Sustainable Cities and Society 63 (2020) 102448

Contents lists available at ScienceDirect

Sustainable Cities and Society

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

Towards resilient and smart cities: A real-time urban analytical and geo-visual system for social media streaming data

Fang Yao [a], Yan Wang [b,*]

[a] Department of Urban and Regional Planning, College of Design, Construction and Planning, University of Florida, 1480 Inner Road, Gainesville, FL, 32601, USA
[b] Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, University of Florida, P.O. Box 115706, Gainesville, FL, 32611, USA

ARTICLE INFO

Keywords:
Geo-visualization
Real-time
Resilience
Smart city
Social media
Urban analytics

ABSTRACT

Cities worldwide are vulnerable to unpredictable extreme events such as disasters and public health crises. Urban big data and data-driven technologies have played an increasingly important role in building smart and resilient cities that can respond rapidly to these perturbations. However, many existing approaches had limited capabilities for processing big data, which has led to time-consuming and costly decision-making. Thus, we develop a real-time data-driven analytical and geo-visual system to enable smart and rapid responses to urban extreme events. The system is built on ArcGIS's GeoEvent Server and Apache Spark and processes streaming data from social media with high speed, massive volume, and multiple modalities. The system employs online topic modeling and domain-adaptive sentiment analysis to track small-scale, undefined events, visualizes their spatial and semantic dynamics, and provides early alerts for crises and emergencies via an interactive online GIS platform. The proposed system has been applied during a large-scale hurricane and demonstrated effectiveness and agility in tracking and reporting emerging small-scale crises. The developed system can be applied in various urban scenarios to enable timely situation awareness and rapid response. This research contributes to the smart city safety and building rapidity of resilient cities.

1. Introduction

Cities are complex and dynamic systems that contain infrastructures, information, and innovation and house the majority of the world's population (Batty, 2008). Cities are also exposed to a variety of unforeseeable extreme events, such as disasters and infectious diseases, which have at times caused tremendous economic and social losses (Arafah & Winarso, 2017; Zhu, Li, & Feng, 2019). Disasters have affected more than a third of the world's population (1.5 billion) and cost more than US$1.3 trillion in economic losses (UN DESA Population Division, 2018). In recent years, influenza epidemics have caused up to 56,000 deaths annually in the United States and have had substantial financial costs (McGowan et al., 2019). To respond actively to such events, researchers and practitioners from multiple disciplines develop theories and approaches to help cities prepare for unexpected perturbations (Woetzel et al., 2018; Zhang & Li, 2018).

In light of the need to building resilient and smart cities in this context, urban studies and practices strive to maintain cities' essential functionality while reducing the adverse effects when disruptions happen (Allam & Newman, 2018; Angelidou et al., 2018; Desouza & Flanery, 2013; Hatuka, Rosen-Zvi, Birnhack, Toch, & Zur, 2018; Leichenko, 2011; Wang, Hulse, Von Meding, Brown, & Dedenbach, 2019). Existing literature body has discussed four critical aspects of resilience: robustness (the ability to withstand stress without suffering degradation or loss of function), redundancy (the extent to which components can be substituted for to recover reduced or lost functionality), resourcefulness (the capacity to identify problems, establish priorities, and allocate resources), and rapidity (the ability to meet priorities and achieve goals promptly) (Bruneau et al., 2003; Godschalk, 2003; Zobel, 2011). However, as cities grow and gain complexity, conventional approaches that treat resilience as a conceptual process and use static data can become ineffective for achieving the conditions outlined above (Meerow, Newell, & Stults, 2016). Thus, it is necessary to implement new methods and technologies to address the challenges of extreme events and promote resilience.

In the context of burgeoning big data and advanced information and communication technologies (ICTs), more "smart" solutions have also been proposed to help cities survive and function under extreme stresses

* Corresponding author.
E-mail addresses: fang.yao@ufl.edu (F. Yao), yanw@ufl.edu (Y. Wang).

https://doi.org/10.1016/j.scs.2020.102448
Received 17 March 2020; Received in revised form 2 July 2020; Accepted 11 August 2020
Available online 28 August 2020

(Palmieri, Ficco, Pardi, & Castiglione, 2016; Soyata, Habibzadeh, Ekenna, Nussbaum, & Lozano, 2019; Yang, Su, & Chen, 2017). A recent article proposed the smart robustness, smart redundancy, smart resourcefulness, and smart rapidity to leverage resilience by embedding smart technologies and systems in the fabric of cities (Desroches & Taylor, 2018). Although rapidity (e.g., speed) in responding to risks is essential to resilient cities (Al Nuaimi, Al Neyadi, Mohamed, & Al-Jaroodi, 2015; Desouza & Flanery, 2013; Palmieri et al., 2016; Platt, Brown, & Hughes, 2016), few studies have focused on this dimension of urban resilience, particularly among urban-scale quantitative studies (Meerow et al., 2016). Within the smart city context, however, rapid or real-time big data applications can mitigate damages' impacts and enhance the capacity to recover from extreme events quickly (Desroches & Taylor, 2018; Malik, Sam, Hussain, & Abuarqoub, 2018). For instance, early detection of crises or emergencies and rapid responses allow cities to collect relevant information, monitor the characteristics of events (e.g., locations, time, types), and provide timely analyses and predictions, and thus to better coordinate relief efforts, assess damages, and restore urban system performance (Desouza & Flanery, 2013; Khan, Anjum, Soomro, & Tahir, 2015; Kitchin, 2014; Kontokosta & Malik, 2018; Woetzel et al., 2018; Zhang, Li, Li, & Fang, 2019).

However, most existing quantitative studies are conceptual rather than operational to enhance resilience with smart rapidity, because it is challenging to design a specific plan for an abstract and complex notion such as resilience (Desouza & Flanery, 2013; Hatuka et al., 2018; Wang, Taylor, & Garvin, 2020). For example, Klein, Koenig, and Schmitt (2017) described a vision and a conceptual framework for monitoring and managing cities' environmental and social dynamics without giving specific methods or plans. Current efforts to create smart and resilient cities also suffer from a mismatch between real-time information resources and delayed decision-making, as well as incompatible algorithms for processing high-volume and -velocity urban streaming data (Al Nuaimi et al., 2015; Khan et al., 2015; Yang et al., 2017). The existing prototypes of urban analytics systems (e.g., Huang, Cervone, & Zhang, 2017; Psyllidis, Bozzon, Bocconi, & Titos Bolivar, 2015) were designed for the analysis and visualization of a diversity of urban topics (human movement patterns, traffic conditions, or place of interests) using periodically updated data or a mixture of static and streaming data. These prototypes did not take full advantage of urban streaming data for smart and rapid resilient city management.

In this research, we propose a real-time urban analytical and visual system that can detect, track, analyze, and visualize small-scale, undefined extreme events clustered in content and space. The system is built on the latest versions of GeoEvent Server and Online GIS for real-time data analysis and visualization and uses geotagged streaming Twitter data. We construct several data-mining and natural language–processing modules within the system, including online topic modeling and sentiment analysis using Apache Spark. The Apache Spark distributed system is especially favorable for online big-data processing with high speed and accuracy. The system is designed for geo-textual streaming data and has the potential to be applied to various urban management scenarios. By leveraging high-volume urban streaming data and smart technologies, we hope to demonstrate the usefulness of our system to understand the dynamics of urban systems, especially during unpredicted perturbations such as disasters. The system demonstrates the analysis results with interactive maps to improve situational awareness and enhance community engagement during extreme events. The system can also be integrated into a holistic, intelligent system to play an active role in future urban planning to achieve smart and resilient cities.

2. Related work

Recently, extracting and interpreting information from streaming data has gained increasing prominence in the data mining domain. In addition, social media platforms, such as Twitter, have brought valuable user-generated behavior-rich data resources in real time, offering a growing number of opportunities to analyze the dynamics of the text streams and topics (Benhardus & Kalita, 2013; Ghani, Hamid, Hashem, & Ahmed, 2019). In urban contexts, these platforms allow people to share the events they perceive, such as nearby crises or urgent needs for specific resources (Leykin, Lahad, & Aharonson-Daniel, 2018; Yoshinaga & Kitsuregawa, 2014). Moreover, these crowdsourced social media data exist at the smallest possible measurement scale and represent the perceptions and emotions of citizens and can be used to engage citizens in two-way communication (Angelidou et al., 2018; Kitchin, 2014; Neirotti, De Marco, Cagliano, Mangano, & Scorrano, 2014; Woetzel et al., 2018).

2.1. Existing methods for real-time streaming text mining and topic derivation

Many methods have been proposed for topic detection and topic evolution over time. For example, Xie, Zhu, Jiang, Lim, and Wang (2016) proposed the TopicSketch framework for detecting bursty topics on Twitter in real time using a sketch-based topic model based on statistical data "sketches" of tweets, such as the acceleration of the number of tweets and words. Hasan, Orgun, and Schwitter (2019) developed the TwitterNews + system to detect local newsworthy events from streaming tweets. This system continually updates the most recent tweets to determine their novelty and cluster tweets into different events.

Other studies have considered both spatial and temporal features of tweets, making them more applicable to urban environments. For instance, Zhang et al. (2017) proposed TrioVecEvent, an online local event-detection method that uses geotagged tweet streams. This method generates topic clusters and selects local events by gathering information on location, time, and content to perform online clustering using a Bayesian mixture model. Yu, Li, Bhuiyan, Zhang, and Huai (2017) presented a real-time emerging-anomaly monitoring system (Ring) to detect anomalies within minutes after the events happened. This system improved on a graph-stream model and was implemented on the Spark distributed data processing system.

These streaming data-based methods were explicitly designed to achieve real-time topic derivation and event detection through the use of low-computation solutions and quickly updating the computation results. These methods were implemented to detect emerging topics or geo-textual clusters without further analyzing the detected events or enabling real-time visualization of the computation results. However, to improve the results of exploring and interpreting massive and complex streaming data, real-time visualization is crucial. Visualization approaches couple human intuition with computational analysis and thus help users understand the patterns of the events and gain insights when making decisions (Chae et al., 2014; MacEachren et al., 2010; Thom, Bosch, Koch, Worner, & Ertl, 2012).

2.2. Real-time event detection and geo-visualization systems

To address the need for real-time analysis and visualization, some systems were designed to not only process social media data streams in real time but demonstrate the results to users interactively. Early systems such as TwitterMonitor (Mathioudakis & Koudas, 2010) could detect trending topics in the Twitter stream and visualize their basic features (e.g., temporal line graphs, keywords) without mapping the topics. More subsequent research was devoted to developing geo-visualization systems or frameworks as prototypes that could be extended to automatically detect anomalies or events and to monitor human spatiotemporal activities in tweet streams in real time. Examples include Terpstra, Stronkman, De Vries, and Paradies (2012)'s Twitcident system, Huang et al. (2017)'s cloud-based framework for disaster monitoring, and Wachowicz, Arteaga, Cha, and Bourgeois (2016)'s workflow of querying space-time activities (STA) via geotagged tweet streams.

Continuing these studies, many systems with similar designs could

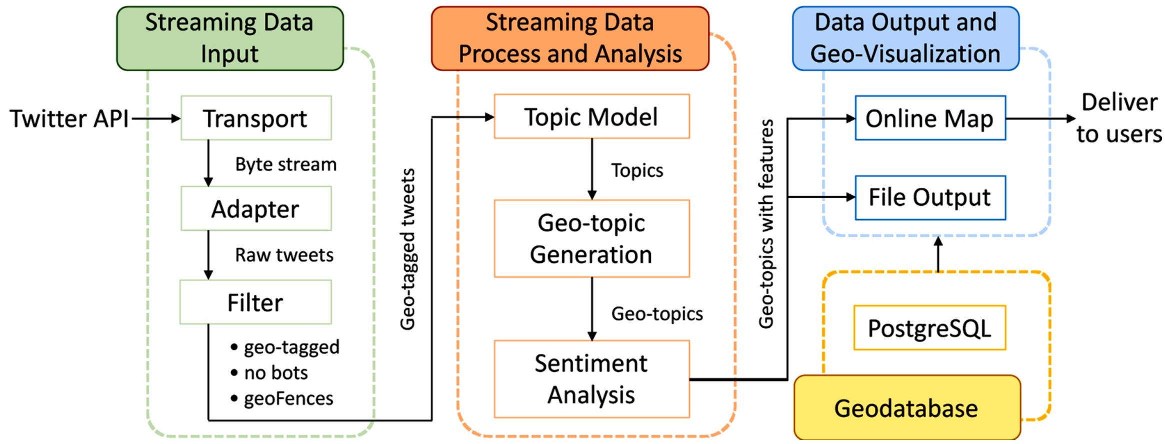


Fig. 1. The architecture and main modules of the real-time analytical and geo-visual system.

achieve real-time data processing and geo-visualization. A large number of these systems were rooted in geospatial visualization and enabled real-time geo-visual analytics for geo-text aggregation, spatial cluster exploration, and crisis discovery, such as SensePlace2 (MacEachren et al., 2010), SensePlace3 (Pezanowski, MacEachren, Savelyev, & Robinson, 2018), and ScatterBlogs (Thom et al., 2012). These methods focused on making sense of places by extracting useful information from geo-textual streams. Some studies (e.g., MacEachren et al., 2010; Middleton, Middleton, & Modafferi, 2014) parsed location information from textual data (e.g., hashtags, name entities, or user account location) and employed reversed geocoding. Although the geocoding approach increased the volume of data from the social media stream that could be crawled, it can be problematic when mapping fine-grained local events. Some other real-time geo-visual analysis systems were developed to recognize the importance of spatial scale and designed to identify small-scale or local events. For example, Boettcher and Lee (2012) proposed the Eventradar system for detecting small-scale local events using a density-based clustering algorithm with sliding time intervals.

In addition to underlining the geographic magnitude of real-time analysis and visualization, some systems also introduced content-relevant features, such as incorporating human perspectives through sentiment analysis or targeting at one type of event by setting search keywords. For example, TwitInfo (Marcus et al., 2011) can automatically identify and label spikes of tweet events and allows users to select and track events through a visualization platform. This platform also shows the positive and negative sentiments surrounding events and the aggregate sentiment of the tweets within the events. Choi and Bae (2015) introduced the Social Big Board, a real-time disaster-monitoring system that can analyze and map disaster-related tweets and their trends. This system also analyzes people's emotional information using pre-defined sentiment words in positive, negative, or neutral sentiment.

Some real-time systems have been proposed to be better for specific types of events, based on the different characteristics of extreme urban events. For instance, Avvenuti, Cresci, Marchetti, Meletti, and Tesconi (2014) developed a real-time alert and report system specifically for earthquake disasters (EARS). Smith, Liang, James, and Lin (2017) presented a real-time modeling framework to identify flooding areas and infer inundation during storm weather. Şerban, Thapen, Maginnis, Hankin, and Foot (2019) introduced a software system SENTINEL that classified health-related tweets to detect disease breaks and provide syndromic surveillance in real time. We also found that a large portion of the existing studies discuss the use of streaming data to extract crisis-relevant information for disaster management and risk control because these events are time-sensitive and require real-time decisions.

In summary, a diverse set of data-driven methods have been designed for mining and mapping information from streaming data. However, only a few studies have been able to pinpoint the locations and scales of detected events and their evolution over time and space at a fine-grained resolution. Existing real-time systems that incorporate both computational analysis and visualization to identify emerging events only show some basic features of the events (e.g., tweet volume change, representative tweets, and spatial clusters), but in-depth information extraction and undefined event detection (without using predefined keywords) are lacking. Still fewer studies have initiated early alerts or other actions after the systems detected emergencies or crises.

3. Developing a real-time urban analytical and geo-visual system

In this section, we demonstrate the design and methods of our real-time urban analytical and geo-visual system that works for streaming geo-textual data (i.e., data with both geographical and content features). This system is built upon Esri's GeoEvent Server (version 10.7). The server provides comprehensive tools and pipelines to support high-volume real-time data input, processing, and output, making it especially suitable for geospatial streaming data. We employ several tools provided by the server and create our customized data-processing tools with the server's Java SDK (Software Development Kit).

Fig. 1 demonstrates the architecture of our proposed system, with analytical modules embedded in the GeoEvent Server. Three analytical and visual modules make up the core of the system: streaming data input, streaming data processing and analysis, and data output and real-time geo-visualization. This system is supported by a large geodatabase - PostgreSQL. The input module ingests streaming data directly from a Twitter streaming API and transforms the stream of information into formatted raw tweets. This module also performs simple tweet screening: collecting geotagged non-bot tweets from within defined spatial areas. The processing and analysis module serves as the primary analytical part of the system. In this module, we develop customized functional models (configured as processors in the server) to extract information of interest from the data stream. The output and geo-visualization module is our platform for real-time mapping and interactions using Online GIS maps.

3.1. Streaming data input module

Tweets generated during extreme events contain rich geographical and content information that is useful for city management. Twitter offers an open API for collecting large amounts of voluntarily reported tweets in real time. In general, about 1% of tweets can be collected by a standard account (Wang & Taylor, 2019; Wang, Wang, & Taylor, 2017). In this module, two tools are used to process the incoming data stream into a format suitable for the GeoEvent server so that the data can be further analyzed. The transport tool connects to the Twitter API and

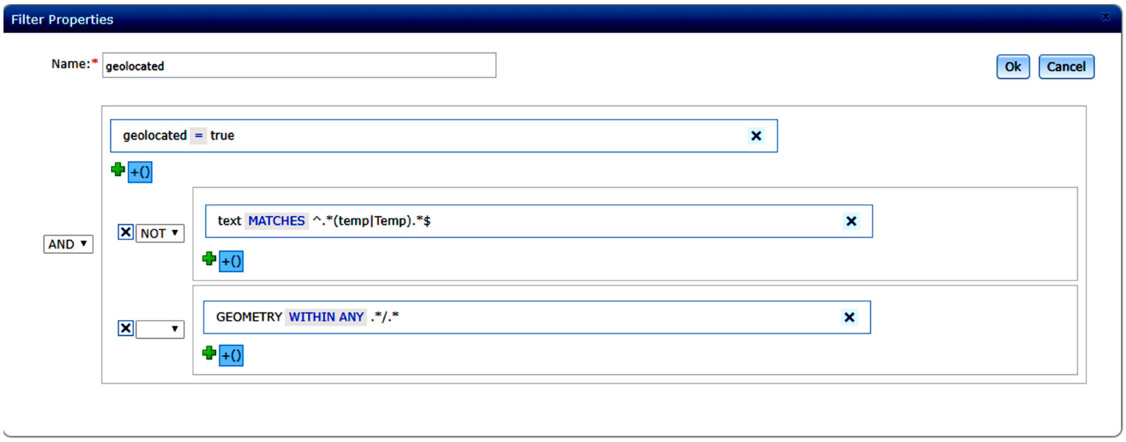


Fig. 2. The interface of filter settings.

receives data as a raw byte stream, and the adapter tool converts this stream into formatted tweet objects that can be processed by the system. Each collected tweet is configured as a geo-textual object that contains both the text of the tweet and its metadata, such as unique ID, user ID, timestamp, latitude, and longitude.

Then we set up a filter tool to remove tweets generated by robot accounts (bots), collect geotagged tweets, and define the geographic study area using GeoFences in the server. Fig. 2 shows an example of our filter settings. We refine a word list from a previous study that contains terms usually used by bots, such as "temp" and "barometer"(Yao & Wang, 2020a), and we use regular expressions to match the incoming raw tweets and exclude any that contain the words from the list. We also exclude tweets without geolocations and filter the tweets to select those from a predefined study area using GeoFences settings. GeoFences can determine the spatial relationship between a geographic boundary (polygon) and a geotagged tweet. We do not set keywords to filter tweets, as we intend to design a system for detecting *unexpected* extreme events that meet our design assumptions. We develop the filter tool that processes data in a computationally simple and fast way in order to remove unnecessary data volume for the computationally expensive process in the following steps, thus enhancing the speed and efficiency of the system.

3.2. *Streaming data process and analysis module*

This analytical module analyzes a stream of geotagged tweets and generates the outcomes for a dynamic map. We first apply the online topic model to identify the topic distribution for each tweet in the stream. Then we use a clustering method to synthesize the topics and generalize geo-topics using additional geospatial information embedded in the tweets. In this study, geo-topics are topics clustered in space that are used to represent urban events, such as urban activities, local news, and emergencies. The assumption is that if tweets are collected in a short period and contain similar words and topics, they are likely to be clustered spatially and related to a specific local event (Wang & Taylor, 2019). In the third step, we compute the sentiment score of each tweet and average the sentiment score of each geo-topic. These scores can be used as indicators of potential extreme events, such as emergencies or crises. We use the Java SDK provided by the GeoEvent Server to create customized processors corresponding to these data-analysis methods.

3.2.1. *Topic modeling of streaming text data based on Online LDA*

One major challenge for real-time data analysis is handling the sheer volume of data rapidly. Topic modeling is a data-mining method for discovering hidden semantic structures (topics) in large text documents, such as tweets. The most commonly used topic models are probabilistic ones, such as probabilistic latent semantic indexing (PLSI) (Hofmann, 1999) and latent Dirichlet allocation (LDA) (Blei, Ng, & Jordan, 2003). These models represent each document as a mixture of topics and each topic as a distribution over words. Most probabilistic models run offline and do not incorporate the temporal aspect of documents (Gao et al., 2020; Yu et al., 2017). Some methods began to exploit timestamps jointly with topic detection and topic evolution, such as Topic Over Time (TOT) (Wang & McCallum, 2006) and Dynamic Topic Model (DTM) (Blei & Lafferty, 2006). In social media environments, tweets arrive continuously and topics change dynamically. Topic models that to be used online or in real time must consider this temporal aspect and be updated to capture topic changes over time rapidly.

We employ Online LDA (Hoffman, Blei, & Bach, 2010), an online variational Bayes algorithm for LDA, to process streaming data and generate topics. Variational Bayes (VB) is a method of variational inference used in a Bayesian hierarchical model. VB is based on Bayes'

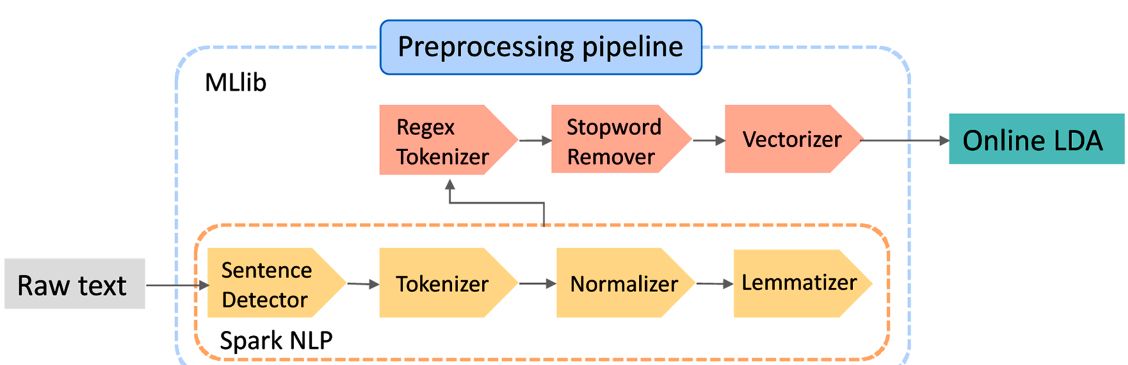


Fig. 3. The process of data cleaning and normalization.

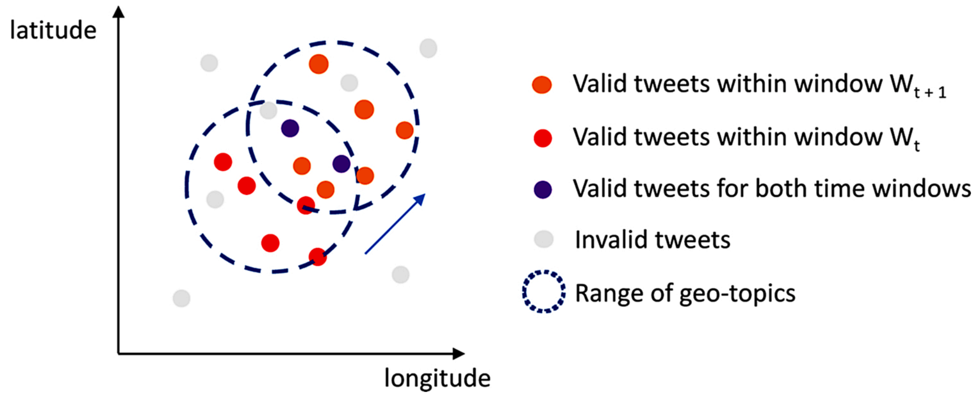


Fig. 4. Demonstration of spatial clustering patterns of geo-topics in streaming data.

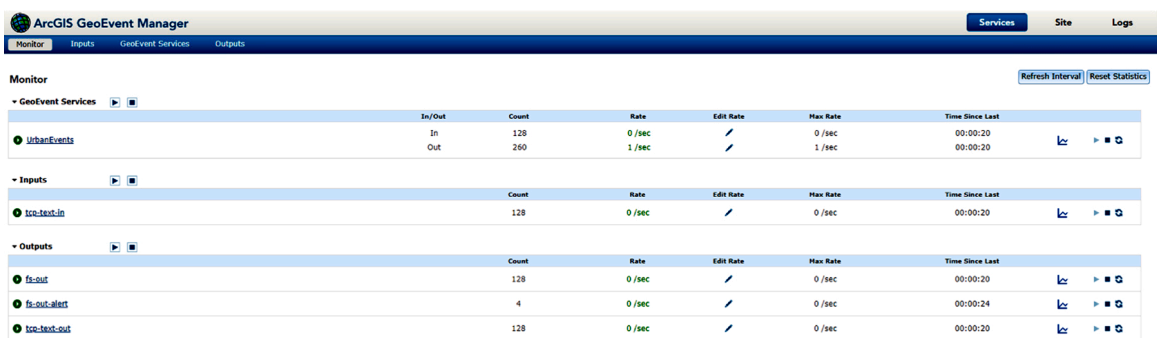


Fig. 5. The interface of GeoEvent Server for managing data and analysis.

theorem and is used to approximate the true posterior by minimizing the Kullback-Leibler divergence to the true posterior. Online LDA has advantages in handling massive collections of documents because the method allows documents to be examined as they arrive in-stream and discarded after a single look to reduce the delays. The Online LDA model is defined by some essential parameters: $\kappa \in (0.5,\ 1]$ is an exponential decay that controls the rate at which old topics are forgotten; its value range can guarantee an asymptotic convergence. $\tau_0 \geq 0$ represents a positive learning offset that slows down the early iterations of the algorithm. Minibatch t is used to consider multiple observations per update to reduce noise. We also need to set the number of topics k before running Online LDA.

We use Apache Spark's machine learning library (MLlib) in the Java programming language to implement Online LDA. Apache Spark is an open-source distributed analytics system designed for big data processing that has exceptionally high performance. We choose Java to meet GeoEvent Server's requirements so that we can create a processor with Java SDK. Before running the online topic modeling, we clean and normalize the texts of tweets using the SparkNLP library and MLlib: we remove web links, @ mentions, and stopwords, make all the words lowercase, and tokenize the texts into single terms (Fig. 3).

3.2.2. Generating geo-topics with spatial features

Because topics in the data stream change over time, they cluster dynamically across space as geo-topics, and our system is designed to monitor these spatial changes. The system uses sliding time windows with statistical metrics to analyze the spatiotemporal dynamics of geo-topics. Sliding time windows are often used in the online mode of topic modeling (e.g., Boettcher & Lee, 2012; Lau, Collier, & Baldwin, 2012). Fig. 4 shows the spatial clustering patterns of tweets for generating geo-topics based on sliding time windows. We set the window size to one hour, and when new tweets arrive, the window moves forward, so geotagged tweets received more than an hour ago are discarded. All the tweets within the window are valid, and only valid tweets are aggregated and clustered to generate geo-topics and have their corresponding features (centers and ranges) calculated. As the window slides onward, the valid tweets and features of geo-topics change. We use statistical metrics to represent the spatial centers and ranges of geo-topics. For

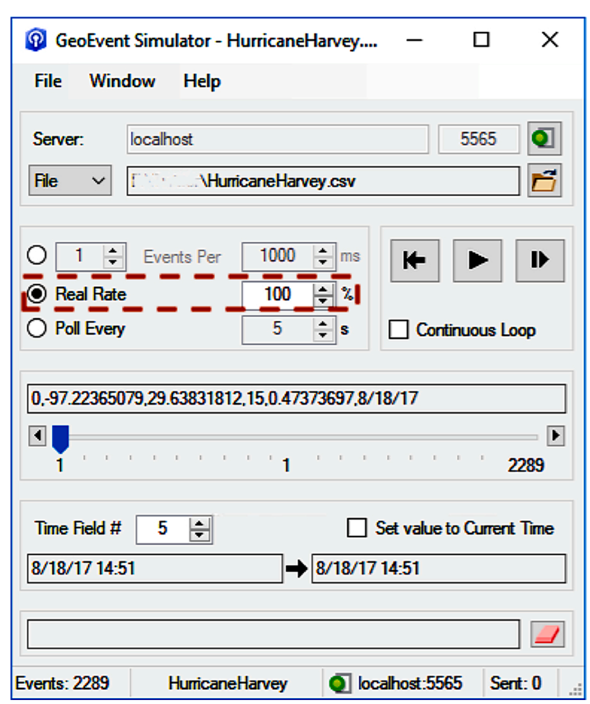
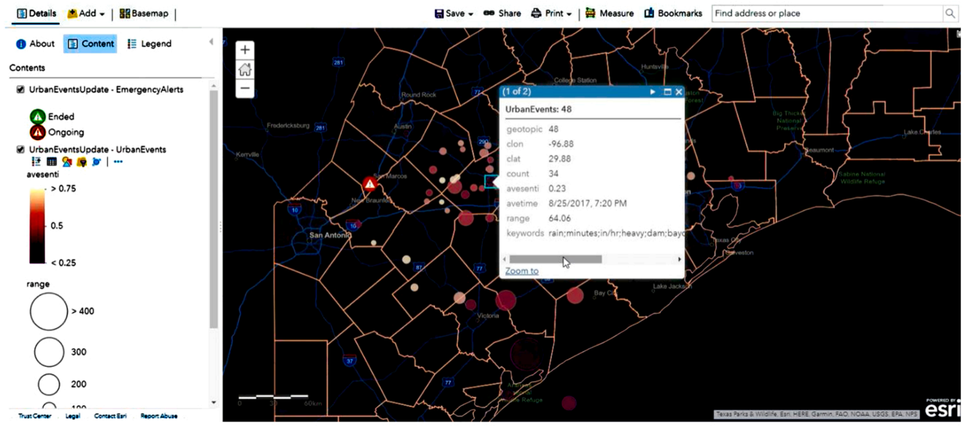
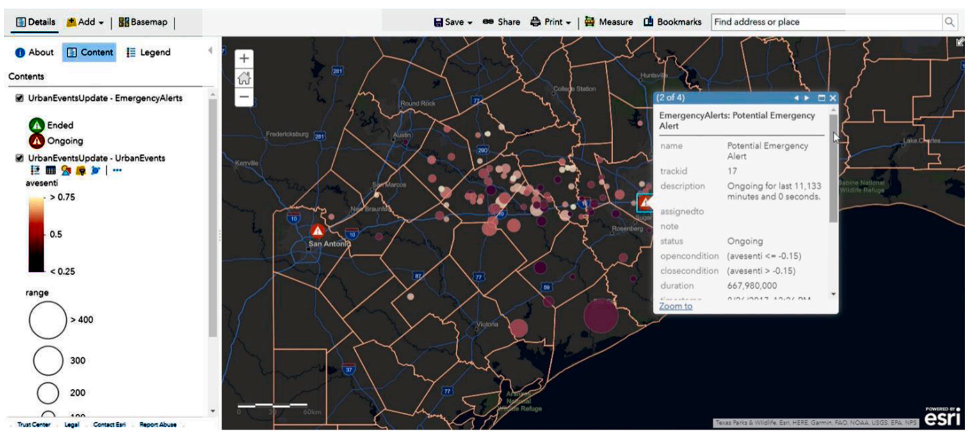


Fig. 6. The interface of GeoEvent Simulator.



(A) Urban events pop-up window



(B) Potential emergency alerts pop-up window

Fig. 7. The user interface with pop-up windows.

each geo-topic k ($k = 0, 1, 2, \ldots$), the spatial center (ϕ_c^k, λ_c^k) is represented by the arithmetic mean of the latitudes and longitudes of the valid tweets on that geo-topic (ϕ_i^k, λ_i^k), $i \in [1, n]$:

$$\phi_c^k = \frac{1}{n}\sum_{i=1}^{n}\phi_i^k, \quad \lambda_c^k = \frac{1}{n}\sum_{i=1}^{n}\lambda_i^k$$

where ϕ_c^k is the latitude and λ_c^k is the longitude of the center of geo-topic k, and n is the number of tweets belonging to geo-topic k. The geographic range of a geo-topic γ^k is the variance in the tweets' latitude or longitude converted to distance in kilometers:

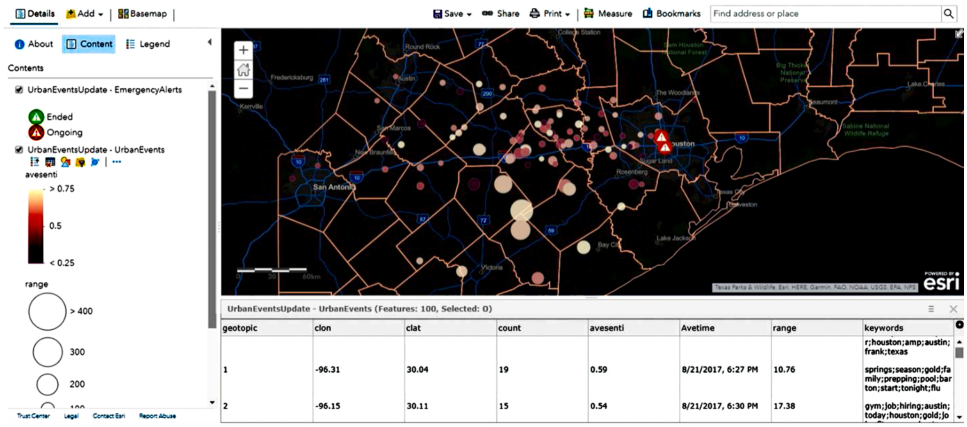
$$\gamma^k = min\left(\frac{\alpha}{n}\sum_{i=1}^{n}\left(\phi_i^k - \phi_c^k\right)^2, \frac{\beta}{n}\sum_{i=1}^{n}\left(\lambda_i^k - \lambda_c^k\right)^2\right)$$

where α and β are the coefficients for converting degrees of latitude and longitude into kilometers. Variance can measure how geotagged tweets spread out from their geographic center. If a geo-topic contains only one valid tweet, its range value is 0 (zero).

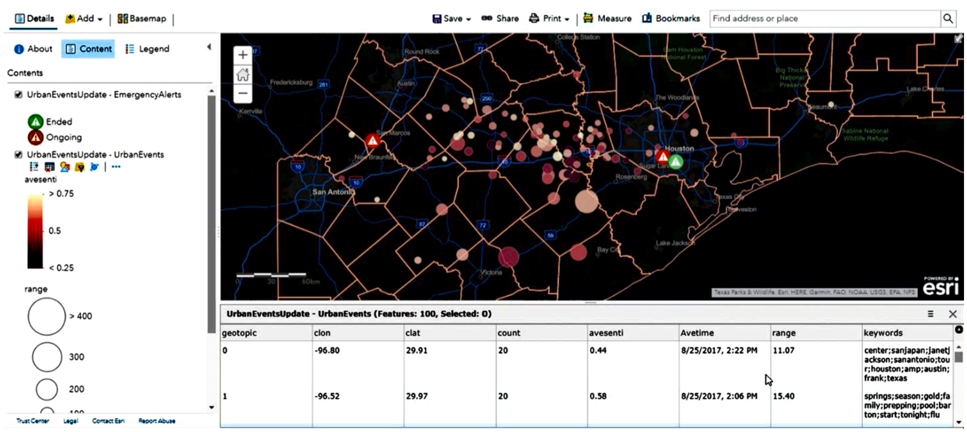
3.2.3. *Domain-specific sentiment analysis for extreme event detection*

We calculate the sentiment scores of geotagged tweets and geo-topics to identify potential extreme events that merit attention. This analytical procedure is based on the assumption that very negative sentiments are likely to indicate crises or accidents (Caragea, Squicciarini, Stehle, Neppalli, & Tapia, 2014; Lu et al., 2015). We employ a pre-trained, domain-adapted sentiment-analysis classifier to predict the sentiment score of each tweet in the data stream (Yao & Wang, 2020b). The sentiment classifier represents a domain-adversarial neural network (DANN) (Ganin et al., 2016) that is built on a recurrent neural network (RNN) with an additional domain-adversarial component. RNN is specifically useful in text mining and processing sequence data, such as streaming texts. The domain-adversarial component is appended to a standard RNN learning process in the backpropagation steps, thus the learned representations are invariant across different domains. This method can achieve high accuracy in classification and performs robustly in distinct domains (e.g., disasters, news, and lifestyles) when analyzing tweets. This domain-adaption feature makes the method exceptionally suitable for streaming tweets and changing geo-topics. The DANN method calculates the sentiment scores of individual tweets as real numbers from –2 (most negative) to 2 (most positive), with 0 as neutral sentiment. The sentiment score of each geo-topic is the arithmetic mean of sentiment values of the valid tweets belonging to the geo-topic.

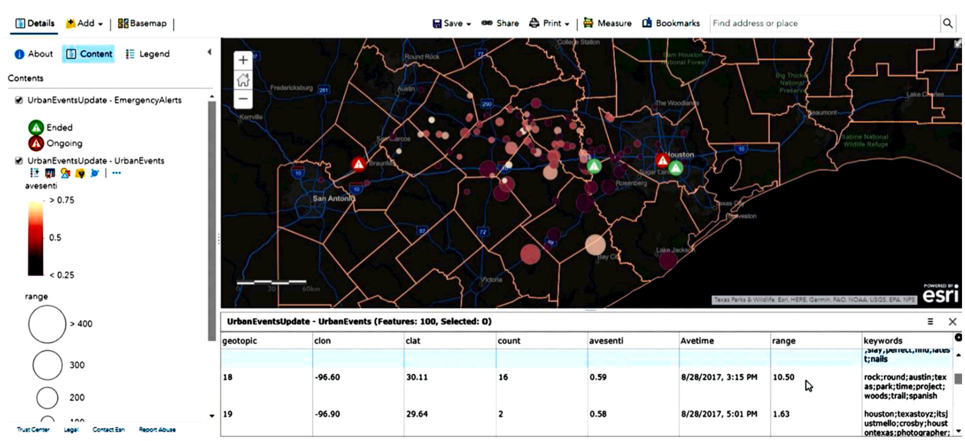
We then use an emergency-detector tool to find abnormally negative geo-topics and provide early alerts when potential emergencies are identified. This tool works by setting up a sentiment threshold, usually a negative number between -2 and 0, to trigger the alerts. If the sentiment of a geo-topic at any time is equal to or below the threshold, the system generates a potential emergency alert and indicates it dynamically on the map. After the alert is detected, its status becomes ongoing, and the tool tracks the total duration of the alert. If the sentiment increases to



(A) before the hurricane (August 21, 2017)



(B) during the hurricane (August 25, 2017)



(C) after the hurricane (August 28, 2017)

Fig. 8. Geo-visualization of sentiment on different dates.

above the threshold, the alert stops and the ongoing status ends.

3.3. Data output and real-time geo-visualization module

In the final output module, the geotagged tweets and the analytical results are stored as tables in the geodatabase. The analytical results are also visualized via a cloud-based GIS mapping platform - ArcGIS Online. These results include the geo-topics with their spatial, temporal, textual semantic, and sentiment features. The online maps can display and monitor changes in such features of detected urban events over both space and time. The maps can also show potential emergency alerts and their locations and status changes. The system employs a large-scale geospatial database (PostgreSQL) to support real-time geospatial data analysis and visualization. The PostgreSQL geodatabase stores and manages a collection of geographic datasets based on a relational database management system (RDBMS). During the real-time visualization process, the geodatabase stores maps published by ArcGIS as a feature service. This feature service contains datasets with spatial

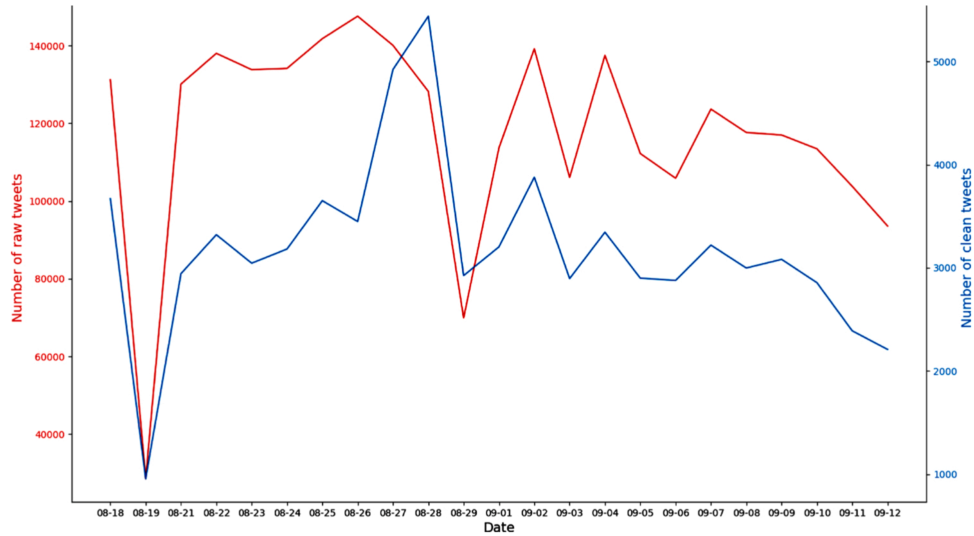


Fig. 9. Daily data volume processed by the system.

information that can be used to generate map layers. The datasets are updated simultaneously when our system is running. The feature service also contains GIS map templates with symbology that is used by Online GIS for data visualization.

4. Experiments and results

4.1. Study case and system settings

We applied the proposed visual urban analytical system in a simulated real-time scenario of Hurricane Harvey, one of the most destructive disasters to happen in the U.S. in the past decade. The hurricane caused more than a hundred billion in damage and made landfall in a densely populated area in south-central Texas. We use geotagged tweets collected by a Twitter streaming API. Our study period ran from August 18 to September 12, 2017, covering the time before (August 18–24), during (August 25–26), and after (August 27 to September 12) the disaster. Our study area was the counties spatially overlapped by the wind swaths of Hurricane Harvey. The area was set as the GeoFences polygon in our system. In general, our proposed system can provide the GeoEvent Server interface for collecting real-time streaming tweets using the adapter, transport, and filter described in Section 3.1 (Fig. 5). For this case study, we simulated the collected tweets at a real rate. The rate was computed from the tweets' timestamps using the simulator in the GeoEvent Server (Fig. 6). This simulation can precisely capture the temporal features of tweets and mirror the real-time collection of streaming data. In addition, we set the topic number to 100 for online topic modeling to cover the most possible topics. We also set the sentiment threshold to –0.15 to detect potential emergency alerts.

4.2. Geo-visualization of the analytical system

4.2.1. User interface

The video (uploaded as supplementary "Video Still") and the screenshots below (Figs. 7 and 8) show the user interface of our system and the analysis results for Hurricane Harvey. The center is the online map showing the analysis results in real time and refreshing every few seconds. The light pink polygons are county boundaries of the study area. The circles are local urban events represented by geo-topics detected by our system. The size of each circle represents the affected range of the geo-topic: the larger the radius, the greater the affected area. The colors of the circles represent the average sentiment scores of different events: the darker the color, the more negative the geo-topic is. The map also shows potential emergency alerts. The red symbol represents ongoing emergencies and the green one represents emergencies that have ended. The left side of the user interface shows the legend for the map, including the study area county boundaries, spatial ranges, sentiment scores of urban events, and potential emergency alerts. The bottom of the map is the table area, which shows detailed information on individual map layers. For example, the urban event layer updates the event number, the location of the event center with latitude and longitude, the number of tweets representing the events, the average sentiments of the events, the approximate times the events appeared, and the spatial ranges of the events. The emergency alert layer updates the ID, description, status, duration, and condition of each alert. Our system also uses pop-up windows (Fig. 7). Users can click urban events or potential emergency alerts on the map to check all the information provided in the bottom table.

4.2.2. Spatiotemporal patterns of sentiment

We examined the general spatiotemporal patterns of sentiment for detected urban events. Fig. 8 shows screenshots of urban events (geo-topics) on different dates. Before the hurricane made landfall (Fig. 8A on August 21), the overall sentiment of the geo-topics was rather positive (see lighter colors of geo-topics). The negative sentiment was mainly distributed around the metropolitan areas of Houston, Austin, and San Antonio. During the hurricane (Fig. 8B on August 25), the sentiment became more negative, shown by darker colors on the map. The spatial area of negative sentiment also expanded to surrounding areas, such as the coast. After the hurricane (Fig. 8C on August 28), the sentiment was even more negative, and the spatial area was more sprawled out because of the accumulated damage and bad weather.

We also found that the Houston metropolitan area had the highest number of potential emergency alerts during the hurricane. These included traffic delay (geo-topic #17), bayou flood flow (geo-topic #62), and heavy rain that might affect citizens' health (geo-topic #81). San Antonio had three traffic-related emergencies on three separate days (e.g., geo-topic #23). In addition, the detected potential emergency alerts also indicated the emergency locations reasonably well. For example, traffic-related emergencies were found near major roads, and bayou flood emergencies were found near Buffalo bayou. We also found that traffic delays and accidents were the most common and long-lasting emergencies during the hurricane (geo-topic #17, 23, 44). This type of emergency can affect social media for a period from half an hour to almost nine hours. Our system found 12 potential emergencies at different times and locations during the hurricane. We detected one false alarm (geo-topic #85), about fitness and training, and it lasted 25 min. The precision of emergency detection in this study case is 91.67 %.

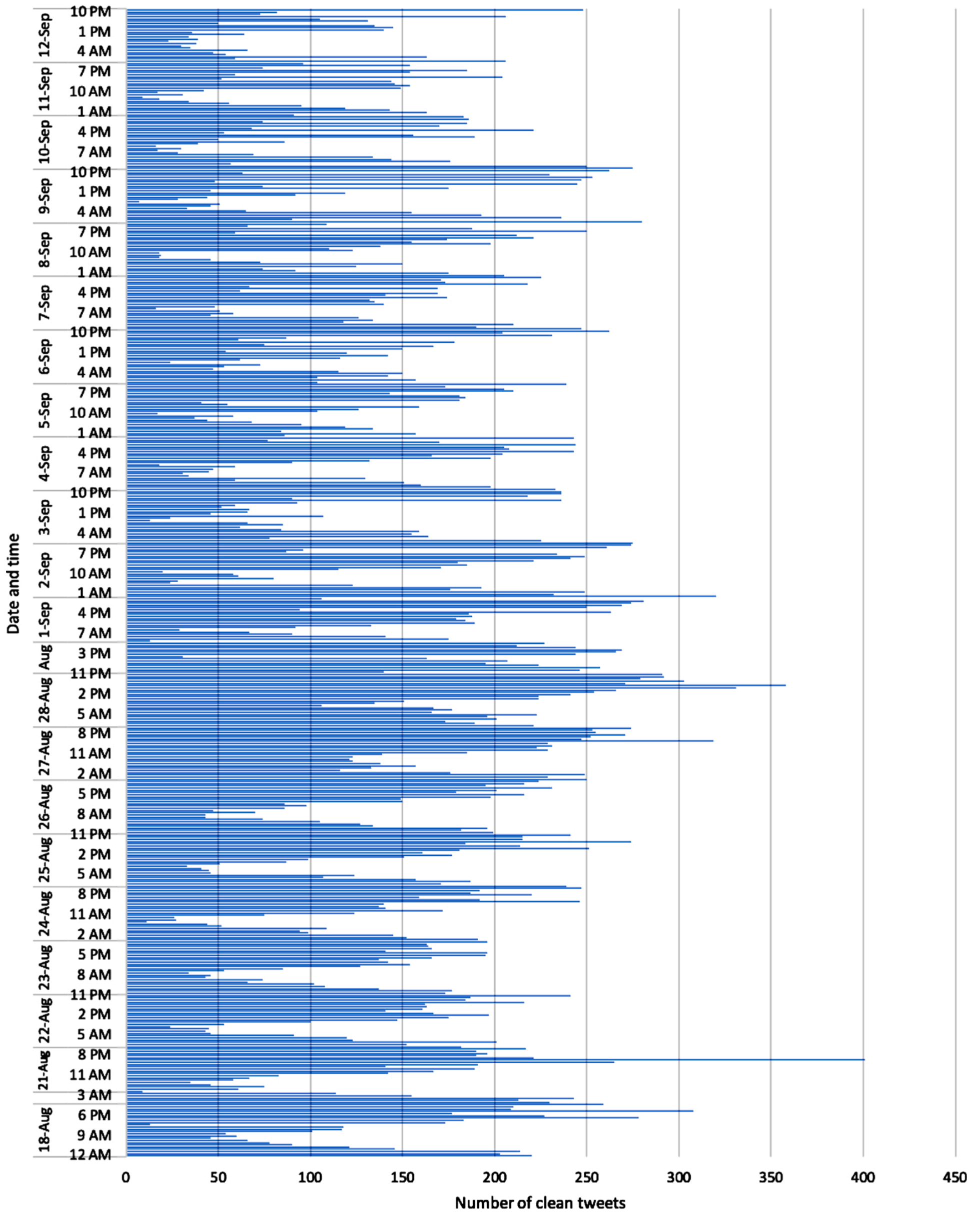


Fig. 10. Hourly clean data volume processed by the system.

4.3. System performances and data statistical features

4.3.1. Streaming data volume and temporal patterns

Our system processed about 2,707,346 tweets during the study period. The daily numbers of tweets before and after the data filtering are shown in Fig. 9. Before the filtering, the Twitter streaming API received 100,000 to 140,000 tweets a day. The peak throughput (volume) of our system is 147,589 tweets per day, and its average throughput is 117,710.

Focusing on the specifically hurricane-affected area, the number of clean tweets (no bots and geotagged) had a maximum of 5,439 and an average of 3,189 per day. These tweets were considered useful for improving situational awareness during disasters. The cleaned tweets flowed to the following modules for analysis and visualization. After the hurricane made landfall on August 25, Twitter activity measured by the number of clean tweets increased dramatically and reached its peak on August 28, then became relatively stable after the disaster.

Fig. 10 shows the hourly change in Twitter activity over the study period. The clean tweets increased and reached a peak then decreased every day, both before and after the hurricane. During the hurricane-affected period (after August 25 until September 1), the daily peaks were greater than before or after the hurricane. The hourly peak throughput of cleaned data for the subsequent analysis was 401 tweets, and the average number was 142 tweets.

4.3.2. Feature patterns of geo-topics

Our system generated geo-topics with features such as centers, ranges, and sentiment scores over time. The results showed that the system can detect a wide range of topics, including disaster-relevant events and daily life events, and monitor their features over time. Although many disaster-relevant events did not trigger emergency alerts, they still provided rich enough information to improve situational awareness.

We aggregated the daily temporal patterns of several geo-topics with counts of their represented geotagged tweets (Fig. 11). Table 1 lists a few of these geo-topics and their top keywords, which have the highest probabilities of representing those geo-topics. For example, during the hurricane period, our system detected multiple hurricane-relevant geo-

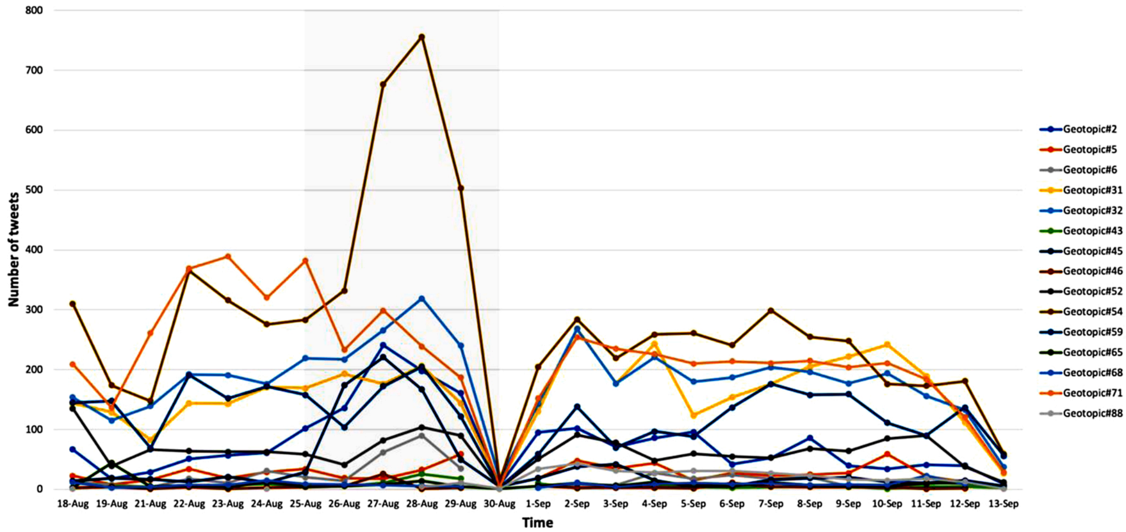


Fig. 11. Temporal patterns of selected geo-topics and the count of their represented tweets.

Table 1
Selected geo-topics of Hurricane Harvey and their top keywords.

Geo-topic	Top keywords
2	flood, flash, county, report, free, warning, giveaway, include, storm, public
31	birthday, drink, happy, college, great, bar, photo, food, restaurant, show
32	harvey, today, come, sanantonio, hurricaneharvey, time, hurricane, last, austin, tonight
45	tornado, warning, lafayette, continue, lakes, city, mission, parish, weston, brookshire
54	traffic, stop, accident, lane, delay, hwy, block, high, water, rock
71	austin, nowplaying, pool, palmillabeach, atx, downtown, livingston, near, airport, long

topics (e.g., geo-topic #2, 32, and 54) with keywords such as “hurricane”, “Harvey”, “flood” and “accident”, and multiple locations of hurricane-affected areas such as “Austin”, “San Antonio”, “downtown”, or “highway”. Geo-topic #2 indicated flash flooding and stormwater conditions during the hurricane, geo-topic #54 was related to traffic accidents on the highways, and geo-topic #32 was about the general situation of Hurricane Harvey. The system also found that such geo-topics reached their volume peaks in the numbers of represented tweets during the hurricane-affected period (from August 25 to August 29) and then gradually disappeared afterward.

Our system also detected some hurricane-irrelevant events and tracked their changes. For example, geo-topic #31 had keywords such as “birthday”, “bar”, and “restaurant”. The volumes of such geo-topics were relatively stable over time, unlike hurricane-relevant geo-topics that reached peak volume. We also found that some geo-topics appeared only in parts of the study period. For example, geo-topic #88 with keywords such as “jog”, “partner”, and “produce” appeared only after the hurricane.

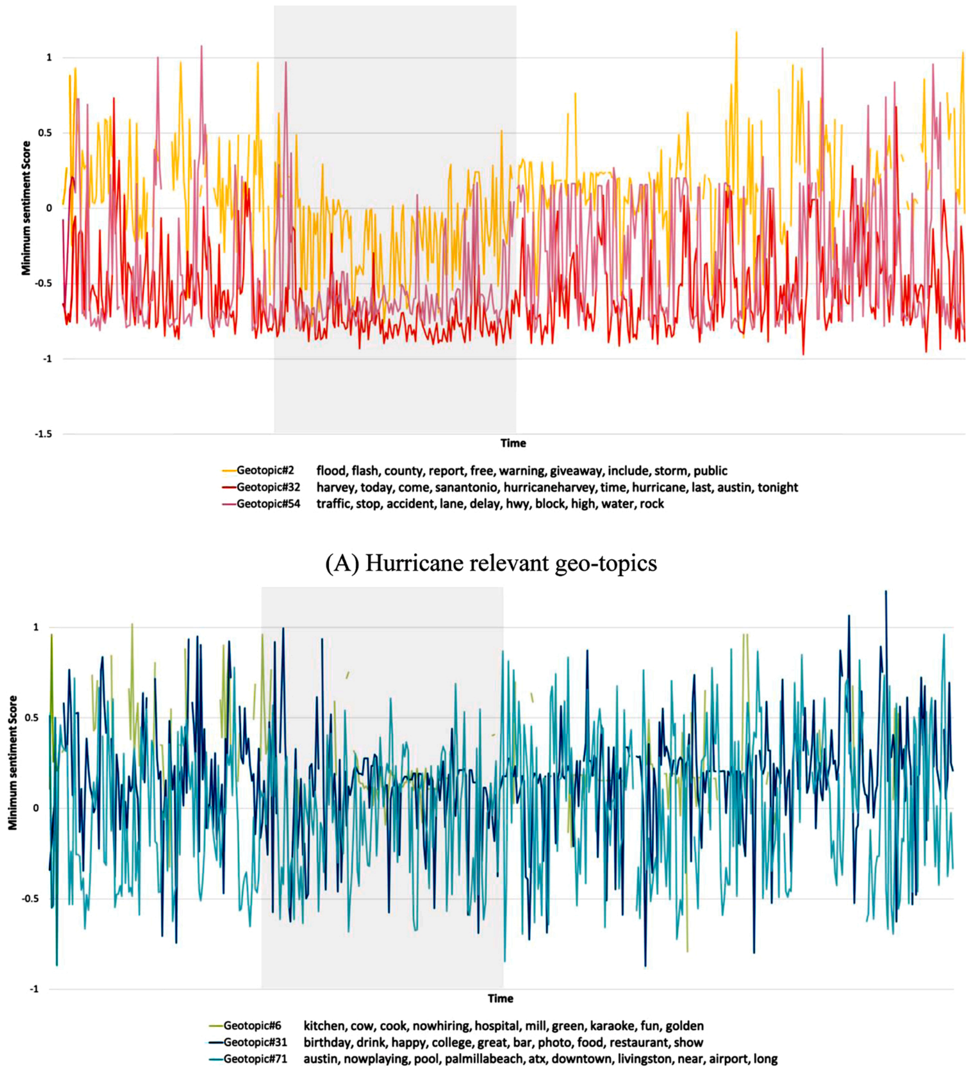
We also analyzed the sentiment changes on each geo-topic over time. To further demonstrate the sentiment patterns and their relationships to different geo-topics, we selected several geo-topics that represented hurricane-relevant and -irrelevant events and then compared their sentiment patterns over time (Fig. 12). We used the minimum sentiment scores among the tweets belonging to the geo-topic because minimum scores revealed more apparent temporal trends. We found that hurricane-relevant geo-topics (Fig. 12A) had more negative sentiment overall throughout the study period, and the most negative sentiments during the period when the hurricane hit the study area. By contrast, hurricane-irrelevant geo-topics (Fig. 12B) had more positive and stable sentiment and were affected less by the hurricane.

5. Discussion

In this paper, we have presented an urban analytical and geo-visual system that automatically collects geotagged tweets and performs urban event detection and visualization in real time. Through the application in the simulated streaming data from a large-scale hurricane, the system has been demonstrated to provide useful and timely information on emergencies and crises during disasters. The system is specifically designed to address the smart rapidity aspect of resilient cities by enabling real-time analysis and geo-visualization to decipher the dynamics of urban environments and systems. Early detection and tracking of the urban events help provide early alerts to the residents and assist city managers and first responders to rapidly respond to the adverse effects.

Processing and analyzing social media streaming data in real time is exceptionally challenging because the speed and volume of data stream require methods to update rapidly to capture the topic changes (Goyal et al., 2019; Hasan, Orgun, & Schwitter, 2018; Nugroho, Paris, Nepal, Yang, & Zhao, 2020). Our proposed system improves on previous research by implementing online data-analysis methods that enable the rapid detection and processing of streaming data (Yao & Wang, 2020a). We used a variational inference method—Online LDA (Hoffman et al., 2010)—to process and infer the topics of incoming data using Bayes' theorem. In general, topic models can be computationally complex and time-consuming (Xie et al., 2016; Yu et al., 2017). The system reduces the redundant computation for topic modeling by applying multiple preprocessing methods that are much faster than topic modeling. By filtering and cleaning streaming raw tweets, and only triggering the later topic modeling and sentiment analysis functions when necessary, we improve the accuracy and efficiency of the overall system. We also exploit the Apache Spark distributed analytics system to improve the scalability and speed of data processing. All these methods and settings can help our system increase data throughput and reduce latency in data processing to achieve a real-time analysis and visualization system.

Our developed system can continuously track multidimensional information from the data stream, such as location, timestamp, semantics, and sentiment changes of detected events. This property expanded on previous research that offered simple place or place–time information reports (e.g., Bifet, Holmes, & Pfahringer, 2011; MacEachren et al., 2010; Wachowicz et al., 2016) or spatial clustering results (Pezanowski et al., 2018; Smith et al., 2017). Some earlier methods used bursty keywords as prerequisites for event detection and tracking (Avvenuti



(A) Hurricane relevant geo-topics

(B) Hurricane irrelevant geo-topics

Fig. 12. Temporal patterns of sentiment for selected geo-topics.

et al., 2014; Li, Sun, & Datta, 2012). Our system does not set keywords for event detection, which gives it the potential to detect a variety of unforeseeable urban events related to disasters and other unexpected perturbations. Instead of using keywords, our system employs sentiment analysis to automatically identify potential emergencies or crises at an early stage. Sentiment analysis in existing systems was classified into two or three categories and demonstrated only in basic statistical features, such as the percentage of each category of sentiment (Choi & Bae, 2015; Marcus et al., 2011). Our system can calculate sentiments in finer-gradations and can be used to trigger early alerts about potential emergencies to help with decision-making.

Although we have used the system to demonstrate one study case, the system is open to other data resources and application scenarios due to its flexible and adaptable design. Supporting by the GeoEvent Server, the system can ingest streaming API from different resources and provide additional analysis by properly changing the adapters, filters, or processors within the system modules. The users of our system can also customize the geo-visualization effects (colors, legends, base maps) as needed through Online GIS maps. Thus, our system is advantageous for transforming extracted information into a broad spectrum of applications, ranging from extreme events such as disaster management, epidemics tracking, and crime monitoring to business-as-usual situations such as place recommendation (Gao et al., 2020; Ghani et al., 2019; Nugroho et al., 2020; Pezanowski et al., 2018).

Additionally, our system uses a dynamic online map to visualize the latest analysis results update every few seconds. Managing extreme events in cities can be complicated, users of our system may not be familiar with data analysis but still need specific information to improve situational awareness, allocate resources, or take action. This geo-visualization feature allows users to learn intuitively about important events happening in their areas and whether people should be aware of those events during unpredictable extreme situations. Using crowd-sourced social media data generated by the public, our system also provides citizens opportunities to be actively involved in the system and promotes community engagement.

However, the proposed system has several limitations that can be addressed in our future research. Currently, the system uses Twitter as its only data source. However, people have differencing preferences in using social networking platforms. Future systems can consider streaming data from multiple sources to become more integrated and reduce the data bias caused by single data source. Additionally, the proposed system performs streaming data analyses based on pre-defined parameters, such as the number of topics and the sentiment threshold. These parameters need to be set case by case due to the complexity in contents and languages online. Future studies will focus on testing and tuning parameters across urban events with different study areas and periods, and propose a method for system users to set parameters for best performances. Lastly, it is challenging to compare the outcomes (e.g., small-scale local events) of the proposed system with the ground truth because these detected events may not be reported by officials, which makes the data unavailable. Although the system is constructed with well-developed data analysis methods, the usage of social media data requires further credibility checks. The future system can also be improved with additional event analysis modules when increasing the volume and versatility of streaming data becomes available.

6. Conclusion

Building resilient cities requires smart solutions, and achieving smart rapidity is one of the most important approaches to enhancing urban resilience when promoting smart cities. We develop a real-time urban analytical and geo-visual system for social media streaming data to track small-scale undefined urban extreme events and provide early emergency alerts. The system has demonstrated the effectiveness and rapidity in processing large volumes of data with low latency. The system has the potential to incorporate streaming data from more sources and to be involved in cities' emergency management tasks, such as improving situational awareness, assisting rapid damage assessments, monitoring emergent incidents, and supporting collaborative decision-making for multiple stakeholders. The research also contributes to developing smart city technologies that can be integrated into holistic urban surveillance systems and achieving more safe, resilient, and smart future cities.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1760645, Grant No. 2028012, Grant No. 1951816, and University of Florida DCP Research Seed Grant Initiative. 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 or University of Florida.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.scs.2020.102448.

References

Al Nuaimi, E., Al Neyadi, H., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications, 6*(1), 1–15.
Allam, Z., & Newman, P. (2018). Redefining the smart City: Culture, metabolism and governance. *Smart Cities, 1*(1), 4–25.
Angelidou, M., Psaltoglou, A., Komninos, N., Kakderi, C., Tsarchopoulos, P., & Panori, A. (2018). Enhancing sustainable urban development through smart city applications. *Journal of Science and Technology Policy Management, 9*(2), 146–169.
Arafah, Y., & Winarso, H. (2017). Redefining smart city concept with resilience approach. *IOP Conference Series: Earth and Environmental Science, 70*(1), 12065.
Avvenuti, M., Cresci, S., Marchetti, A., Meletti, C., & Tesconi, M. (2014). EARS (earthquake alert and report system): A real time decision support system for earthquake crisis management. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1749–1758.
Batty, M. (2008). The size, scale, and shape of cities. *Science, 319*(5864), 769–771.
Benhardus, J., & Kalita, J. (2013). Streaming trend detection in Twitter. *International Journal of Web Based Communities, 9*(1), 122–139.
Bifet, A., Holmes, G., & Pfahringer, B. (2011). MOA-TweetReader: Real-time analysis in twitter streaming data. *International Conference on Discovery Science*, 46–60.
Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *ACM International Conference Proceeding Series, 148*, 113–120.
Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.
Boettcher, A., & Lee, D. (2012). EventRadar: A real-time local event detection scheme using Twitter stream. *IEEE International Conference on Green Computing and Communications*, 358–367.
Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., et al. (2003). A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra, 19*(4), 733–752.
Caragea, C., Squicciarini, A., Stehle, S., Neppalli, K., & Tapia, A. (2014). Mapping moods: Geo-mapped sentiment analysis during hurricane sandy. *ISCRAM 2014 Conference Proceedings*, 642–651.
Chae, J., Thom, D., Jang, Y., Kim, S., Ertl, T., & Ebert, D. S. (2014). Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Computers and Graphics (Pergamon), 38*(1), 51–60.
Choi, S., & Bae, B. (2015). The real-time monitoring system of social big data for disaster management. *Computer science and its applications* (pp. 809–815). Berlin, Heidelberg: Springer.
Desouza, K. C., & Flanery, T. H. (2013). Designing, planning, and managing resilient cities: A conceptual framework. *Cities, 35*, 89–99.
Desroches, R., & Taylor, J. E. (2018). The promise of smart and resilient cities. *Bridge, 48* (2), 13–20.
Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., et al. (2016). Domain-adversarial training of neural networks. *Journal of Machine Learning Research, 17*, 1–35.
Gao, W., Peng, M., Wang, H., Zhang, Y., Han, W., Hu, G., et al. (2020). Generation of topic evolution graphs from short text streams. *Neurocomputing, 383*, 282–294.
Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior, 101*, 417–428.
Godschalk, D. R. (2003). Urban hazard mitigation: Creating resilient cities. *Natural Hazards Review, 4*(3), 136–143.
Goyal, P., Kaushik, P., Gupta, P., Vashisth, D., Agarwal, S., & Goyal, N. (2019). Multilevel event detection, storyline generation, and summarization for tweet streams. *IEEE Transactions on Computational Social Systems, 7*(1), 8–23.
Hasan, M., Orgun, M. A., & Schwitter, R. (2018). A survey on real-time event detection from the Twitter data stream. *Journal of Information Science, 44*(4), 443–463.
Hasan, M., Orgun, M. A., & Schwitter, R. (2019). Real-time event detection from the Twitter data stream using the TwitterNews+ Framework. *Information Processing and Management, 56*(3), 1146–1165.
Hatuka, T., Rosen-Zvi, I., Birnhack, M., Toch, E., & Zur, H. (2018). The political premises of contemporary urban concepts: The global city, the sustainable city, the resilient city, the creative city, and the smart city. *Planning Theory and Practice, 19*(2), 160–179.
Hoffman, M. D., Blei, D. M., & Bach, F. (2010). Online learning for Latent Dirichlet Allocation. *Advances in Neural Information Processing Systems*, 856–864.
Hofmann, T. (1999). Probabilistic latent semantic indexing. *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 50–57.
Huang, Q., Cervone, G., & Zhang, G. (2017). A cloud-enabled automatic disaster analysis system of multi-sourced data streams: An example synthesizing social media, remote sensing and Wikipedia data. *Computers, Environment and Urban Systems, 66*, 23–37.
Khan, Z., Anjum, A., Soomro, K., & Tahir, M. A. (2015). Towards cloud based big data analytics for smart future cities. *Journal of Cloud Computing, 4*(1), 1–11.
Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1–14.
Klein, B., Koenig, R., & Schmitt, G. (2017). Managing urban resilience: Stream processing platform for responsive cities. *Informatik-Spektrum, 40*(1), 35–45.
Kontokosta, C. E., & Malik, A. (2018). The resilience to emergencies and disasters index: Applying big data to benchmark and validate neighborhood resilience capacity. *Sustainable Cities and Society, 36*, 272–285.
Lau, J. H., Collier, N., & Baldwin, T. (2012). On-line trend analysis with topic models: Twitter trends detection topic model online. *24th International Conference on Computational Linguistics - Proceedings of COLING 2012: Technical Papers*, 1519–1534.
Leichenko, R. (2011). Climate change and urban resilience. *Current Opinion in Environmental Sustainability, 3*(3), 164–168.
Leykin, D., Lahad, M., & Aharonson-Daniel, L. (2018). Gauging urban resilience from social media. *International Journal of Disaster Risk Reduction, 31*, 393–402.
Li, C., Sun, A., & Datta, A. (2012). Twevent: Segment-based event detection from tweets. *ACM International Conference Proceeding Series*, 155–164.
Lu, Y., Hu, X., Wang, F., Kumar, S., Liu, H., & Maciejewski, R. (2015). Visualizing social media sentiment in disaster scenarios. *WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web*, 1211–1215.
MacEachren, A., Robinson, A. C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J., et al. (2010). Geo-twitter analytics: Applications in crisis management. *Proceedings of the 25th International Cartographic Conference*, 1–8.
Malik, K. R., Sam, Y., Hussain, M., & Abuarqoub, A. (2018). A methodology for real-time data sustainability in smart city: Towards inferencing and analytics for big-data. *Sustainable Cities and Society, 39*, 548–556.
Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., & Miller, R. C. (2011). TwitInfo: Aggregating and visualizing microblogs for event exploration. *Conference on Human Factors in Computing Systems - Proceedings*, 227–236.
Mathioudakis, M., & Koudas, N. (2010). TwitterMonitor: Trend detection over the twitter stream. *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 1155–1157.
McGowan, C. J., Biggerstaff, M., Johansson, M., Apfeldorf, K. M., Ben-Nun, M., Brooks, L., et al. (2019). Collaborative efforts to forecast seasonal influenza in the United States, 2015–2016. *Scientific Reports, 9*(1).
Meerow, S., Newell, J. P., & Stults, M. (2016). Defining urban resilience: A review. *Landscape and Urban Planning, 147*, 38–49.

Middleton, S. E., Middleton, L., & Modafferi, S. (2014). Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems, 29*(2), 9–17.
Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current trends in smart city initiatives: Some stylised facts. *Cities, 38*, 25–36.
Nugroho, R., Paris, C., Nepal, S., Yang, J., & Zhao, W. (2020). A survey of recent methods on deriving topics from Twitter: Algorithm to evaluation. *Knowledge and Information Systems*, 1–35.
Palmieri, F., Ficco, M., Pardi, S., & Castiglione, A. (2016). A cloud-based architecture for emergency management and first responders localization in smart city environments. *Computers and Electrical Engineering, 56*, 810–830.
Pezanowski, S., MacEachren, A. M., Savelyev, A., & Robinson, A. C. (2018). SensePlace3: A geovisual framework to analyze place–time–attribute information in social media. *Cartography and Geographic Information Science, 45*(5), 420–437.
Platt, S., Brown, D., & Hughes, M. (2016). Measuring resilience and recovery. *International Journal of Disaster Risk Reduction, 19*, 447–460.
Psyllidis, A., Bozzon, A., Bocconi, S., & Titos Bolivar, C. (2015). A platform for urban analytics and semantic data integration in city planning. *Communications in Computer and Information Science, 527*, 21–36.
Şerban, O., Thapen, N., Maginnis, B., Hankin, C., & Foot, V. (2019). Real-time processing of social media with SENTINEL: A syndromic surveillance system incorporating deep learning for health classification. *Information Processing & Management, 56*(3), 1166–1184.
Smith, L., Liang, Q., James, P., & Lin, W. (2017). Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. *Journal of Flood Risk Management, 10*(3), 370–380.
Soyata, T., Habibzadeh, H., Ekenna, C., Nussbaum, B., & Lozano, J. (2019). Smart city in crisis: Technology and policy concerns. *Sustainable Cities and Society, 50*, Article 101566. https://doi.org/10.1016/j.scs.2019.101566.
Terpstra, T., Stronkman, R., De Vries, A., & Paradies, G. L. (2012). Towards a real-time Twitter analysis during crises for operational crisis management. *ISCRAM 2012 Conference Proceedings*, 1–9.
Thom, D., Bosch, H., Koch, S., Worner, M., & Ertl, T. (2012). Spatiotemporal anomaly detection through visual analysis of geolocated Twitter messages. *IEEE Pacific Visualization Symposium 2012, PacificVis 2012 - Proceedings*, 41–48.
UN DESA Population Division. (2018). *The world's cities in 2018*. Retrieved from: https://www.un.org/en/events/citiesday/assets/pdf/the_worlds_cities_in_2018_data_booklet.pdf.
Wachowicz, M., Arteaga, M. D., Cha, S., & Bourgeois, Y. (2016). Developing a streaming data processing workflow for querying space–time activities from geotagged tweets. *Computers, Environment and Urban Systems, 59*, 256–268.
Wang, X., & McCallum, A. (2006). Topics over time: A non-Markov continuous-time model of topical trends. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006*, 424–433.
Wang, Y., & Taylor, J. E. (2019). DUET: Data-driven approach based on Dirichlet allocation topic modeling. *Journal of Computing in Civil Engineering, 33*(3), Article 04019023.
Wang, Y., Wang, Q., & Taylor, J. E. (2017). Aggregated responses of human mobility to severe winter storms: An empirical study. *PloS ONE, 12*(12), Article e0188734.
Wang, Y., Hulse, D., Von Meding, J., Brown, M., & Dedenbach, L. (2019). Conceiving resilience: Lexical shifts and proximal meanings in the human-centered natural and built environment literature from 1990 to 2018. *Developments in the Built Environment*. https://doi.org/10.1016/j.dibe.2019.100003.
Wang, Y., Taylor, J. E., & Garvin, M. J. (2020). Measuring resilience of human–spatial systems to disasters: Framework combining spatial-network analysis and fisher information. *Journal of Management in Engineering, 36*(4), Article 04020019.
Woetzel, J., Remes, J., Boland, B., Lv, K., Sinha, S., Strube, G., et al. (2018). *Smart cities: Digital solutions for a more livable future*. McKinsey & Company.
Xie, W., Zhu, F., Jiang, J., Lim, E. P., & Wang, K. (2016). TopicSketch: Real-time bursty topic detection from twitter. *IEEE Transactions on Knowledge and Data Engineering, 28*(8), 2216–2229.
Yang, C., Su, G., & Chen, J. (2017). Using big data to enhance crisis response and disaster resilience for a smart city. *IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*, 504–507.
Yao, F., & Wang, Y. (2020a). Tracking urban geo-topics based on dynamic topic model. *Computers, Environment and Urban Systems, 79*, Article 101419. https://doi.org/10.1016/j.compenvurbsys.2019.101419.
Yao, F., & Wang, Y. (2020b). Domain-specific sentiment analysis for tweets during hurricanes (DSSA-H): A domain-adversarial neural-network-based approach. *Computers, Environment and Urban Systems, 83*, Article 101522. https://doi.org/10.1016/j.compenvurbsys.2020.101522.
Yoshinaga, N., & Kitsuregawa, M. (2014). A self-adaptive classifier for efficient text-stream processing. *Proceedings of the 25th International Conference on Computational Linguistics, 2014: Technical Papers*, 1091–1102.
Yu, W., Li, J., Bhuiyan, M. Z. A., Zhang, R., & Huai, J. (2017). Ring: Real-time emerging anomaly monitoring system over text streams. *IEEE Transactions on Big Data, 5*(4), 506–519.
Zhang, X., & Li, H. (2018). Urban resilience and urban sustainability: What we know and what do not know? *Cities, 72*, 141–148.
Zhang, C., Liu, L., Lei, D., Yuan, Q., Zhuang, H., Hanratty, T., et al. (2017). TrioVecEvent: Embedding-based online local event detection in geo-tagged tweet streams. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 595–604.
Zhang, F., Li, Z., Li, N., & Fang, D. (2019). Assessment of urban human mobility perturbation under extreme weather events: A case study in Nanjing, China. *Sustainable Cities and Society, 50*, Article 101671.
Zhu, S., Li, D., & Feng, H. (2019). Is smart city resilient? Evidence from China. *Sustainable Cities and Society, 50*. https://doi.org/10.1016/j.scs.2019.101636.
Zobel, C. W. (2011). Representing perceived tradeoffs in defining disaster resilience. *Decision Support Systems, 50*(2), 394–403.