity foci in the OSN of day 6. (*b*) Densities of daily networks (black dots) and activity foci (boxes and colourful dots). (*c*) The proportion of existing social ties in OSNs during the disasters (see Material and methods). (*d*) The distribution and mean of weights of social ties in each OSN. There are two types of social ties with regard to the neighbourhoods of the connected users: the social ties connect the users from the same neighbourhood (orange), and the social ties connect the users from different neighbourhoods (blue). The weights of social ties were obtained by the frequency of communications between two users.

appendix, table S5). In addition, the number of edges connecting users from different neighbourhoods is greater than the number of edges connecting users from the same neighbourhood. This phenomenon is more significant in the days when a disruption event occurred (e.g. day 1 when the hurricane started to land in the area, day 3 when water release from the reservoirs started, day 7 when the flood started to recede and day 12 when the flood water finally receded). This indicates that user activities immediately after triggering events go beyond the boundaries of their own neighbourhoods. Because the users in disasters are not self-contained or self-sufficient for information processing [10], they rely on information shared by others as well as on the reactions and sentiments of others in processing information regarding disaster threats. Thus, when the disaster situation evolves (i.e. triggering events occur), online users would create more weak social ties to gather new situational information from different users and examine their sentiments and reactions. This finding indicates that weak social ties connecting a large number of users from different neighbourhoods play a primary role in emergent social cohesion. Hence, unlike regular social cohesion with strong social ties [35], the weak ties are the building blocks of emergent social cohesion.

2.4. Performance modality

As discussed earlier, activity themes enable emergent social cohesion in groups of users (a.k.a., communities or activity foci) and create weak social ties between the users discussing the same topics. In this section, we investigate the network structural properties as a result of emergent social cohesion to gain deep insights into the role of neighbourhood-tagged social media in improving information propagation.

Mixing patterns in networks is an important approach to study the tendency for nodes in networks to be connected to other nodes that are similar (or dissimilar) to them in terms of selected node attributes such as node degree and user profile features (see Material and methods) [36]. To dissect the structure of emergent social cohesion and understand how the cohesion structure contributes to the information propagation, we examined the mixing patterns based on degree (figure 3*a*) and neighbourhood (figure 3*b*) attributes of the nodes and analysed the differences in the mixing patterns at two levels: network level and activity focus level.

2.4.1. Degree assortativity

Degree assortativity measures the extent to which nodes with similar degrees are connected to each other (e.g. whether gregarious people tend to associate with other gregarious people) [36]. Figure 3*a* shows that most of the activity foci in OSNs across the 12 days of the hurricane and flooding are disassortative. That is, the activity foci most often paired unlike nodes in which their degrees are quite different. In addition, the degree assortative mixing patterns also vary during the evolvement of the disaster situation, especially when new triggering events happened. For example, the mean of degree assortativity for the activity foci reached 0 (non-assortative) on day 3 when Hurricane Harvey passed and the flooding from the reservoirs started, and on day 7 when the flooding started to recede. During these days, when new events happened, activity foci lost the mixing tendencies. At the network level, however, the daily OSNs

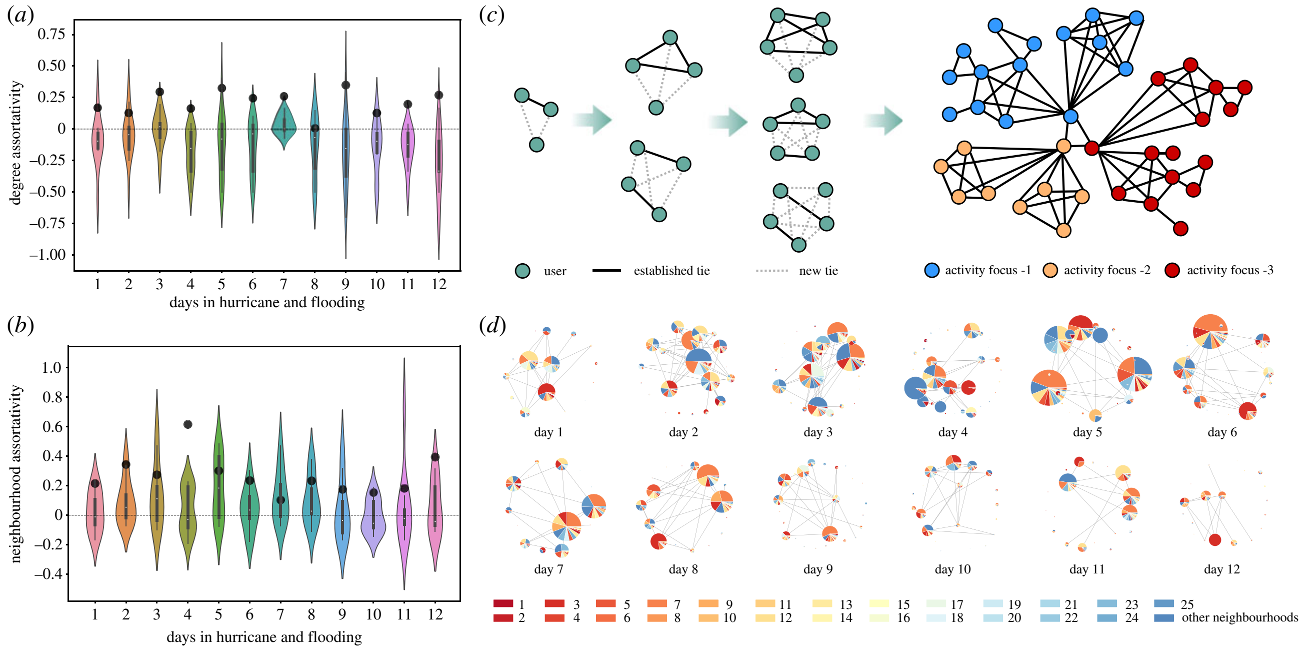

Figure 3. (*a*) Degree assortativity for daily OSNs (black dots); and the mean and distribution of the activity foci (colourful violin plot) in each day during the hurricane and flooding. (*b*) Neighbourhood assortativity for daily OSNs (black dots); and the mean and distribution of the activity foci (colourful violin plot) in each day during the hurricane and flooding. (*c*) Structural anatomy of daily OSNs for understanding emergent social cohesion and information propagation. (*d*) Activity focus-level OSNs with the proportions of various neighbourhoods by counting the number of users from the same neighbourhoods. Each pie chart represents an activity focus. The size of a pie chart is consistent with the number of users in the activity foci. The connections between communities depend on the connections between the users in both communities. There are three neighbourhoods without any active users on social media during disasters because there are only a few residents and corresponding registered users on Nextdoor. Meanwhile, a small group of people outside the investigated neighbourhoods joined the communication in disasters. This figure labels these neighbourhoods as 'other neighbourhoods'.

exhibit a tendency of assortative mixing (i.e. users with a similar degree connect to each other) (figure 3*a*). In evaluating the degree assortativity, for each daily network, the social ties among different activity foci in OSNs are considered. These ties cross the boundaries of activity foci and bridge the gaps among users in different activity foci and discussing different themes. In doing so, the degree assortativity of the networks increased by about 0.4 from the means of degree assortativity of the activity foci in each daily network.

Building upon this result, we can examine the structures of the social network emerging on Nextdoor by assembling the cliques in accordance with the mixing patterns (figure 3*c*). The posting and commenting functionality on Nextdoor enables users to form cliques in which users who commented in the same post are fully connected with each other. Thus, the social networks formed on Nextdoor are the result of assembling the cliques with certain mixing patterns. Initially, cliques emerged when disaster-related posts were generated by users and attracted the attention of other users to comment. Different users joined different posts based on their theme of interests. Users with similar interests in the themes form cohesive activity foci built upon multiple cliques. The activity foci gravitate more users with varying levels of degrees. This assembling mode contributes to the formation of the hierarchy in social networks with various degrees of users.

2.4.2. Neighbourhood assortativity

To assess the assortative mixing pattern for neighbourhoods (the extent to which activity foci in OSNs include users from the same neighbourhood), we calculated the assortativity coefficient of activity foci and examined the distribution and means of the assortativity over time (figure 3*b*). As the result shows, the means of neighbourhood assortativity coefficients for activity foci during Hurricane Harvey and flooding are around 0. The maximum value for the mean of assortativity coefficients is only 0.18. This result indicates that the majority of the activity foci in the emergent OSNs are non-assortative regarding neighbourhoods. That is, activity foci involved users from different neighbourhoods to generate and share information in disasters. To further support this finding, we examined the proportion of neighbourhoods in each activity focus in these 12 OSNs (figure 3*d*). Each pie chart was embedded in a node representing an activity focus. As shown in figure 3*d*, the majority of activity foci in OSNs are composed of multiple different neighbourhoods, regardless of different sizes of the activity foci. This result also provides evidence that, instead of the geographical boundaries, cohesive activity foci are more driven by the information needs of users related to triggering events and their impacts. Existing studies on Twitter show that social media has not affected geographical homophily (i.e. individuals from the same location tend to connect with each other) and recent empirical research on OSNs found that people still tend to connect more often to geographically close people [37,38]. These studies did not consider geographical homophily at the neighbourhood scale. Our results indicate that neighbourhood-tagged social media can enable users to break the physical boundaries of neighbourhoods and achieve cross-neighbourhood communication for information sharing and seeking to form cohesive activity foci.

Analysing the neighbourhood assortativity for the entire daily OSNs, we find that the majority of the daily OSNs have the neighbourhood assortativity coefficients 0.2–0.6 higher than 0 and the mean values of neighbourhood assortativity coefficients for their activity foci (figure 3*b*). The average value for the neighbourhood assortativity coefficients for these 12 days is 0.3, which signifies a weak assortativity for the neighbourhoods. In addition, the changes in the neighbourhood assortativity coefficients are consistent with the unfolding of disruption events. Specifically, day 3 (when Hurricane Harvey affected these neighbourhoods) and day 7 (the last day with high water level in neighbourhoods) both have low neighbourhood assortativity coefficients, which implies that users tend to get information from different neighbourhoods as part of their protective action information seeking.

Combining the findings at the activity focus level and network level, we can identify an important mixing pattern. The neighbourhood-tagged social media (Nextdoor) enabled users from different neighbourhoods to form cohesive activity foci to share and seek the information they needed. This also confirmed the role of activity foci in emergent social cohesion in OSNs, i.e. the primary reason for building social ties on social media is communication activities which motivate users to seek/share information regarding a triggering event. However, the social ties that cross the boundaries of activity foci and enable the spread of information across different activity foci tend to be created by users from the same neighbourhood. These users play an important role as boundary spanners contributing to the emergent social cohesion within and across neighbourhoods.

3. Discussion

Our proposed theoretical network reticulation framework uncovers the underlying modalities and structural network properties affecting emergent social cohesion in the context of community disruptions in disasters. Specifically, the findings in this study show that community disruption triggers user activities on social media, and users form activity foci for communicating information related to different themes. Then, weak social ties bridge the communication instances (i.e. activity foci) to enable the reticulation in networks, which subsequently shows a non-assortative mixing of users to promote information propagation across physical and online community boundaries. Although the case study focused on how residents cope with a hurricane and flooding event, the theoretical framework and many of our findings could be generalized to other crises and geographical contexts.

One key finding was that disaster events trigger the emergence and evolvement of human protective activities in seeking information on social media. This information seeking and sharing behaviour creates activity foci that gravitate additional users and leads to the creation of new links giving rise to emergent social cohesion. This finding implies the important role that neighbourhood-tagged social media, such as Nextdoor, plays in the formation of activity foci related to different triggering events for which residents seek information, and hence improves the individual and collective protective action of the residents. This finding also has implications for online influential users or community leaders to initiate activity themes/foci and enhance cohesiveness of users [39,40]. These individuals with boundary-spanning weak ties between neighbourhoods play an important role in scaling up the communication among neighbourhoods. Future studies can examine the bridging and boundary-spanning roles of these individuals to inform the efficient and effective spread of information among online users from different neighbourhoods.

Another key finding was that emergent social cohesion arises from cohesive activity foci which focus on specific disaster-related events and are constructed by weak social ties. In the disaster context, activity foci become the gravitation centres that absorb users (create weak ties) across different neighbourhoods with interest/need for information about the activity theme. Meanwhile, we observed that weak social ties were derived from the activity themes, instead of communication history or physical proximity. This finding further confirms that people use social media during disasters for information seeking as part of their protective action. In other words, the fundamental function of social media changes for people during disasters. Also, unlike regular social cohesion with strong social ties and physical homophily, the formation of inclusive activity foci and weak ties triggered by disaster events are the building blocks of emergent social cohesion. This result is also consistent with the principles in homophily/heterophily theory that weak social ties are more heterophilous than strong ties [41,42]. Accordingly, creating weak social ties not only contributes to the inclusiveness of activity foci, but also has a significant influence on emergent social cohesion and information spread. This is because heterophilous communication facilitates the flow of information between diverse segments of a social network [42,43].

The third key finding was that different assortative mixing patterns for OSNs and activity foci improve the information spread. This happens because the assemblage of multiple activity foci makes OSNs assortatively mixed by creating social ties between users with similar degrees and from the same neighbourhoods. At the same time, activity foci themselves are composed of the ties between users with different degrees and from different neighbourhoods. The findings illustrate the assemblage process of OSNs and the formation of the hierarchy in terms of the node degree and neighbourhood attributes. This self-organized mixing pattern could give rise to emergent influential users who could be identified and used to seed urgent safety-related information and speed up the diffusion of information [44]. Identifying users who bridge the boundaries of activity foci using the degree and neighbourhood features and seeding the information to these users can optimize information propagation and help users to better perceive disaster situations.

With the increased use of social media in disasters, emergency managers, public officials and community leaders would need to optimize their strategies to improve information seeking protective action in communities. Understanding the structure and dynamics of social networks could provide information about better intervention strategies to improve the spread of situational information, the situation awareness of the affected population and the response and recovery of neighbourhoods in disasters. Our theoretical network reticulation framework and the findings may provide important insights for public officials, emergency managers and community leaders regarding social media selection and usage policies for improving the capacity of communities in coping with disasters. Additionally, the proposed integrative framework can be further adopted in other studies and contexts to examine the dynamics of OSNs based on the characteristics and relationships among the four modalities. Therefore, the next critical question is how the identified emergent social cohesion varies from one disaster to another. Moreover, user behaviours on multiple social media tools could be examined in future studies to provide further information about emergent social cohesion in disasters. A challenging problem, however, is to collect quality data from different social media platforms for the same areas. Moreover, future studies can combine the data about phone calls with social media data to further examine social network interactions (specifically, when there are disruptions in Internet and telecommunication services). Nevertheless, emergent social cohesion would primarily arise as a result of the formation of weak social ties on OSNs (phone calls are usually made to existing contacts, which reinforces existing social ties rather creating new communication instances).

4. Material and methods

The data were gathered by volunteers in different neighbourhoods one week after Hurricane Harvey. The neighbourhoods in this study are defined by Nextdoor. Since Nextdoor allows users to access all their historical data, the volunteers could gather communications for the period before the disaster as well. The volunteers used their personal Nextdoor accounts to gather communications among users in their own and nearby neighbourhoods using the public posts. A public post is a message that the user consents to be publicly available for all users from nearby neighbourhoods rather than only to his designated users. Communications among users through public posts were gathered anonymously. Nextdoor only allows users to show their physical addresses, and no organizational or association identities are allowed to be disclosed. Social/institutional identity cannot be observed and collected from user profiles. Hence, there are no institutional mechanisms that affect communication among neighbourhoods.

The neighbourhoods were selected based on two criteria. First, these neighbourhoods were flooded when the water was released from the reservoirs in west Houston during Hurricane Harvey. The water release was a major disruption and caused a significant impact; it thus provided an ideal setting for examining emergent social cohesion. Severe damage happened in the downstream area (see electronic supplementary material, appendix, figure S4). Since the impact of flooding was almost the same across different neighbourhoods, we assumed that the use of online communication was not differentiated by the extent of impact experienced by these neighbourhoods. Second, Nextdoor has some constraints related to the neighbourhoods that a user can communicate with. The neighbourhoods that we selected are within the geographical area affected by the release of water from the reservoirs whose users could communicate with each other. The sizes of the neighbourhoods are shown in electronic supplementary material, appendix, table S1.

4.1. Activity foci and network modularity

An activity focus is a structural cluster in which users are connected to each other more closely than they are connected to users outside the activity focus. The idea of detecting activity foci was named community detection in computer science. Modularity maximization (i.e. Louvain heuristic [45]) is one of the widely used approaches for detecting communities which measure the modularity of a network to examine how well a network is partitioned into communities. This approach compares the number of edges within a certain group with the expected number of edges in a null model. The modularity for a network is formulated as [36,46]

$$Q = \frac{1}{2m} \cdot \sum_{ij} \left[A_{ij} - \frac{d_i d_j}{2m} \right] \cdot \delta(d_i, d_j), \tag{4.1}$$

where m is the number of edges in the network, A_{ij} is the adjacency matrix of the network, and d_i and d_j are the degrees of the nodes i and j. The higher the modularity (closer to 1), the more edges within the module that we expect by chance [47]. Generally, a modularity higher than 0.3 means a significant community structure. In this study, we adopted the Louvain algorithm to detect communities by maximizing the modularity of networks.

4.2. Communication history

Communication activities among users on social media before the hurricane approached the locations are considered to be the communication history. In this study, the communication history is identified using the existing social ties established during the 7 days before Hurricane Harvey. Then, we can measure the effect of communication history on the communication behaviours during the hurricane and flooding by

$$p = \frac{|E'|}{|E|}, \tag{4.2}$$

where $|E'|$ is the number of social ties that have been created before Hurricane Harvey and $|E|$ is the total number of social ties in the OSN. A low proportion of social ties created previously can indicate the low effect of communication history on emergent social cohesion during disasters.

4.3. Physical proximity

To examine the effects of physical proximity on social cohesion in OSNs, physical proximity was defined to measure the distance between different neighbourhoods [48]. This study developed three levels of physical proximity: 'within a neighbourhood', i.e. both of the users live in the same neighbourhood; 'nearby neighbourhoods', i.e. the users live in neighbourhoods bordering each other; and 'distant neighbourhoods', i.e. the users live in neighbourhoods which do not border each other.

4.4. Communication frequency

Communication frequency within and between neighbourhoods was a measurement of the cohesiveness between two neighbourhoods. The matrix of communication frequency between neighbourhoods was used to calculate the correlation with physical proximity and illustrates the effects of physical proximity on emergent social cohesion. The communication frequency was computed by

$$f = \frac{\sum_{u \in U, v \in V} \text{weight}(u,v)}{|U| \cdot |V|}, \tag{4.3}$$

in which U and V are collections of users in neighbourhoods. U and V can be the same neighbourhood (i.e. $|V| = |U| - 1$), in which the results show the communication frequency of users within the same neighbourhood. The communication frequency of users in different neighbourhoods can be obtained when U and V represent the users from different neighbourhoods. Whether users come from the same neighbourhood or not, as shown in equation (4.3), we deem each communication to be from two neighbourhoods (they may be the same). This approach can overcome the discrepancy that arises from the differences in the number of users in different neighbourhoods.

4.5. Content coding

Two researchers manually coded each post for message content in our dataset by using a content-coding approach proposed in prior research [49]. Based upon the content themes presented in [1], this study developed a new coding ontology, which was fitted to the content categories on our dataset, including eight themes: housing and properties, i.e. information about damage to housing, loss of property and casualties; infrastructure status, i.e. information about infrastructure facilities; evacuation/shelter, i.e. information about pre-evacuations, mandatory evacuations and sheltering information; disaster descriptions, i.e. descriptions of the disaster itself and its scales; relief/advisory information, i.e. relief information and response tips; requests for help, i.e. actions about requesting neighbours' on-site help; volunteer, i.e. information about volunteering or providing individual help to neighbours; and off topic, i.e. the posts are not within the scope of the disaster-related topics. The specific definition and examples of posts are available in electronic supplementary material, table S2. A set of posts were split-recorded by two coders and they exchanged and checked the codes for intercoder agreement. The coders agreed on theme codes in 96% of cases and the disagreements were resolved through discussion and consensus. The proportions of each theme in the daily communications were calculated by taking into account the number of comments belonging to the posts coded as the theme. This computing approach is beneficial to exhibit the contribution of themes to the formation of social ties as well as to emergent social cohesion.

4.6. Assortative mixing in networks

A good measure of the extent to which nodes with similar degrees or attributes connect to each other is assortativity. To quantify the level of assortative mixing of attributes in a community, this study employed equation (4.4) to compute the assortativity coefficient [36]

$$r = \frac{Tr(M) - \|M^2\|}{1 - \|M^2\|}, \tag{4.4}$$

where M is the mixing matrix of the attribute and $\|M^2\|$ is the sum of all elements of the matrix M^2. Here, for the purpose of our analysis, the examined attribute is users' neighbourhoods. The assortativity coefficient ranges from −1 to 1. The larger the assortativity coefficient, the more perfect the assortative mixing. Specifically, $r = 1$ when there is perfect assortative mixing, $r = 0$ when there is no assortative mixing and r is negative when there is disassortative mixing [36].

Data accessibility. Owing to Nextdoor's policy, we cannot publicly share the raw dataset used in this study. However, aggregated results from which the analyses can be recreated are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.gqnk98sht [50].

Authors' contributions. All the authors designed the research; C.F. and Y.J. performed the research and analysed the data; and C.F. and A.M. wrote the paper with comments from all authors.

Competing interests. The authors declare no competing interest.

Funding. We received no funding for this study.

Acknowledgements. This material is based in part upon work supported by an Amazon Web Services (AWS) Machine Learning Award. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Amazon Web Services.

References

1. Sutton J, Gibson CB, Phillips NE, Spiro ES, League C, Johnson B, Fitzhugh SM, Butts CT. 2015 A cross-hazard analysis of terse message retransmission on Twitter. *Proc. Natl Acad. Sci. USA* **112**, 14 793–14 798. (doi:10.1073/pnas.1508916112)
2. Zhang C, Fan C, Yao W, Hu X, Mostafavi A. 2019 Social media for intelligent public information and warning in disasters: an interdisciplinary review. *Int. J. Inf. Manage.* **49**, 190–207. (doi:10.1016/j.ijinfomgt.2019.04.004)
3. Balcik B, Beamon BM, Krejci CC, Muramatsu KM, Ramirez M. 2010 Coordination in humanitarian relief chains: practices, challenges and opportunities. *Int. J. Prod. Econ.* **126**, 22–34. (doi:10.1016/j.ijpe.2009.09.008)
4. Akhtar P, Marr NE, Garnevska EV. 2012 Coordination in humanitarian relief chains: chain coordinators. *J. Humanit. Logist. Supply Chain Manage.* **2**, 85–103. (doi:10.1108/20426741211226019)
5. Roca CP, Helbing D. 2011 Emergence of social cohesion in a model society of greedy, mobile individuals. *Proc. Natl Acad. Sci. USA* **108**, 11 370–11 374. (doi:10.1073/pnas.1101044108)
6. Misra S, Goswami R, Mondal T, Jana R. 2017 Social networks in the context of community response to disaster: study of a cyclone-affected community in Coastal West Bengal, India. *Int. J. Disaster Risk Reduct.* **22**, 281–296. (doi:10.1016/j.ijdrr.2017.02.017)
7. Lazega E, Wasserman S, Faust K. 2006 Social network analysis: methods and applications. *Rev. Française Sociol.* **36**, 781–783. (doi:10.2307/3322457)
8. Kim YC, Kang J. 2010 Communication, neighbourhood belonging and household hurricane preparedness. *Disasters* **34**, 470–488. (doi:10.1111/j.1467-7717.2009.01138.x)
9. Sekara V, Stopczynski A, Lehmann S. 2016 Fundamental structures of dynamic social networks. *Proc. Natl Acad. Sci. USA* **113**, 9977–9982. (doi:10.1073/pnas.1602803113)
10. Prasad S, Zakaria R, Altay N. 2018 Big data in humanitarian supply chain networks: a resource dependence perspective. *Ann. Oper. Res.* **270**, 383–413. (doi:10.1007/s10479-016-2280-7)
11. Qiu X, Oliveira DFM, Sahami Shirazi A, Flammini A, Menczer F. 2017 Limited individual attention and online virality of low-quality information. *Nat. Hum. Behav.* **1**, 1–7. (doi:10.1038/s41562-017-0132)
12. Vosoughi S, Roy D, Aral S. 2018 The spread of true and false news online. *Science* **359**, 1146–1151. (doi:10.1126/science.aap9559)
13. Morone F, Makse HA. 2015 Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68. (doi:10.1038/nature14604)
14. Altay N, Pal R. 2014 Information diffusion among agents: implications for humanitarian operations. *Prod. Oper. Manage.* **23**, 1015–1027. (doi:10.1111/poms.12102)
15. Wang Q, Phillips NE, Small ML, Sampson RJ. 2018 Urban mobility and neighborhood isolation in America's 50 largest cities. *Proc. Natl Acad. Sci. USA* **115**, 7735–7740. (doi:10.1073/pnas.1802537115)
16. Fan C, Wu F, Mostafavi A. 2020 A Hybrid machine learning pipeline for automated mapping of events and locations from social media in disasters. *IEEE Access* **8**, 10 478–10 490. (doi:10.1109/ACCESS.2020.2965550)
17. Papadopoulos T, Gunasekaran A, Dubey R, Altay N, Childe SJ, Fosso-Wamba S. 2017 The role of Big Data in explaining disaster resilience in supply chains for sustainability. *J. Clean. Prod.* **142**, 1108–1118. (doi:10.1016/j.jclepro.2016.03.059)
18. Fan C, Zhang C, Yahja A, Mostafavi A. 2020 Disaster city digital twin: a vision for integrating artificial and human intelligence for disaster management. *Int. J. Inf. Manage.* 1–12. (doi:10.1016/j.ijinfomgt.2019.102049)
19. Bagrow JP, Liu X, Mitchell L. 2019 Information flow reveals prediction limits in online social activity. *Nat. Hum. Behav.* **3**, 122–128. (doi:10.1038/s41562-018-0510-5)
20. Kryvasheyeu Y, Chen H, Obradovich N, Moro E, Van Hentenryck P, Fowler J, Cebrian M. 2016 Rapid assessment of disaster damage using social media

activity. *Sci. Adv.* **2**, e1500779. (doi:10.1126/sciadv.1500779)

21. Fan C, Jiang Y, Yang Y, Zhang C, Mostafavi A. 2020 Crowd or hubs: information diffusion patterns in online social networks in disasters. *Int. J. Disaster Risk Reduct.* **46**, 101498. (doi:10.1016/j.ijdrr.2020.101498)
22. Sebastian T *et al*. 2017 Hurricane Harvey report: a fact-finding effort in the direct aftermath of Hurricane Harvey in the Greater Houston region. See https://www.researchgate.net/publication/321749153_Hurricane_Harvey_Report_A_fact-finding_effort_in_the_direct_aftermath_of_Hurricane_Harvey_in_the_Greater_Houston_Region.
23. Tang F, O'Neil E. 2017 How the Barker and Addicks dams work. *Houston Chronicle*. See https://www.houstonchronicle.com/local/gray-matters/article/How-the-Barker-and-Addicks-dams-work-12171719.php (accessed on 5 June 2018).
24. Fan C, Mostafavi A. 2019 A graph-based method for social sensing of infrastructure disruptions in disasters. *Comput. Civ. Infrastruct. Eng.* **24**, 1055–1070. (doi:10.1111/mice.12457)
25. Corman SR, Scott CR. 1994 Perceived networks, activity foci, and observable communication in social collectives. *Commun. Theory* **7**, 171–190. (doi:10.1111/j.1468-2885.1994.tb00089.x)
26. Giddens A. 1984 The constitution of society: outline of the theory of structuration. *Cognit. Ther. Res.* **12**, 448. (doi:10.1007/BF01173303)
27. Homans GC. 1950 *The human group*. New York, NY: Harcourt, Brace and Company.
28. Allen K. 2017 Army Corps releases water from 2 Houston dams; thousands of homes to be affected. ABC News. See https://abcnews.go.com/US/army-corps-releases-water-houston-dams-levels-increased/story?id=49462262 (accessed on 18 February 2019).
29. Shilcutt K, Asgarian R. 2017 *The deliberate flooding of west Houston*. Houston, TX: Houstonia. See https://www.houstoniamag.com/articles/2017/10/16/barker-addicks-reservoirs-release-west-houston-memorial-energy-corridor-hurricane-harvey (accessed 6 July 2019)
30. Waltman L, Van Eck NJ. 2013 A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* **86**, 471. (doi:10.1140/epjb/e2013-40829-0)
31. Burt RS. 1984 Network items and the general social survey. *Soc. Netw.* **6**, 293–339. (doi:10.1016/0378-8733(84)90007-8)
32. Newman MEJ, Girvan M. 2004 Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113. (doi:10.1103/PhysRevE.69.026113)
33. Burt RS. 1987 Social contagion and innovation: cohesion versus structural equivalence. *Am. J. Sociol.* **92**, 1287–1335. (doi:10.1086/228667)
34. Jackson JC, Jong J, Bilkey D, Whitehouse H, Zollmann S, McNaughton C, Halberstadt J. 2018 Synchrony and physiological arousal increase cohesion and cooperation in large naturalistic groups. *Sci. Rep.* **8**, 1–8. (doi:10.1038/s41598-017-18023-4)
35. Kawachi I, Berkman L. 2000 Social cohesion, social capital, and health. *Soc. Epidemiol.* **1**, 174–190.
36. Newman MEJ. 2003 Mixing patterns in networks. *Phys. Rev. E* **67**, 13. (doi:10.1103/PhysRevE.67.026126)
37. Kulshrestha J, Kooti F, Nikravesh A, Gummadi K. 2012 Geographic dissection of the Twitter network. In *Proc. of the 6th Int. AAAI Conference on Weblogs and Social Media, Dublin, Ireland, 5–7 June 2012*. Palo Alto, CA: AAAI Press.
38. Choudhury M, Counts S, Czerwinski M. 2011 Identifying relevant social media content: leveraging information diversity and user cognition. In *Proc. 22nd ACM Conf. Hypertext hypermedia - HT '11, Eindhoven, The Netherlands, 6–9 June 2011*. New York, NY: Association for Computing Machinery.
39. Patterson O, Weil F, Patel K. 2010 The role of community in disaster response: conceptual models. *Popul. Res. Policy Rev.* **29**, 127–141. (doi:10.1007/s11113-009-9133-x)
40. Yang Y, Zhang C, Fan C, Yao W, Huang R, Mostafavi A. 2019 Exploring the emergence of influential users on social media during natural disasters. *Int. J. Disaster Risk Reduct.* **38**, 101204. (doi:10.1016/J.IJDRR.2019.101204)
41. Gross R, Acquisti A, Heinz HJ. 2006 Information revelation and privacy in online social networks. In *Proc. of the 2005 ACM Workshop on Privacy in the Electronic Society, Alexandria, VA, 7 November 2005*. New York, NY: Association for Computing Machinery. (doi:10.1145/1102199.1102214)
42. Brown JJ, Reingen PH. 2002 Social ties and word-of-mouth referral behavior. *J. Consum. Res.* **14**, 350–362. (doi:10.1086/209118)
43. Aral S, Walker D. 2014 Tie strength, embeddedness, and social influence: a large-scale networked experiment. *Manage. Sci.* **60**, 1352–1370. (doi:10.1287/mnsc.2014.1936)
44. Aral S, Dhillon PS. 2018 Social influence maximization under empirical influence models. *Nat. Hum. Behav.* **2**, 375–382. (doi:10.1038/s41562-018-0346-z)
45. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. 2008 Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, 1–12. (doi:10.1088/1742-5468/2008/10/P10008)
46. Montes F, Jimenez RC, Onnela J-P. 2017 Connected but segregated: social networks in rural villages. *J. Complex Netw.* **6**, 693–705. (doi:10.1093/comnet/cnx060)
47. Foster JG, Foster DV, Grassberger P, Paczuski M. 2010 Edge direction and the structure of networks. *Proc. Natl Acad. Sci. USA* **107**, 10 815–10 820. (doi:10.1073/pnas.0912671107)
48. Gibbons J, Nara A, Appleyard B. 2018 Exploring the imprint of social media networks on neighborhood community through the lens of gentrification. *Environ. Plan. B Urban Anal. City Sci.* **45**, 470–488. (doi:10.1177/2399808317728289)
49. Sutton J, Spiro ES, Fitzhugh S, Johnson B, Gibson B, Butts CT. 2014 Terse message amplification in the Boston bombing response. In *ISCRAM 2014 Conf. Proc. - 11th Int. Conf. Inf. Syst. Cris. Response Manag., University Park, PA, 18–21 May 2014*, pp. 612–621. University Park, PA: Pennsylvania State University.
50. Fan C, Jiang Y, Mostafavi A. 2020 Data from: Emergent social cohesion for coping with community disruptions in disasters. Dryad Digital Repository. (https://doi.org/10.5061/dryad.gqnk98sht)