A Graph-based Approach for Detecting Critical Infrastructure Disruptions on Social Media in Disasters

Chao Fan
Texas A&M University
chfan@tamu.edu

Ali Mostafavi
Texas A&M University
amostafavi@civil.tamu.edu

Wenlin Yao
Texas A&M University
wenlinyao@tamu.edu

Ruihong Huang
Texas A&M University
huangrh@cse.tamu.edu

Abstract

The objective of this paper is to propose and test a graph-based approach for detection of critical infrastructure disruptions in social media data in disasters. Understanding the situation and disruptive events of critical infrastructure is essential to effective disaster response and recovery of communities. The potential of social media data for situation awareness during disasters has been highlighted in recent studies. However, the application of social sensing in detecting disruptions of critical infrastructure is limited because existing approaches cannot provide complete and non-ambiguous situational information about critical infrastructure. Therefore, to address this methodological gap, we developed a graph-based approach including data filtering, burst time-frame detection, content similarity and graph analysis. A case study of Hurricane Harvey in 2017 in Houston was conducted to illustrate the application of the proposed approach. The findings highlighted the temporal patterns of critical infrastructure events that occurred in disasters including disruptive events and their adverse impacts on communities. The findings also provided insights for better understanding critical infrastructure interdependencies in disasters. From the practical perspective, the proposed methodology study can improve the ability of community members, first responders and decision makers to detect and respond to infrastructure disruptions in disasters.

1. Introduction

Being aware of how a community's situation evolves in extreme events like hurricanes, wildfires, or earthquakes is crucial to effective disaster response and recovery [1]. A key component of community is the critical infrastructure, the damages and failures of which make severe impacts on human, environment, and the economy [2], [3]. Hence, a better understanding of the situation and performance of critical infrastructure is important to effective disaster response and recovery.

With the emergence of Digital Twin paradigm [4] to achieve more resilient and smarter cities, disaster response and recovery should be data-centric and network-centric [5] to provide better situation awareness regarding spatiotemporal fluctuations in disasters. However, the situation of critical infrastructure in disasters changes over time due to the disruptions and cascading failures. Usually, physical and remote sensing techniques (e.g., satellite and UAV platforms) are primary techniques for collecting infrastructure data in normal situations and disasters [6]. However, due to the rapidly evolving situations in the aftermath of a disaster, specific data about the performance of critical infrastructure might not be collected in the required time and space [7]. Alternatively, humans as sensors on social media share information about disaster situations, including critical infrastructure disruptions [8]. For example, "Closed due to flooding, in #WestSide on W Sam Houston Tollway Frontage Rd NB between Westheimer and the I-10 Katy Fwy #traffic". This tweet is related to road closure during Hurricane Harvey. Hence, improving methods for detecting infrastructure disruptions enables automated evaluation of spatiotemporal changes to feed information into the digital twin of smart cities during disasters.

Recent studies have highlighted the importance and applications of social sensing in detecting relevant events and understanding the situation for disaster response and recovery [9], [10]. For example, a study [11] about social sensing for urban crisis management developed an event analysis framework covering human emotions and behaviors in response to Singapore Haze on social media. This study focuses on human networks

URI: https://hdl.handle.net/10125/59637
ISBN: 978-0-9981331-2-6

and their activities that mentioned in social media content. Another example of social sensing for disaster response investigated the dynamics of social networks and the formation and evolution of online communities in the 2011 Japanese Earthquake and Tsunami [12]. This study processed the content of social posts and provided insights for understanding how patterns of human interaction are changed by external attacks. Another branch of studies related to social sensing focus on mapping the flood inundation and social emotions based on the geographic information in social media posts [13]. For example, Fohringer [14] utilized quantitative data derived from photos in social media posts in disasters to support the inundation mapping [14]. The approach enables rapid flood mapping, estimation of flood risks, and determination of response actions.

While the use of social sensing in disasters is growing, little of the existing work focuses on detecting critical infrastructure disruptions. One reason is that effective methodologies for detecting complete and non-ambiguous situational information related to specific critical infrastructure are missing. Such methodology should consider the timing and severity of disruptions, effects of disruptions on people, and actions that are taken to adjust to the disruptions. Thus, to address such methodological gap, this paper proposed a graph-based approach to detect the situation and corresponding changes of critical infrastructure in social media data. A case study related to the situation of Barker and Addicks reservoirs in Houston area during Hurricane Harvey was conducted to illustrate the application and capability of the proposed approach.

2. Background

Detecting situational information including human behaviors and physical events on social media during occurrences of disasters has been studied by a number of researchers. In our survey of the related work, this paper focused on the approaches that could be applied to large-scale datasets such as millions of tweets. The existing techniques for event detection include clustering [15],[16] and network analysis [17].

Most of the techniques for unspecified event detection on social media data is based on clustering. Features including words and phrases extracted from social messages are the primary objects in the processes. For example, Weng and Lee developed an Event Detection with Clustering of Wavelet-based Signal of words (*EDCoW*) [18]. However, this approach treats each word independently. The identified events are likely to be a group of words associated with different events. Thus, the information detected from this approach is not reliable for situation awareness in disasters. Besides, Pohl et al. proposed a two-phase clustering approach to identifying individual sub-events within a crisis automatically [19]. However, the results are summarized into several discrete words which are difficult to be understood by users. To aggregate and categorize significant events, Ritter et al. discovered event categories and classified events based on latent variable models [20]. However, the approach is for open-domain analysis. Thus, the outcomes of this approach are comparatively general and cannot identify specific infrastructure disruptions in the context of disasters.

Some studies conducted event detection on social posts by using graphs or networks [21]. However, some of them only focused on social networks and cannot demonstrate the relationships between social actors and events [22]. For example, Zhao et al. detected events by combining text-based clustering, temporal segmentation and graph cuts of social networks [23]. This study provides insights into exploring temporal and social information together with text content. But, their relationships cannot be validated, and the results are token-based. Thus, this approach is limited to improving situation awareness of specific infrastructure disruptions. Another example is developed by Wang and Taylor [24]. They explored geographical and semantical dimensions of events from tweets. However, the approach can only be applied to geotagged tweets which are very limited in the complete Twitter dataset. As a result, events detected by this approach could be missing important information related to the evolving nature and timing of infrastructure disruptions.

In summary, existing approaches and techniques for event detection on social media data are feature-based methods, which study the distribution of words and detect the events by grouping words together. The users need to infer the logic and meaning behind the words and come up with the scenarios based on their inferences. Such results would lack ground truth evidence, and the findings are mostly subjective. The situational information of critical infrastructure disruption should involve when disruption happened, how it affected residents, what caused this event and what actions were taken. However, such information cannot be obtained by using existing approaches for social media data analysis. Improving event detection for social media data is essential to evaluate infrastructure disruptions in disasters. Hence, the findings from social sensing data for detecting infrastructure disruption events can play an important role in developing a digital twin paradigm for a city impacted by natural disasters [4].

3. Graph-based approach

Tweets, which allow up to 280 characters in one message, tend to deliver complete and non-ambiguous information about critical infrastructure situations. That is because the tweets can include complete sentences and informative phrases talking about situations in disasters [25],[26]. Also, high-impact events and credible information tend to be repeated on tweets and retweeted many times, so the frequency of tweets about the same event is supposed to show a burst in a short period [27]. Based on above observations and assumptions, the underlying premise of this study is to identify and analyze high-impact events and credible tweets that inform about critical infrastructure disruptions. The proposed graph-based approach includes four steps: data filtering, burst detection, content similarity, and graph analysis to identify critical tweets and better understand the situation of critical infrastructure disruption (see Figure 1). The details of our approach are described in the following sections. The outputs of our graph-based approach are critical tweets which contain the complete information about high-impact critical infrastructure disruption events in each detected timeframe.

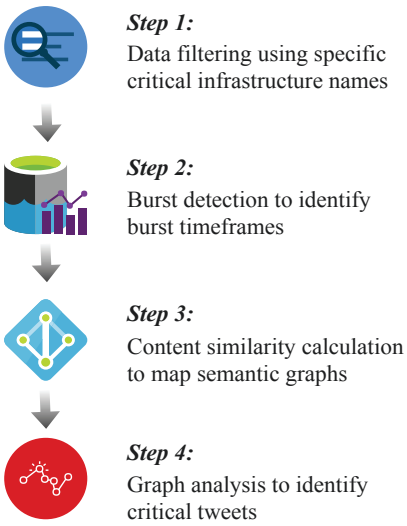

Figure 1. Framework of graph-based approach

3.1. Data filtering

To construct datasets specific to certain critical infrastructure, this paper filtered out the tweets using specific names of critical infrastructure as well as their abbreviations such as "Interstate highway 10 (I-10)", "George Bush Airport (IAH)" and "Barker reservoir (Barker)". This strict restriction of data filtering contributes to the recognition and tagging of specific infrastructure and locations. Therefore, the results of the following processes can be targeted to specific infrastructure and the findings can be used by residents and responders directly. To lessen the effects of uninformative words and characters, we remove the stopwords, punctuations, URL, and emoticons.

3.2. Burst detection of timeframes

A burst of frequencies of tweets indicates a change of situation or an impact of disruption on humans. The frequency means the number of relevant tweets posted in one hour. Disruptions of critical infrastructure often happen in a very short period and lead to human activities on social media, such as reporting damages and complain its adverse effects [28]. Hence, the burst of tweets related to critical infrastructure can indicate the changes of situation and the impacts of such changes on society. During that period, the frequencies of relevant tweets are much more significant than usual, then reach a peak when the situational information is disseminated across the social media platform, and then decrease with the dissipation of adverse effects of this damage. High-impact failures of critical infrastructure and changes of related situation would lead to differences in frequencies of associated tweets over time. The phenomenon is illustrated in Figure 2. Thus, an important step is to identify the burst timeframes based on the trend of tweets' frequencies in the context of disasters.

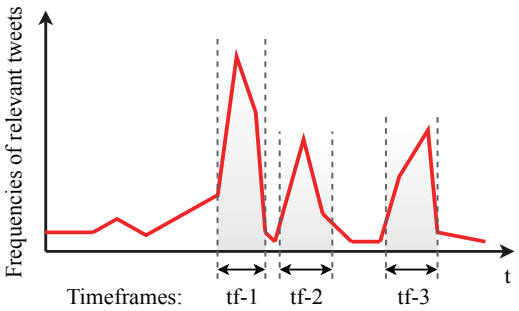

Figure 2. Illustration of burst timeframes

Based on the understanding of bursts related to critical infrastructure on social media, the following equations were derived to identify the burst timeframes (e.g., tf-1, tf-2 and tf-3 in Figure 2). First, the frequencies of relevant tweets stay within a relatively consistent range. The relevant tweets frequency can be computed as:

$$freq(t,S) = \sum_{w \in W} count(w, S_t) \quad (1)$$

$$avg(k, t, S) = \frac{\sum_{j=1}^{t-1} freq(j, S)}{k} \quad (2)$$

where S is a data stream between $t-1$ and t, t is the time slice t in hours, w is a relevant tweet, W is a set of relevant tweets, and k is the number of past time slices. This approach is an improved extension of an existing technique which was developed to identify key moments on social media streams from a set of keywords [29]. The average frequency of relevant tweets is the stable frequency before bursts. The bursts can be identified as:

$$\frac{freq(t, S)}{avg(k, t, S)} > \delta \quad (3)$$

where δ is the threshold of burst frequency in social media streams. The value of δ typically ranges from 2 to 15 and is determined by the requirement of precision [29]. The larger the value, the more distinguishing the burst time-frame. There could be cases that a certain infrastructure component or facility may be rarely or never noticed and discussed on social media, until it is impacted by extreme events causing major disruptions to services. However, the perceived impacts of damages (service disruptions) on users is as important as the damages to components for disaster response and recovery. For example, the information obtained from social sensing may not inform about damages to electricity transformer substations. But, it will help detecting electricity outages as a result or impacts on residents (e.g., inability to cook). Hence, the tweets collected and analyzed by the proposed approach enable capturing the impacts of service disruptions, as well as damages to infrastructure components and facilities. In addition, this approach can identify the discrete time when there is a burst frequency. First, the continuous points are grouped together to form a time-frame. Then, the tweets posted in each time frame are filtered out in each subset.

3.3. Content similarity and graph mapping

Similar content between tweets posted by different users indicate the similarity of situational information that these tweets deliver and the potential relation between them, which can be computed as [30]:

$$cos\theta_{ij} = \frac{\overrightarrow{v_i} \cdot \overrightarrow{v_j}}{\|\overrightarrow{v_i}\| \|\overrightarrow{v_j}\|} \quad (4)$$

where $\overrightarrow{v_i}$ and $\overrightarrow{v_j}$ are the vectors of two different tweets. There are multiple matured methods to convert the list of tokens into vectors in Python for being better processed on computers. This paper employs one of the most common method, tf-idf (i.e., term frequency-inverse document frequency) to obtain numeric matrices of tweets. In this study, each tweet is considered as a document, and a tweet is represented as the frequencies of tokens. Once vectors of tweets are obtained, we can calculate the content similarity between these tweets using Eq. (4). The results vary from 0 to 1, showing the extent of content similarity between two tweets. The higher the value, the greater the content similarity.

Edges in graphs represent the relationships between different nodes. Here, tweets can be considered as nodes and their content similarity can be represented by edges in graphs. The undirected graph G is defined as follows:

$$G = (V(G), E(G), w) \quad (5)$$

where w is the weights of edges. The graph is composed by nodes $v \in V(G)$ and edges $e \in E(G)$. The weights of edges are defined based on the content similarity between two tweets:

$$w(e_k) = [10 \times cos\theta_{ij} + 0.5],\ e_k = (v_i, v_j) \quad 0.2 < cos\theta_{ij} < 0.9 \quad (6)$$

As shown in Eq. (6), an edge cannot be established unless the content similarity is greater than 0.2 and smaller than 0.9. Retweeted tweets, the similarity of which is greater than 0.9, cannot be an evidence of the credibility of the content in such tweets. Nevertheless, those tweets can be the supplementary materials to show the importance of the information delivered by these tweets. Thus, in the proposed approach, the weight of the edges between retweeted tweets is set as 2. Meanwhile, two tweets with an edge weight which is less than 2 are considered to be irrelevant. Thus, we ignore the relations between these tweets and does not establish the edges between them. As shown in Eq. (6), the weights of the edges are integers obtained based on integer conversion. Based on the identified nodes, edges and computed weights, the semantic graph can be mapped in each timeframe. Such representation provides a simplified model of complex relationships among tweets, where network core could contain complete and reliable situation information about certain critical infrastructure.

3.4. Graph analysis with weighted degrees

Critical tweets with the most similarity to other tweets in the same timeframe can contain situation information related to certain critical infrastructure

disruptions. The degree in the semantic graph represents the similarity of a tweet to other tweets. Thus, the tweets with the highest degree can be considered as *critical tweets*. In our weighted graphs, the weights of edges vary depending on the extent of similarity of a tweet to other tweets. Hence, the weighted degree considering the similarity of tweets can be obtained as:

$$deg_{v_i} = \sum_{e \epsilon v_i \times V} \big(w(e)\big)^{\alpha} \quad (7)$$

where $v_i \in V(G)$ are the nodes, and α is a parameter that contributes to reducing the impacts of a large number of low weight degrees. The value of α should be greater than 1 and be determined based on the required precision. The larger the α, the more important the close similarity between tweets for identifying critical tweets. The definition of α indicates the importance of close similarity of a tweet to other tweets. Hence, the identified tweets with close similarity are more representative than the tweets with a number of low similarities. It should be mentioned that there may be some discrete components in a semantic graph. It would be possible that a critical tweet is identified from a small component (rather than a giant component of the network) when a topic is the focus of only a small group of people. To overcome this challenge, we developed an equation to identify the giant components in the graph before identifying critical tweets:

$$g_c = Max\big(\sum_{v_i \in V(H)} deg_{v_i}\big) \quad (8)$$

where g_c represents the giant component in a graph, and $H \subset G$, which is a component of graph G. As shown in Eq. (8), the giant component is obtained from the sum of weighted degree in a component rather than the size of a component. The weighted degree is the representation of content similarity which emphasizes the criticality and credibility of a tweet. Hence, the giant component in the tweets graph is a collection of tweets with critical and credible information. In the giant component, the critical tweet can be obtained as:

$$v_c = Max\big(deg_{v_i}\big),\ v_i \epsilon V(g_c) \quad (9)$$

Based on the definition of critical tweets discussed earlier, a critical tweet is the node with the highest weighted degree in the giant component, which can be identified by Eq. (9). There could be a case that more than one tweets have the highest weighted degree because of retweets. Our algorithm can detect all the tweets with the highest weighted degrees and convert them to their original tweets. The results of this algorithm are informative and effective for understanding the disruption and situation of critical infrastructure in disasters.

4. Case study of Hurricane Harvey

To illustrate the capabilities of this graph-based approach, we conducted a case study of Hurricane Harvey in Houston. In this paper, the investigated critical infrastructure includes Barker and Addicks reservoirs.

4.1. Context and data filtering

Hurricane Harvey, a category four tropical storm landed in Texas from August 25th to August 29th. Harvey caused significant infrastructure disruptions. For example, more than 200 road sections were closed due to flooding, all flights were suspended at Houston Intercontinental Airport System, and the water level in Addicks and Barker reservoirs reached their maximum capacity, which led to water release from the reservoirs [31].

We collected all tweets which are around 21 million, over Houston area from August 22nd to September 30th. This dataset includes the tweets that posted by the users whose profiles have a location of Houston, or the tweets that are geotagged in the bounding box of Houston. People whose profiles mentioned Houston might have been in other places. They may or may not comment the state of infrastructure (e.g., what they learned from news pages or word of mouth). The collection of tweets with localized information about infrastructure is a subset of our total dataset. Hence, the requirement for data collection proposed in this study would need acquiring a complete dataset of tweets with critical situational information and improve the accuracy of our findings.

To demonstrate the application of the proposed graph-based approach, we investigated the disruptions in the reservoirs (i.e., Barker and Addicks) in West Houston, filtered out 14728 tweets related to these reservoirs from August 22nd to September 4th in our entire dataset, and mapped the trend of frequencies for the relevant tweets (see Figure 3). In this case, the reservoirs were safe. However, because of the heavy rainfall from Hurricane Harvey, the reservoirs were full, a large amount of water was released, and subsequently the nearby neighborhoods were flooded.

As shown in Figure 3, there are some distinguishing crests and bottoms. The labels are designated as "year_month_day_time (in hour)". For example, there is a burst at 9 PM on August 26th when Hurricane landed in Texas. After that, bursts of relevant tweets at 2 AM on August 28th, at 1 PM on August 29, at 1 PM on August 31 and at 10 PM on September 2 appeared

because of the changes in the situation of the reservoirs and their adverse impacts on the residents. The following sections will examine how the situation changed and what were the specific events related to the reservoirs according to the graph-based approach.

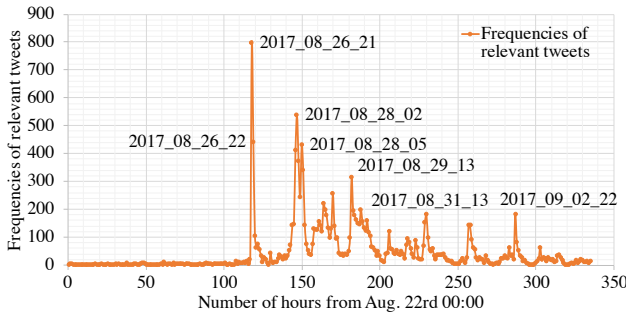

Figure 3. Frequencies of relevant tweets during disasters

4.2. Results

Sixteen burst time-frames were detected by applying the graph-based approach to the filtered dataset (see Figure 4). The spans of the time-frames varied because of the duration of the disruption and the impacts on people. The comparison of time-frames with Figure 3 shows that the detected timeframes were consistent with the crests of the frequencies of relevant tweets. Then, the tweets posted in the detected timeframes were grouped into sixteen clusters. The content similarities between the tweets in the same clusters were implemented for the establishment of the graph edges. Based on their content similarities, the semantic graph in each time frame was mapped to exhibit the correlations between tweets. Figure 5 shows the representation of nodes and edges for each graph. As shown in Figure 5, the numbers of nodes and edges in the first four graphs were much greater than the other graphs. The corresponding time-frames for these four graphs were from August 26th to September 30th. After Harvey passed, the weather went better, and the flooding receded. The water levels in the reservoirs were dropping down. Thus, fewer people talked about the situation of the reservoirs. The number of nodes and edges in the other graphs were expectedly less than the nodes and edges in the first four graphs.

The next step includes examining the weighted degrees of tweets. Figure 6 shows the graphs in two burst timeframes: No.5 and No.15, as examples of semantic similarity among tweets. The weighted degrees vary from 0 to 750, which are distinguished by the intensity of the color and size of the nodes in both graphs. Distinctly, in Figure 6(a), there should be one tweet having the highest similarity to other tweets. In Figure 6(b), multiple tweets have the highest similarity to other tweets. Our approach can identify all such tweets and investigate if their close-similarity was caused by retweeting. Meanwhile, a number of tweets with low similarity to other tweets can be seen in the graphs. These low-similarity tweets were associated with the situation of reservoirs, but they may not deliver complete and informative messages related to the high-impact disruptive events.

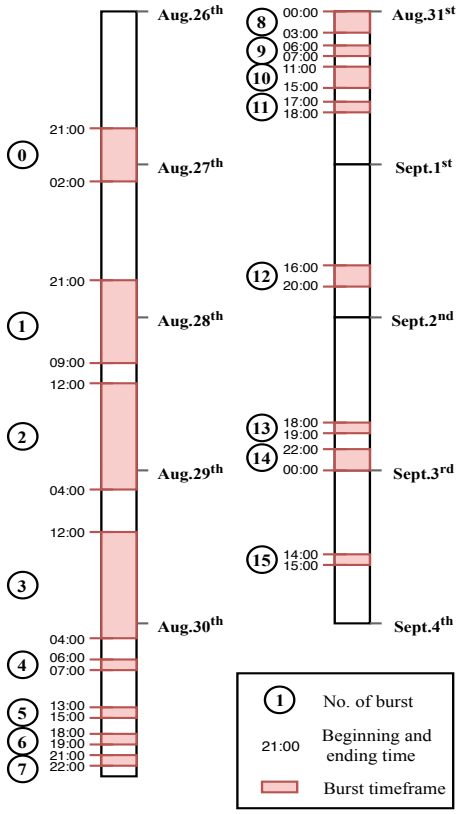

Figure 4. Identified burst timeframes

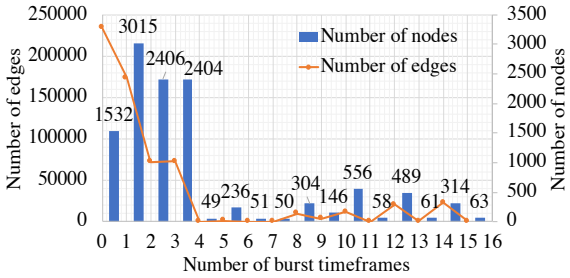

Figure 5. Graph information in burst timeframes

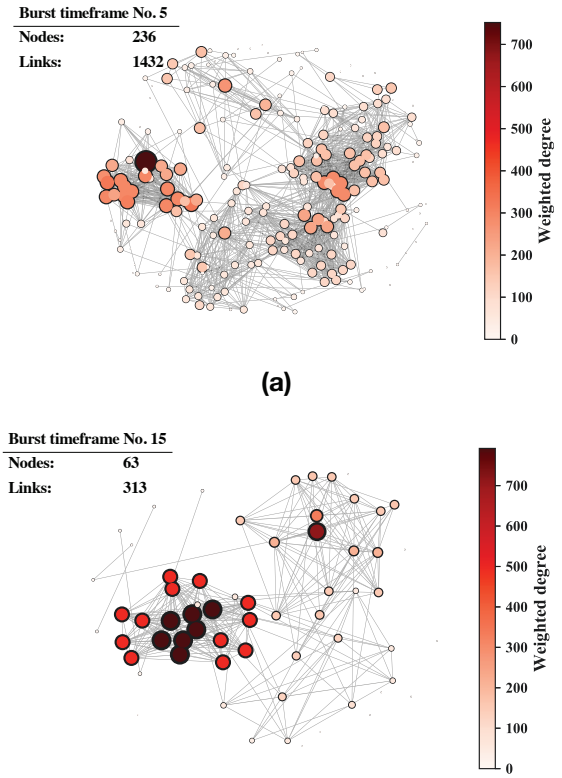

Figure 6. Semantic graphs in two burst timeframes: (a) burst timeframe No.5; and (b) burst timeframe No.15

The critical tweets in each burst time-frame were identified based on their weighted degrees by using Eq. (9). Table 1 displays all original critical tweets that users posted. As shown in these tweets, the situation of the reservoirs can be fully interpreted and understood by humans because it is presented completely and logically. According to the identified critical tweets, we deduced the temporal unfolding of events related to Barker and Addicks reservoirs. At the beginning of Hurricane Harvey, there was an early warning on social media, which indicated that an extreme weather condition would occur affecting the reservoirs (i.e., No.0 in Table 1.). Then, when Harvey landed in Houston, the reservoirs had to release water to eliminate possible breach (dam safety) (i.e., No.1 in Table 1.). These events led to flooding in nearby neighborhoods (in the downstream of Addicks and Barker reservoirs) and affected other critical infrastructure such as Interstate Highway 10 (i.e., No.6 in Table 1.), and Buffalo Bayou (i.e., No.7 in Table 1.). Emerging information spreaders (such as Twitter user Jeff Linder) updated the conditions of reservoirs and eased residents' concerns about the danger of dam spills (i.e., No.8 in Table 1.). Meanwhile, to reduce the adverse impacts of the water release, response actions (e.g., mandatory evacuation) were taken by stakeholders (e.g., Harris County Flood Control) (i.e., No.9, No.12 and No.14 in Table 1.). Finally, when Hurricane Harvey ended, and flooding receded, water levels in the reservoirs declined and water release stopped (i.e., No.15 in Table 1.).

The information shared via the critical tweets guided residents to adequately respond to the adverse effects of critical infrastructure disruptions, helped voluntary organizations to provide relief efforts to worst-hit areas, and supported responders to distribute relief resources and personnel. Thus, the critical tweets provided reliable and useful situational information for community members and first responders to better understand the disruptions of critical infrastructure during the occurrence of disasters.

Table 1. Identified critical tweets in timeframes

No.	Identified critical tweets
0	'RT @abc13houston: A possible tornado was reported in northwest Harris County on Barker Cypress '
1	'Harris County Flood Control says the release of water from the Addicks and Barker Reservoirs has started. khou11 '
2	'KATY - Many shelters have reached capacity, but you can go to the Berry Center, 8877 Barker Cypress and they can get you to an open shelter.'
3	'Addicks and Barker dams have NOT breached! '
4	'UPDATE on Addicks and Barker Reservoirs. @KPRC2 '
5	'RT @JeffLindner1: News conference at 800am on watershed updates and Addicks/Barker and Buffalo Bayou #houwx #hounews'
6	'Crews are getting creative along I-10 to dam up rising water from the Addicks and Barker Reservoir releases. '
7	'NEW: Buffalo Bayou continues to rise because of releases from Addicks reservoir, says @JeffLindner1 from Harris Co. Flood Control district'
8	'RT @JeffLindner1: Addicks and Barker Dams are completely safe and are NOT in danger of failing #houwx #hounews #txwx'
9	'New Mandatory Evacuation for Barker Reservoir Area Communities - '
10	'There are no problems with Addicks and Barker Reservoirs, the dams are safe and operating as designed #houwx #hounews #txwx'

11	'NOW: @JeffLindner1 with @hcfcd joins us for updates on Barker & Addicks reservoirs #Harvey #houstonflooding #hounews'
12	'Mayor asking residents w/ water in homes to leave if S of I-10, W of Gessner, N of Briarforest, E of reservoirs. '
13	',ö†Ô∏ÈALERT: Addicks Dam will NOT break!! '
14	'Ordering MANDATORY evacuation of West Houston dwellings already flooded by release of water from Addicks and Barker-Cypress reservoirs.'
15	'RT @JeffLindner1: Barker is currently at 99.14 ft and falling #houwx #hounews '

4.3. Validation

To demonstrate the reliability of our graph-based approach and the credibility of the identified situational information, we plotted the weighted degree distribution on log-log scales for semantic graphs and referred to some new articles to validate the content of critical tweets.

Figure 7 shows the weighted degree distribution of two semantic graphs in burst timeframes No.5 and No.15. As shown in Figure 7(a), only one node has the highest degree, which indicates there is only one tweet has the most complete information based on the definition of the weighted degree in our approach. The weighted degrees of tweets varied among the nodes in the graph. In Figure 7(b), however, there are multiple tweets with the highest weighted degree, which may indicate that the critical tweets with most complete situational information were retweeted several times in that timeframe. Our approach can identify all those critical tweets. Like Figure 7(a), the weighted degrees are various from 0 to 750, which are consistent with the semantic graphs in Figure 6.

To validate the content of detected critical tweets, we examined the critical tweets comparing to recorded information from news articles. For example, "*It took Harris County officials until late Sunday, Aug. 27, to begin issuing similar warnings for communities upstream of both Barker and Addicks.*" [32] This is consistent with critical tweet No.1 in Table 1. Also, as shown in Figure 8, Interstate Highway 10 Barker and Addicks reservoirs, and water was released from the reservoirs to Buffalo Bayou. This information can validate the physical interdependencies between the reservoirs, I-10 and Buffalo Bayou and provides evidence regarding the reliability of the information in critical tweets No.6 and No.7. Another news article stated, "*Late Sunday night, local officials issued voluntary evacuation notices for residents around the reservoirs*" [33] This information can validate that the mandatory evacuation identified in critical tweet No.9. "*HCFCD officials said the reservoir levels peaked on Aug. 30 at 109 feet at Addicks and 101.5 at Barker. With the releases, the reservoir levels are dropping.*" [34] The information that reported in this article can demonstrate the event of declining water level identified in the critical tweet No.15 and No.11. Another news article informed that "*Water levels in the two reservoirs had already reached record levels Monday evening, measuring 105 feet at Addicks and 99 feet at Barker*" [33]. The water level reported in critical tweet No.15 is consistent with the statistics in the above news post.

Based on the comparison between the news articles and the identified critical tweets, the situational information in the critical tweets related to Barker and Addicks reservoirs were verified. Also, no other important situational information related to the water release from reservoirs is found missing in identified critical tweets, and the timing of the burst events are corresponding to the information posted in news articles. Hence, the results and findings obtained from the graph-based approach are credible and reliable for community members and first responders to better understand the situation in disasters.

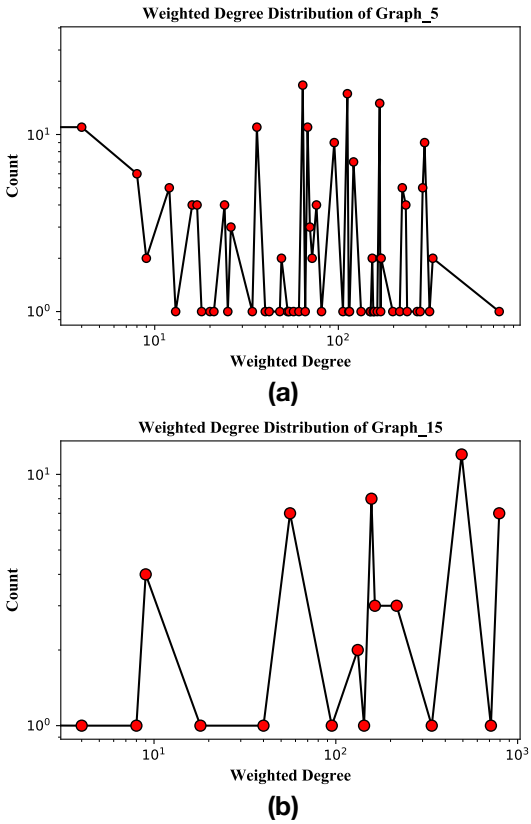

Figure 7. Weighted degree distribution on log-log scales in two timeframes: (a) burst timeframe No.5; and (b) burst timeframe No.15

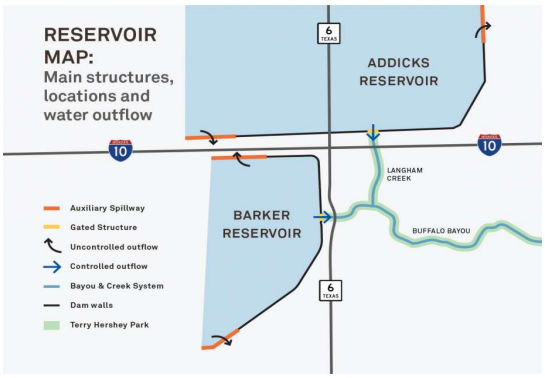

Figure 8. Reservoir map: main structures, locations and water outflow [34]

5. Concluding remarks

This paper proposed a graph-based approach to detect the critical infrastructure disruptions from social media data in disasters. The application of the proposed approach was demonstrated in a case study of Hurricane Harvey in Houston. The results showed that the primary disruptive events related to the flood control reservoirs and their adverse impacts on communities could be detected from the critical tweets. The findings also showed that the situational information related to the critical infrastructure on social media provided reliable insights regarding the spatial and temporal unfolding of infrastructure disruptions.

From the methodological perspective, the proposed graph-based approach provided useful insights into complete and non-ambiguous event detection. All the identified critical tweets contained detailed information about the situation of critical infrastructure. Hence, the results are more reliable than groups of tokens based on word frequencies in existing studies (e.g., topic modeling and clustering). In addition, this approach distinguishes the credibility of online information. For example, retweets do not play an important role in detecting disruptive events. Rumors with sensitive words (e.g., "reservoirs will spill", which will affect large-scale of residents) would also be retweeted many times. To overcome this challenge, the proposed graph-based approach put more weight on the content similarity between the tweets posted by multiple users. Hence, this method can improve the credibility of the critical tweets as well as the reliability of infrastructure disruption findings.

From a practical point of view, the findings of this study identified the occurrence of the disruption, the location of the critical infrastructure, how severe the events are, what agents are involved, and what response actions are taken for a response. This spatial and temporal information can be integrated into smart city digital twins to examine the disruptions in infrastructure services using social sensing. The actors in cities can achieve better information sharing and communication by using such situational information, and make better decisions such as locating shelters, distributing resources, and mobilizing relief crews. In addition, this approach enables effective evaluation of infrastructure performance in the aftermath of disasters to inform future hazard mitigation planning and infrastructure prioritization. As such, the cities can become smarter and more resilient, relying on human-sensor data-driven approaches (i.e., social sensing) for detecting infrastructure disruptions and corresponding societal impacts.

This approach can be further extended in the following areas: (1) improving the algorithm of burst detection to achieve burst prediction in a near real-time manner; (2) enhancing the algorithm of vectorizing in tweets to make it more feasible in disaster and critical infrastructure domain; and (3) developing algorithms to identify the relevant entities in critical tweets and exploring their relationships for better disaster planning and response.

6. Acknowledgments

This material is based in part upon work supported by the National Science Foundation under Grant Number IIS-1759537 and the 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 the National Science Foundation and Amazon Web Services (AWS).

7. References

[1] J. Zhu and A. Mostafavi, "Enhancing Resilience in Disaster Response: A Meta-Network Analysis Approach," *Constr. Res. Congr. 2018*, no. 1, pp. 2250–2259, 2018.

[2] F. Kadri, B. Birregah, and E. Châtelet, "The impact of natural disasters on critical infrastructures: A domino effect-based study," *J. Homel. Secur. Emerg. Manag.*, vol. 11, no. 2, pp. 217–241, 2014.

[3] A. Mostafavi, "A System-of-Systems Approach for Integrated Resilience Assessment in Highway Transportation Infrastructure Investment," *Infrastructures*, vol. 2, no. 4, p. 22, 2017.

[4] N. Mohammadi and J. E. Taylor, "Smart City Digital Twins," in IEEE Symposium Series on Computational Intelligence (SSCI), 2017, pp. 1–5.

[5] C. Fan, C. Zhang, and A. Mostafavi, "Meta-Network Framework for Analyzing Disaster Management System-of-Systems," in *2018 13th Annual Conference on System of Systems Engineering (SoSE)*, 2018, pp. 372–378.

[6] B. Jongman, J. Wagemaker, B. Romero, and E. de Perez, "Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals," *ISPRS Int. J. Geo-Information*, vol. 4, no. 4, pp. 2246–2266, 2015.
[7] G. Cervone, E. Sava, Q. Huang, E. Schnebele, J. Harrison, and N. Waters, "Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study," *Int. J. Remote Sens.*, vol. 37, no. 1, pp. 100–124, 2016.
[8] C. Fan, A. Mostafavi, A. Gupta, and C. Zhang, "A System Analytics Framework for Detecting Infrastructure-Related Topics in Disasters Using Social Sensing," in *Advanced Computing Strategies for Engineering*, 2018, pp. 74–91.
[9] Z. Li, C. Wang, C. T. Emrich, and D. Guo, "A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods," *Cartogr. Geogr. Inf. Sci.*, vol. 45, no. 2, pp. 97–110, 2018.
[10] W. Cui *et al.*, "An algorithm for event detection based on social media data," *Neurocomputing*, vol. 254, pp. 1339–1351, 2017.
[11] C. K. Scollon, "Social Sensing for Urban Crisis Management: The Case of Singapore Haze Social Sensing for Urban Crisis Management: The Case of Singapore Haze," pp. 478–491, 2013.
[12] X. Lu and C. Brelsford, "Network structure and community evolution on Twitter: Human behavior change in response to the 2011 Japanese earthquake and tsunami," *Sci. Rep.*, vol. 4, pp. 1–11, 2014.
[13] L. Smith, Q. Liang, P. James, and W. Lin, "Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework," *J. Flood Risk Manag.*, vol. 10, no. 3, pp. 370–380, 2017.
[14] J. Fohringer, D. Dransch, H. Kreibich, and K. Schröter, "Social media as an information source for rapid flood inundation mapping," *Nat. Hazards Earth Syst. Sci.*, vol. 15, no. 12, pp. 2725–2738, 2015.
[15] H. Lee, J. Kihm, J. Choo, J. Stasko, and H. Park, "iVisClustering: An Interactive Visual Document Clustering via Topic Modeling," *Comput. Graph. Forum*, vol. 31, no. 3pt3, pp. 1155–1164, 2012.
[16] D. Alvarez-Melis and M. Saveski, "Topic Modeling in Twitter: Aggregating Tweets by Conversations," *$Icwsm16*, no. Icwsm, pp. 519–522, 2016.
[17] S. Misra, R. Goswami, T. Mondal, and R. Jana, "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.*, vol. 22, no. February, pp. 281–296, 2017.
[18] J. Weng, Y. Yao, E. Leonardi, and F. Lee, "Event Detection in Twitter," *Development*, pp. 401–408, 2011.
[19] D. Pohl, A. Bouchachia, and H. Hellwagner, "Automatic identification of crisis-related sub-events using clustering," *Proc. - 2012 11th Int. Conf. Mach. Learn. Appl. ICMLA 2012*, vol. 2, pp. 333–338, 2012.
[20] A. Ritter, Mausam, O. Etzioni, and S. Clark, "Open domain event extraction from twitter," *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '12*, p. 1104, 2012.
[21] D. T. Nguyen and J. E. Jung, "Real-time event detection for online behavioral analysis of big social data," *Futur. Gener. Comput. Syst.*, vol. 66, pp. 137–145, 2017.
[22] W. Hu, H. Wang, C. Peng, H. Liang, and B. Du, "An event detection method for social networks based on link prediction," *Inf. Syst.*, vol. 71, pp. 16–26, 2017.
[23] Q. Zhao and P. Mitra, "Event Detection and Visualization for Social Text Streams," *Event London*, pp. 26–28, 2007.
[24] Y. Wang and J. E. Taylor, "Urban Crisis Detection Technique: A Spatial and Data Driven Approach Based on Latent Dirichlet Allocation (LDA) Topic Modeling," in *Construction Research Congress 2018*, 2018, pp. 428–438.
[25] J. F. Rosser, D. G. Leibovici, and M. J. Jackson, "Rapid flood inundation mapping using social media, remote sensing and topographic data," *Nat. Hazards*, vol. 87, no. 1, pp. 103–120, 2017.
[26] M. Imran, P. Mitra, and C. Castillo, "Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages," *Proc. Tenth Int. Conf. Lang. Resour. Eval. (LREC 2016)*, pp. 1638–1643, 2016.
[27] Y. Kryvasheyeu *et al.*, "Rapid assessment of disaster damage using social media activity," *Sci. Adv.*, vol. 2, no. 3, pp. e1500779–e1500779, 2016.
[28] A. Weiler, M. Grossniklaus, and M. H. Scholl, "Situation monitoring of urban areas using social media data streams," *Inf. Syst.*, vol. 57, pp. 129–141, 2016.
[29] C. Buntain, J. Lin, and J. Golbeck, "Learning to Discover Key Moments in Social Media Streams," 2015.
[30] X. Zhou and L. Chen, "Event detection over twitter social media streams," *VLDB J.*, vol. 23, no. 3, pp. 381–400, 2014.
[31] "How Badly Has Hurricane Harvey Damaged Texas Infrastructure? | 2017-08-28 | ENR," *Engineering News-Record*, 2017. [Online]. Available: https://www.enr.com/articles/42639-how-badly-has-hurricane-harvey-damaged-texas-infrastructure. [Accessed: 12-Feb-2018].
[32] L. Olsen, "Record reservoir flooding was predicted even before Harvey hit Houston - Houston Chronicle," *Chron*, 2018. [Online]. Available: https://www.houstonchronicle.com/news/houston-texas/houston/article/barker-addicks-dams-flooding-predicted-army-corps-12632041.php. [Accessed: 05-Jun-2018].
[33] E. Wax-Thibodeaus, A. Horton, and A. B. Wang, "Houston dam spills over for the first time in history, overwhelmed by Harvey rainfall - The Washington Post," *The Washington Post*, 2017. [Online]. Available: https://www.washingtonpost.com/news/post-nation/wp/2017/08/28/houston-releases-water-from-two-dams-in-attempt-to-prevent-uncontrolled-overflow/?noredirect=on&utm_term=.706a45cf29cd. [Accessed: 05-Jun-2018].
[34] F. Tang and E. O'Neil, "How the Barker and Addicks dams work - Houston Chronicle," *Houston Chronicle*, 2017. [Online]. Available: https://www.houstonchronicle.com/local/gray-matters/article/How-the-Barker-and-Addicks-dams-work-12171719.php. [Accessed: 05-Jun-2018].