GEORGIA DOT RESEARCH PROJECT 18-34

Final Report

SOCIAL MEDIA-INFORMED URBAN CRISIS DETECTION

Georgia Department of Transportation

Office of Performance-based Management and Research

600 West Peachtree NW
Atlanta, GA 30308

TECHNICAL REPORT DOCUMENTATION

1. Report No. FHWA-GA-20-1834	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A
4. Title and Subtitle Social Media-Informed Urban Crisis Detection		**5. Report Date** July 2020 **6. Performing Organization Code** N/A
7. Author(s) Rachel Samuels Neda Mohammadi, Ph.D. John E. Taylor, Ph.D.		**8. Performing Organization Report No.** 18-34
9. Performing Organization Name and Address School of Civil and Environmental Engineering Georgia Institute of Technology 790 Atlantic Drive NW Atlanta, GA 30332		**10. Work Unit No.** N/A **11. Contract or Grant No.** Task Order No. 2014-86
12. Sponsoring Agency Name and Address Georgia Department of Transportation Office of Performance-based Management and Research 600 W. Peachtree St. NW Atlanta, GA 30308		**13. Type of Report and Period Covered** Final Report (May 2019 – July 2020) **14. Sponsoring Agency Code** N/A
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.		

16. Abstract
This research project examines the potential utility of a social media-informed crisis detection system to aid GDOT in the identification of hazards across the state. To assess that utility, the research team first analyzed two case studies with the goal of identifying the relevant information contained within posted social media data. Case study data were drawn from hundreds of thousands of social media postings in Georgia, and include: 1) the winter storm that impacted north Georgia January 16-17, 2018 (N = 436 after processing) and 2) the flooding that occurred across the state of Georgia from Tropical Storm Irma from September 10-17, 2017 (N = 910 after processing). Nearest Neighbor Ratio analysis identified that postings were geographically clustered in areas of higher populations, but not statistically significantly clustered in terms of either sentiment or primary topic. The research team then analyzed sentiment as a method of ranking the social media data relevance and determined that neutral sentiment can function as a secondary filtration method for relevance. Finally, the team specified an approach to integrating this data visually that aligns with GDOT processes and with existing Waze data streams currently utilized by GDOT. The findings from this research and, in particular, the specification of a social media-informed crisis detection system, provide GDOT and other state agencies with a deeper understanding of the role social media can serve in crisis detection, tracking, and visualization, as well as providing a set of specifications to design and implement such a system.

17. Key Words Crisis detection; Crisis response; Disaster and emergency operations; Information management; Social media		**18. Distribution Statement** No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.	
19. Security Classification (of this report) Unclassified	**20. Security Classification (of this page)** Unclassified	**21. No. of Pages** 79	**22. Price** Free

GDOT Research Project 18-34

Final Report

SOCIAL MEDIA-INFORMED URBAN CRISIS DETECTION

by

Rachel Samuels
Graduate Research Assistant

Neda Mohammadi
City Infrastructure Analytics Director

John E. Taylor
Frederick L. Olmsted Professor

Name of Contractor
Georgia Tech Research Corporation
Georgia Institute of Technology

Contract with
Georgia Department of Transportation

In Cooperation with
U.S. Department of Transportation Federal Highway Administration

July 2020

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Recent natural and man-made crises have impacted the lives of Georgia residents. In 2017 alone, natural disasters due to the Atlantic hurricane season and the major infrastructure crisis of the Interstate 85 bridge collapse in Atlanta generated significant social, environmental, and economic damage. Hurricane Irma placed Atlanta under its first ever tropical storm warning and required mass evacuations of all areas east of Interstate 95 in coastal areas of Georgia, which lead to inland evacuation orders for 540,000 coastal residents. In such crises, residents are increasingly turning to social media for help and information, and many disaster response organizations are incorporating social media platform data into their existing crisis identification and emergency management processes and systems. The focus of this research is addressing the current gaps in our ability to effectively detect crises through social media and extending our research to the context of hazards in the Georgia transportation system.

Systems that can detect and track crises in near real-time can be critical components of rapid crises identification and response deployment decisions. Identified crises can inform GDOT managers and first responders when and where an emergency is, the emergency details and sentiment level, and how the emergency evolves over time. Crucially, a more immediate detection and response system can provide information to reduce potential casualties and damages and improve allocation of scarce resources. In this research project, we examined the potential of a social media-informed crisis detection system that takes advantage of the aforementioned capabilities to aid GDOT in the identification of hazards across the state. The high-level objective of this preliminary research project was to identify the potential utility of social media data to detect, track, and visualize such hazards. To assess that utility, the research team first performed two case studies, one for a winter storm event and one for an episode of extreme flooding. Within

each case study, the goal was to identify the relevant information contained within posted social media data, evaluate the relevance and importance of that data and what can be determined through Natural Language Processing (NLP) techniques, and specify a method to integrate this into GDOT operations that could be developed in a potential future research project. The case studies that were analyzed include: 1) the winter storm that impacted north Georgia on January 16th and 17th, 2018 (number of sample points (N) = 436 after processing) and 2) the flooding that occurred across the state of Georgia as a result of Tropical Storm Irma from September 10th to the 17th, 2017 (N = 910 after processing).

The research team worked with the GDOT Technical Implementation Team for this research project to identify the key issues for each case study that were of primary relevance to GDOT operations. Georgia produces the 5th highest number of daily Tweets of the states within the United States (Stirtz n.d.) and roughly 2 million Twitter users, and these users generate 130,000 Tweets each day. Of those Tweets, the ones which were geo-located with a specific address and posted on the case study days were filtered down using NLP and topic analysis techniques, including stop words and regular expressions. The remaining social media postings were then assessed for their spatial proximity and how well clustered the topics were. The Nearest Neighbor Ratio analysis identified that the Tweets were geographically clustered in areas of higher populations, as expected; however, the Tweets were not statistically significantly clustered in terms of either sentiment or primary topic. This indicates that there are certain areas in which social media will be a more likely indicator of an event (such as cities), and also that individual social media postings are likely addressing distinct concerns and should primarily be analyzed on the individual scale for relevance. Additionally, on review of the subject topics within city-clusters, there was a higher percentage of postings relevant to GDOT concerns outside of city limits than within them. This indicates that valuable information is still available outside of urban populations, if less of it, and that distinct filtration methods may be necessary for urban versus rural areas.

The research team then analyzed sentiment as a method of ranking the social media data relevance. The general sentiment derived from the case studies was normally distributed with a slight right skew with a peak at the 0.0 mark, indicating little if any positive or negative sentiment. Nevertheless, the research team determined that most of these postings were describing specific incidents relevant to the broader crisis event and contained valuable information. We additionally investigated the sentiment distributions of each keyword for each case study and determined that neutral sentiment can function as a secondary filtration method for social media data relevance. These results additionally identified the potentially beneficial use of sentiment in determining which topics are useful to GDOT.

The final stage of the research was to take the findings from the first two stages and specify a method of integrating this data systemically and visually such that it aligns with GDOT processes. On comparison of the amount of relevant data, the extent of the available social media data, feedback from the GDOT Technical Implementation Team and the Social Media Coordination Officers, and research into existing social media visualization and interaction platforms, the team determined that designing a method for integrating the social media data with existing Waze data streams currently utilized by GDOT would function best at providing GDOT additional, relevant information on newly developing crises. Merging two user volunteered information streams instead of generating a new platform diminishes disruption to existing data streams and takes advantage of existing information visualization platforms at the GDOT State Operations Center. In specifying the system functionality, we focused primarily on the format of the data integration. The research team designed a framework for producing a value for Twitter social media data that aligns with Waze data's confidence, location, and topic outputs. The team then developed specifications for the design of the combined social media data (Twitter) and Waze system and how it would be displayed to the end user (the GDOT Social Media Communications Officer). The findings from this research and, in particular, the specification of a social media-informed crisis detection system, provides GDOT and other state agencies with a deeper understanding of the role social media can serve in crisis detection, tracking, and visualization, as well as a set of specifications to design and implement such a system.

1 INTRODUCTION

Recent natural and man-made crises have had major impacts on the life of Georgia residents. In 2017 alone, natural disasters due to the Atlantic hurricane season, which caused heavy rains and flooding across the state, and a major infrastructure crisis such as the Interstate 85 bridge collapse in Atlanta, have generated significant social, environmental, and economic damage. Hurricane Irma placed Atlanta under its first ever tropical storm warning and required mass evacuations of all areas east of Interstate 95 in coastal areas of Georgia. Hurricane Irma resulted in a total of three fatalities and inland evacuation orders for 540,000 coastal residents. Notwithstanding the authorities' orders and warnings, many residents failed to follow the procedures, which exacerbated a range of quickly-developing crises, including: stranded residents in rising storm surge, shortages of gas, water and food, power outages, infrastructure failures, fires, traffic jams, traffic incidents, evacuation barriers, looting, and other crises that require rapid response and emergency assistance. Under such circumstances, where the crisis is too time critical for a 911 call to be effective (e.g., people cannot make phone calls, emergency telephone hotlines are jammed, or the emergency responders are unable to assess the relative gravity of one crisis over another), residents are increasingly turning to social media for help, even posting their full addresses in desperation. This highlights the significance of timely emergency communication, and the need to incorporate social media platforms (e.g., Twitter) into existing crisis identification and emergency management systems.

Especially as our cities become smarter and more connected, it is important now more than ever to lay the groundwork for how to incorporate non-traditional forms of information and communication into our response structures. The systems that we use to identify, evaluate, and respond to hazards need to be adaptable in order to be comprehensive. As social media is one of the emerging forms of citizen communication with both each other and response agencies that is clearly becoming more widespread and utilized, understanding the merit in combining this non-traditional form of information with the existing Georgia Department of Transportation (GDOT) system is critical. The focus of this research, then, is to evaluate the best methods for incorporating social media activity and posts into the event identification and event response framework currently utilized by GDOT.

Unfortunately, social media as a sensor is often messy, and there is a large amount of data that needs to be filtered down to what is relevant not simply to emergency response but specifically to the roadways under GDOT's jurisdiction. Recent developments in Natural Language Processing (NLP) and topic detection need to be utilized to identify the most efficient methods of filtration. Ranking the relative information value of each Tweet will also be critical in identifying what needs urgent attention and what is supplemental information for broader event knowledge. As the formatting of social media changes, identifying the most useful aspects of social media data for GDOT will also be necessary in prioritizing filtration methods and the parsed data. Finally, it will be necessary to work closely with the various moving parts of GDOT to understand what has already been done within the organization, what the current information analysis framework is, and what the least disruptive and most beneficial method of data incorporation would be for GDOT as a response and resource distribution agency.

In view of the various stages of data collection, analysis, and system design outlined above, such research will lay the groundwork for developing methods of incorporating multiple separate and new data streams into the existing system that GDOT is utilizing to identify, assess, and respond to roadway incidents and citizen needs. Although the goal of this research project is to specify a social media-informed event tracking and visualization system, we expect that the frameworks produced in the scope of this research will provide an overall sense of the requirements for utilizing streams of social media data, the extent of that data's value, the most current methods for automatically assessing and extracting valuable information from that data, and, finally, the optimal path for incorporating additional data streams into the systems that are already in existence. It is our hope that this research can serve as a beneficial reference for the development of systems capable of seamlessly incorporating the expanding quantities of social media data into GDOT's event detection system, especially as citizens produce more of that data—and expect more from GDOT's response in return.

2 LITERATURE REVIEW

The purpose of this research is to investigate the ability of user volunteered information (UVI), especially in the form of social media, to augment crisis detection and monitoring on state-owned roads. Although the field of crisis informatics has been seeking to develop methods for parsing information from social media for more than a decade, the types of available information, their format, and the tools available for analysis have been constantly evolving (Reuter and Kaufhold 2017). The situations that have been investigated using UVI also vary as widely as social media topics do. As such, our preliminary literature review focused on three main areas of study: how UVI has been used in crisis identification and response, what kinds of UVI analytics have been incorporated into existing monitoring platforms, and, finally, what kinds of UVI are both available and relevant to GDOT.

2.1 Social Media Usage in Crisis Response

UVI has been recognized as a potential source of actionable information since the use of social media and personal communication devices surged in the late-2000s. Users creating social media posts can include information about their whereabouts, what they are seeing at a given moment, what they are concerned about, what help they need, and what help they are able to give. Additionally, many of the social media sites provide certain amounts of their data through Application Programming Interfaces (APIs) at no charge in order to market that additional data for corporations. Researchers have additionally been able to access those social media streams for use in improving emergency situational awareness (Yin et al. 2015). This idea of using citizens as sensors--mobile, active sensors in places where people are being impacted by an emergency--has been able to aid real-time analysis of emergencies, group coordination, and collective action (Purohit et al. 2014).

The research community has developed applications that automatically filter and analyze these posts, and Twitter is one of the primary sources used by these applications for emergency detection and tracking (Imran et al. 2014b; a; Purohit et al. 2014) Professional social media analytical applications are currently being used by organizations such as the American Red Cross and the United Nations Office for the Coordination of Humanitarian Affairs (Imran et al. 2014a). As such, this non-traditional form of information can provide valuable insight into emergencies happening on the ground (Wang and Taylor 2015). However, the scope, generalizability, and direct relevance of social media analysis to state and federal personnel beyond catastrophes is still underdeveloped.

Case studies for the utilization of social media have included analyses of meteorological disasters such as flooding (de Albuquerque et al. 2015), winter storms (Wang et al. 2017) and hurricanes (Wang and Taylor 2014), providing forewarning for the occurrence of earthquakes for cities further from the epicenters (Sakaki et al. 2010), analyzing text sentiment and its use in assessing human mobility after earthquakes (Wang and Taylor 2018), tracking the extent of power outages (Jennex 2012), and terrorism events like the Boston Marathon Bombing (Starbird et al. 2014). These research endeavors have sought to answer the primary questions raised by response organizations, the 5 W's: Where, What, When, Who, and Why (Kropczynski et al. 2018). Research has also sought to rank the relative importance and severity of events, and the confidence with which those events have been identified, to further delineate the relevance of the information to the end user (Imran et al. 2016).

The ultimate goal that crisis informatics researchers have been pursuing is the derivation of actionable data for responders. As detailed above, what qualifies information as "actionable" changes from organization to organization, and so multiple different techniques have been developed within the field or altered from another. Analytical methods in crisis informatics have evolved alongside developments in machine learning, natural language processing (NLP), and geographic information science (GIS). Within our research, we sought to utilize each of these to produce actionable data from social media for GDOT.

2.2 Social Media Analytical Methods

Actionable information is one of the most important yet lacking resources in the middle of a crisis. Understanding who and what is at risk in the shortest time possible, with the most accurate and comprehensive details possible, is the goal for any emergency manager. However, there is often a tradeoff between information timeliness and information accuracy or relevance (Reuter 2018). With social media, searching for needed information in the "firehose" of the data stream can extremely challenging due to the large volumes of data involved. Fortunately, as a direct result of that volume, it has been found to lead to a number of nuanced and surprising insights. Crisis informatics research has primarily focused on automated methods of removing the extraneous, irrelevant information and condensing the relevant into formats that decision-makers in response organizations can use. Within our research, we focused on event detection, cluster analysis/spatial autocorrelation, and sentiment analysis.

Event detection techniques founded on clustering-based approaches often use co-occurrences of keywords for semantic examinations, and thus, are unable to uncover the latent structure of topics underlying the text corpora. For example, SigniTrend, a scalable detection technique developed by Schubert et al. (Schubert et al. 2014) clusters the detected keywords into larger topics after detecting trending word pairs based on their co-occurrences, the significance of the words using hashing technique. GeoBurst (Zhang et al. 2016) identifies candidate events based on geographical and semantic impact between each pair of Tweets, and ranks the candidates according to their spatial and temporal burstiness to extract local events from streams of geotagged Tweets in real time. Probabilistic topic models such as Latent Dirichlet Allocation (LDA) (Chaney and Blei 2012), however, can uncover underlying themes of a document by generating a probabilistic distribution of words under a topic. LDA has been used to identifying thematic content in social networks event detection. For example, Semantic Scan (Nobles et al. 2019) identifies new topics in text streams from Yelp using a contrastive LDA-based topic modeling approach as well as statistical scanning to spatially localize events. Topicsketch (Xie et al. 2016) generates topics based on sketch-based topic modeling using Singular Value Decomposition of word pair frequency matrices or tensor decomposition of word triple frequency matrices integrated with a hashing-based dimension reduction technique to detect bursty topics from Twitter.

LDA was originally developed for large static corpora and has weaknesses in analyzing spatially localized and temporally sequenced data such as social media postings form Twitter (Wang et al. 2012). Current supervised LDA-based event detection methods primarily integrate temporal and semantic dimensions and often require a pre-defined number of topics for both background corpus and foreground topics, or are based on the assumption that each Tweet is only related to one latent topic. In the events of disaster emergency in urban areas, however, it is critical to detect crisis events with respect to associated spatial patterns and geographic dimension as well as the intensity of negative sentiments. A novel general event detection technique (without a known or specific target event) developed by the PI (Wang and Taylor 2019) detects crises in near real-time occurring in specific geographic locations and with unknown characteristics within the context of a larger emergency (e.g., a hurricane) and is the basis of the event detection approach employed in this project. The technique takes both spatial and textual information from social media into consideration, and integrates semantic correlation and the change of intensity of negative sentiment to filter the events.

In terms of that spatial information, the spatial autocorrelation of data is critical in understanding the underlying spatial structure of various features of an emergency event; for example, assessing the spatial heterogeneities of hurricane flood to identify risk hotspots (Sajjad et al. 2020), or examining spatiotemporal community resilience to natural hazards across scales (Cutter and Derakhshan 2018). Evidently, determining the spatial distribution and aggregation of social media data is significant to situational awareness of first responders and critical to emergency response, reduction and prevention.

As for the latter concern, the sentiment analysis (Nasukawa 2003) of the text concerns computational and Natural Language Processing (NLP)-based approaches in identifying and characterizing subjective information (Liu 2012) such as emotions and opinions in textual data. Lexicon-based approaches to sentiment analysis explore the relation between importance of events and sentiment intensity of textual information (Thelwall et al. 2011). Caragea et al. (2014) classified the sentiment of geo-located Tweets into three classes of positive, negative and neutral to determine the spatial distribution of public mood. Identifying the dynamic polarity of social media sentiments (positive, negative or neutral) over the course of natural disasters and emergencies can help first responders to improve situational awareness and crisis management (Beigi et al. 2016). For example, upon classifying and tracking the emotions of affected people using Tweets, fear and anxiety have been recognized as the main emotions after an earthquake in Japan, while calm, and unpleasantness were detected during severe earthquakes (Vo and Collier

2013). Verma et al. (2011) introduced machine learning based sentiment analysis approaches for disaster management. Exploring the collective sentiment of Tweets containing the word ‘‘climate’’, Cody et al. (2015) discovered an association between change of happiness and climate-change-related topics. Neppalli et al. (2017) found unique spatial Tweeting patterns including increasing clustering tendency for both positive and negative sentiment during Hurricane Sandy. They observed closer proximity to the Hurricane for negative sentiment clusters, which was dispersed in the days following the Hurricane's maximum impact. Bai and Yu (2016) discovered aftershocks and potential public crises post Ya'an earthquake and implementing an incident monitoring approach based on crowd negative sentiment of Chinese short blogs from Weibo. In this project, we further classify sentiment of social media postings during different events of emergency through examining both spatial and temporal dynamics of the sentiment.

2.3 Social Media Applicability to GDOT

2.3.1 Existing UVI for Traffic and Road Hazards

Waze, one form of UVI, is currently being utilized in GDOT's existing Advanced Traffic Management System (ATMS.) Waze contains crowdsourced smartphone application data that is reported by drivers from the road. The data comes with a geolocation and a time stamp, both of which tend to be fairly accurate within a few minutes and a radius of approximately one mile (Amin-Naseri 2018). The information sent through the application also contains a hazard type and a descriptor. Often, user reports come into the application describing the same event, so an ATMS is used to reconcile and congregate the event data through clustering and similarity-matching the reports. Waze has been found to record and report a substantial amount of traffic incidents first among detection systems; for instance, Iowa's ATMS found that Waze initially recorded 13.4%

of congestion and crashes in 2018. It was, however, found to be less reliable during the nighttime, when fewer people were on the road (Amin-Naseri 2018).

Although Twitter has also been tested for usage in identifying congestion and traffic (Li 2012, D'Andrea 2015), filtering the Tweets for relevant text and accurate geolocation data has been difficult even as it is constantly improving (Gu 2016).

2.3.2 Transportation Reports Related to this Research

Related Transport Research International Documentation (TRID) studies in this area, including the records from Transportation Research Board (TRB) Transportation Research Information Services (TRIS), includes several TRB projects. For example, *Social Media Practices in Traffic Safety* (RiP 01629715) explores how State Highway Offices (SHOs) may use social media to promote safety through qualitative and quantitative scans of social media platforms as well as interviews with nine States. *Social Media Guidebook for Emergency Management* (RiP 01642763) developes a guidebook to help airports leverage social media for emergency management and crisis communication. *Utilize Crowd-Sourced Data and Machine Learning Technology to Enhance Planning for Transportation Resilience to Flooding* (RiP 01674188) extends from this to develop a decision support system (DSS) that combines non-traditional, crowdsourced big-data with traditional data to enhance transportation readiness for quick response decisions in urban flooding. And *Emergency Management Agencies: Pilot for a Crisis Communication Analysis Assessment Test* (RiP 01460030) examines issues such as attitudes toward the use of social media, public information, and Emergency Management Agency (EMA) web sites in emergency management. There are also a few related TRB projects recently completed. These include; *Improving Emergency Preparedness and Crisis Management Capabilities in Transportation* (RP 01467319) explored whether and how one significant functional area—surface transportation—developed the capabilities to effectively fulfill

U.S. commitments for developing a comprehensive, integrated emergency management system, *Modeling Disaster Operations from an Interdisciplinary Perspective in the New York-New Jersey Area* (RP 01566476) used social media in addition to survey data to understand overall demand, destination type choice, and route choice decisions in the aftermath of Hurricane Irene, and *Big Data During Crisis: Lessons from Hurricane Irene* (RP 01556674) characterized the potential of big data from social networks and Natural Language Processing (NLP) methods in creating actionable information in a crisis in the event of Hurricane Irma. These projects have established the initial steps towards integrating social media information and communication data into various dimensions of DOT's crisis and emergency management strategies; however, the need for integrating social and community data into current systems for (near) real-time crisis communication and response remains unfulfilled.

2.3.3 Georgia-Specific Crises

Weather patterns are becoming increasingly erratic and harder to predict across the globe, and Georgia has not been spared the impacts of increasingly intense and frequent extreme weather events (Noy 2016). Although in Georgia, the city of Savannah has experienced numerous hurricanes and tropical storms, the city of Atlanta which is hundreds of miles inland was placed under its first Tropical Storm warning due to Hurricane Irma in 2017. Hurricane Irma resulted in a total of three fatalities and inland evacuation orders for 540,000 coastal residents. Notwithstanding the authorities' orders and warnings, many residents failed to follow the procedures, which exacerbates a range of quickly-developing crises, including: stranded residents in rising storm surge, shortages of gas, water and food, power outages, infrastructure failures, fires, traffic jams, traffic incidents, evacuation barriers, looting, and other crises that require rapid response and emergency assistance. Under such circumstances, where the crisis is too time critical for a 911 call to be effective (e.g., people cannot make phone calls, emergency telephone hotlines are jammed, or the emergency responders are unable to assess the relative gravity of one crisis over another), residents are increasingly turning to social media for help, even posting their full addresses in desperation.

Additionally, in terms of meteorological events, Atlanta is particularly challenged by snow and ice events. Previous winter storm events have shut the city down, generated large amounts of panic, and trapped citizens on icy roads for hours at a time. As most Atlanta citizens have relatively little experience driving on icy roads and understanding the dangers associated with them, it can be critical for GDOT to address icy roads before a driver can encounter the hazard.

2.4 Summary

Ultimately, social media is currently an untapped resource with respect to its ability to identify on-road emergencies, provide additional information such as severity and human impact for existing incidents, and its generalizability for many different kinds of events. In detecting and responding to emergencies, any information that can increase reaction time or level of preparedness is valuable, and social media and other forms of UVI have been repeatedly shown to be able to provide it. For Georgia, and for Atlanta in particular, there is a particular need for improving the identification of specific risks during ice/snow events and for flooding events. These are both events with widespread impact with geographic pockets of extreme severity and risk, and thus two of the most suitable applications for the widespread network of “human sensors” that can be tapped through social media.

3 DATA COLLECTION

3.1 Overview of Data Collection Methods

At the start of the project, we focused on data collection, processing, storage details, and filtration techniques. Our lab set up two Twitter Streaming Application Programming Interfaces (APIs) during the course of the project. One was designed to pull representative Tweets that contained a geolocation tag, which includes a set of latitude and longitude coordinates, and one was designed to pull representative Tweets that contained one of a set of keywords. As described earlier, we had identified in our proposal that ice/snow and flooding events would be ideal test cases for the potential of social media application in GDOT's response activities. The keywords for each test case set varied based on what type of event, and we deployed these keywords based on the time of year (ice/snow in winter and flooding in spring and summer). We collected Tweets that contained the words "ice" or "icy" during the winter, and we changed our tracking stream to collect Tweets that contain the words "flood", "flooded", or "flooding" for the summer in the state of Georgia.

While storing the data, we determined that we would need to set up a more reliable data repository that was more secure, had more storage space, and would be more resilient to power and network outages. We established a Structured Query Language (SQL) Server at the School of Civil and Environmental Engineering for database management of the project, and identified a method of porting our retrospective data (originally stored in a Network Attached Storage (NAS)) into a new SQL database in order to increase the speed of data transfer and better facilitate data access, processing, documentation, analysis, and future integration with GDOT systems. Additionally, to increase our data robustness, we updated our download compression format from pickle data format to JavaScript Object Notation (JSON) because JSON is a more standardized, language-independent file format, a more lightweight and faster format, and does not have the known security risks associated with the .pickle format. We then created a new framework for the storage of the new data in a database on the server. Error handling (for dropped connections and data formatting errors) for the streaming process was made substantially more robust.

Twitter data has notoriously few context clues for the subject matter of its Tweets due to their relatively short length (280 characters). As such, it was necessary for the team to develop STOP words to remove Tweets irrelevant to the subject of interest. As keywords are a signal to the search algorithm that the topic may be of interest to the research, so STOP words are words that signal to the algorithm that the topic is *not* of interest. We identified the STOP words necessary for non-relevant data removal (in the case of a word such as 'ice', these STOP words include 'cream', to remove posts about 'ice cream', and 'immigration', to remove Tweets concerned with 'Immigration Control and Enforcement'.)

Following the shift in data format to SQL, the shift in data storage from an NAS to a server, and the collection of Tweets posted in Georgia concerned with snow and ice across the winter of 2018/2019, we found that there were no major snow/ice events that winter. Although this was certainly good for Atlanta citizens, we were forced to switch to a secondary data source: using historic data from the snow/ice events in the winter of 2017/2018. Additionally, because the original timeline of the project had been shifted forward several months, the project would need to be completed prior to our ability to collect flooding events in the spring. As such, because we were using data from 2017 for the snow/ice events, we agreed with our GDOT Technical Implementation Team to investigate flooding in Georgia during Hurricane Irma (also 2017) for our flood test cases. Data was collected and processed using the natural language processing method outlined above for the week of Hurricane Irma (for flooding) and for the three ice events that occurred in the 2017-2018 winter.

4 GEO-TOPIC DETECTION

4.1 Identification of Necessary Topics

Prior to beginning our investigation of the data accumulated above, we met with GDOT personnel to discuss the types of topics they would be interested in monitoring during both snow/ice and flooding events. On review of the Tweets gathered from the test cases and further discussion with the GDOT team regarding situations of concern, the team decided to use the following keywords search for ice events: "ice", "icy", "frozen", "power", and "road"; for flooding events, the words used were: “flood”, “flooding”, “heavy rain”, “road”, “power”, or “Irma”. Stop words—words that prevent a Tweet from being included for analysis—were used to prevent false positives (such as someone Tweeting about “ice cream”). Semantic notation were also utilized; for “ice”, regular expressions were used to prevent words such as “service” and “advice” from being incorporated. The results of these national language processing methods are discussed below.

4.2 Ice Events in Atlanta

As there were no significant ice events this past winter, we used data from three ice events that occurred in the 2017-2018 winter. Ice was observed on the roads of Atlanta on December 9th, 2017; the night of January 5th, 2018 and on January 16th and 17th, 2018. Within the historic data that was available to us, we pulled Tweets geolocated within Georgia and a 30-mile buffer around the state line from a day before and after the icing event. The 30-mile buffer was included on the advisement of the GDOT Technical Implementation Team, which informed us that disaster events affecting neighboring state roads often required GDOT personnel to assist. An ice event on the other side of the border can back up traffic into Georgia, and often Georgia resources are allocated to neighboring troubled states.

Twitter data were identified that were geolocated in the state of Georgia and a 30-mile buffer for the following days: Dec. 8-10, 2017; Jan 5-6, 2018; and Jan 15-18, 2018. The Tweet counts following the filtration methods outlined above are included in **Table 4-1**.

Table 4-1. Tweet counts by day for ice events in Atlanta, GA (2017-2018).

	8-Dec	9-Dec	10-Dec	5-Jan	6-Jan	15-Jan	16-Jan	17-Jan	18-Jan
Counts	59	85	31	24	35	14	33	94	61

As with any sensor network, the amount of information and the clarity of that information is dependent on the number of sensors available in any given location. For Twitter, the number of Tweets that can be expected in an area experiencing a negative event is dependent on the number of Twitter users in that area and the propensity and ability of that user to Tweet about a given event. In order to identify which areas have a higher number of Twitter users with the propensity and ability to Tweet about ice events, we generated a map of the counties that contained Tweets during the day with the largest number of Tweets recorded for an ice event, January 17th. That map is depicted in **Figure 4-1**.

Following the acquisition of relevant data, we analyzed the focus of the information and its relevance to GDOT's activities. Each of the filtered Tweets were assigned a primary keyword based on the topic that was used as the subject of the text. For example, consider this Tweet posted on January 17th, 2018:

> "Snow day 2018 - it's COLD - the roads in my neighborhood are pretty icy and I am really looking forward to Sunday when it's going to be 60 degrees here! #snowday #snowintheatl #digitallearningday."

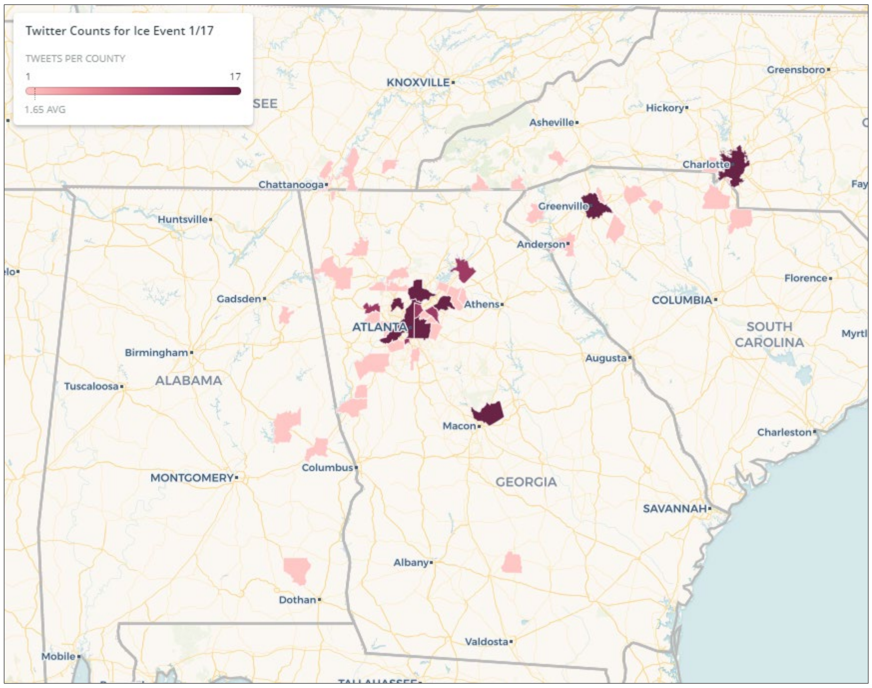


Figure 4-1. Distribution of Tweets during an ice event with the largest number of Tweets recorded across the counties in the state of Georgia (January 17, 2018).

The primary concern in this Tweet is the status of the roads. However, this Tweet contains two of the listed keywords. In cases where Tweets contained more than one of the keywords, secondary keywords were assigned. The roads, as the subject of the Tweet, are assigned the adjective “icy”. Because the second keyword is not the subject but is still relevant to the Tweet’s content, it was deemed to be a secondary concern. Although the focus of the spatial and temporal analyses included in this research focused on the primary keywords, the juxtaposition of primary and secondary keywords were analyzed as well. Within this dataset, the most common secondary keyword was “icy”, and the second-most common was “frozen”.

The following in-depth results in this report focus on the ice event recorded on January 17th, 2018, as the greatest number of relevant Tweets were recorded on that day, and because it was the second day of an ice event that lasted two days. The numbers and percentages of Tweets with different primary keywords are represented in the pie chart in **Figure 4-2**.

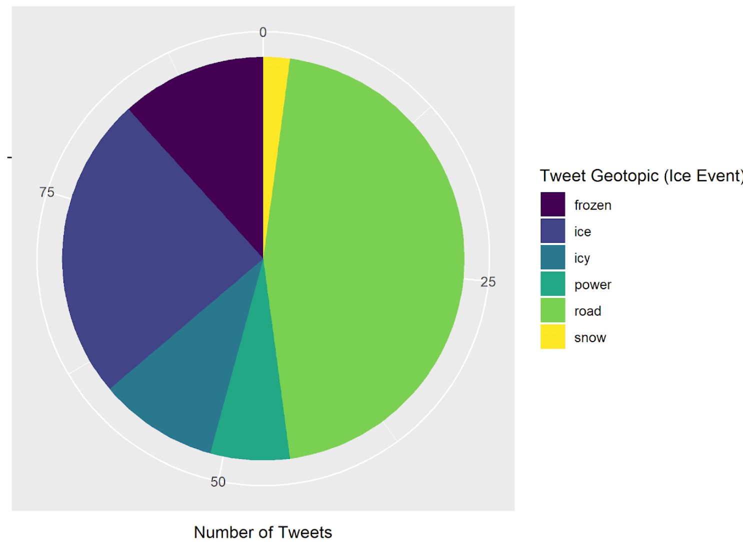


Figure 4-2. Distribution of Tweets by geo-topic during an ice event in the state of Georgia (January 17, 2018).

With respect to relevance to GDOT response personnel, the GDOT Technical Implementation Team informed us that GDOT was primarily concerned with the expressways, major roads, and state highways. The team informed us that alerts outside of their jurisdiction would be forwarded to the presiding districts' authorities. With that in mind, we additionally analyzed the proximity of the filtered Tweets to the GDOT roadways. We determined the length of state highways that were within half of a mile of a relevant Tweet, which are pictured in **Figure 4-3**.

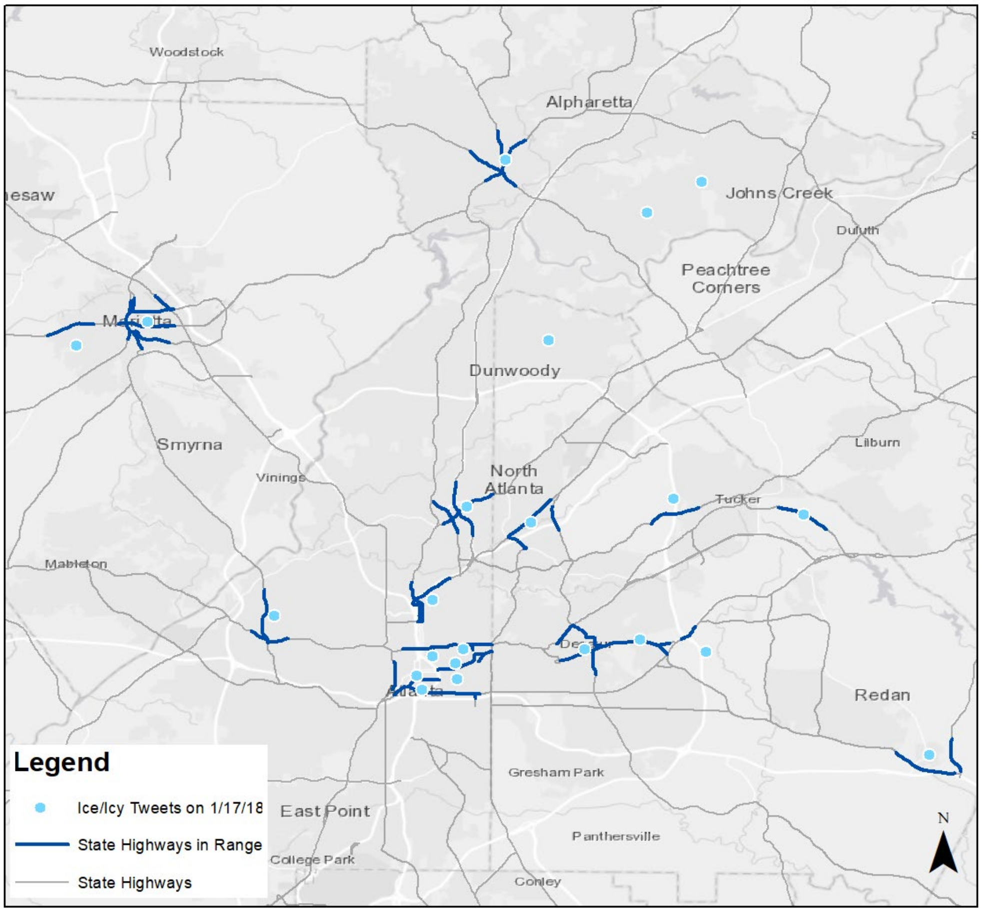


Figure 4-3. Distribution of a total of 100 miles state highways within ½ mile proximity of Tweets during an ice event in the state of Georgia (January 17, 2018).

This is equivalent to a bit more than 100 miles of state highway. In terms of the proximity of the Tweets to state highways, we determined that, of the 62 Tweets within the state of Georgia (out of 94 within the state of Georgia and the 30-mile buffer), 36 were within 0.5 miles of state highways; 21 were within that range for major roads, and 20 were within that range for expressways. If that search radius is expanded, we find that 51 Tweets were within 1 mile of state highways; 50 were within 1 mile of major roads, and 28 were within 1 mile of an expressway.

Additionally, it is necessary to further understand when this information is available. Previous research identified information gaps in UVI during the night (i.e. from 12am to 6am), when fewer people are on the roads (Amin-Naseri et al. 2018). In order to assess when this data might be most viable, the density of information availability at each hour was plotted in the graph in **Figure 4-4**.

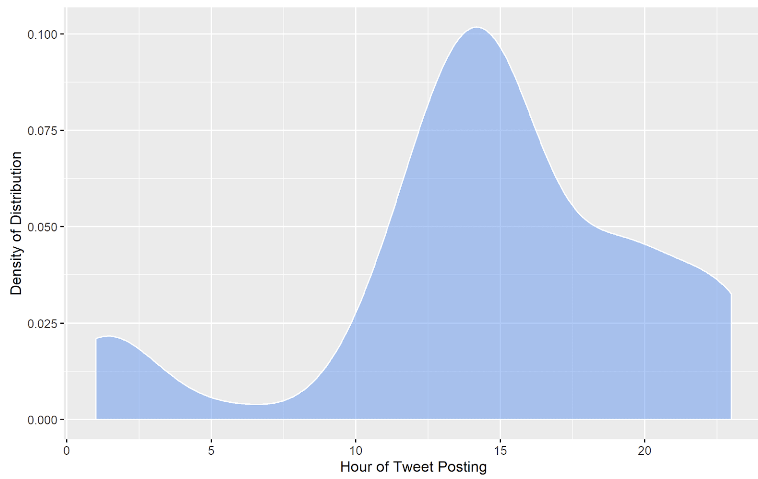


Figure 4-4. Hour of Tweet posting distribution during an ice event in the state of Georgia (January 17, 2018).

From the above analysis, it is apparent that the Twitter data available for ice events mirrors the data collected to analyze traffic data. Twitter data for ice events appears to increase as people are beginning to leave for work and finding the streets too icy to drive. The fewest Tweets are available from 2am to 6am, and the most Tweets are available from 12pm to 5pm. Although the lack of data when the ice is likely forming is concerning, these results do indicate that the most data is available when the most people are looking at the roads, and thus, possibly, when the most people are in danger from driving in wintry conditions.

4.3 Flood Events in Atlanta

For historical cases of flooding, we utilized a week during which we knew excessive amounts of flooding occurred: the week that Hurricane (then Tropical Storm) Irma impacted Georgia. Hurricane Irma resulted in a total of three fatalities and inland evacuation orders for 540,000 coastal residents. Notwithstanding the authorities' orders and warnings, many residents failed to follow the procedures, which exacerbates a range of newly-developing crises, including: stranded residents in rising storm surge, shortages of gas, water and food, power outages, infrastructure failures, fires, traffic jams, traffic incidents, evacuation barriers, looting, and other crises that require rapid response and emergency assistance. As such, it was deemed to be a prime test case for the reaction of Georgia citizens to almost unprecedented amounts of rain and flooding.

In a manner similar to the collection of the ice event data, Twitter data geolocated in the state of Georgia and a 30-mile buffer around the state line was collected for the week of Hurricane Irma (Sept. 10-17, 2017) to provide context for the flooding case. The Tweet counts following our filtration methods are presented in **Table 4-2**.

Table 4-2. Tweet counts by day during Hurricane Irma in the state of Georgia in addition to a 30-mile buffer around the state boarder (Sept. 10-17, 2017).

	10-Sep	11-Sep	12-Sep	13-Sep	14-Sep	15-Sep	16-Sep	17-Sep
Counts	202	319	144	71	52	46	49	27

We observed these Tweets in the counties displayed in **Figure 4-5**.

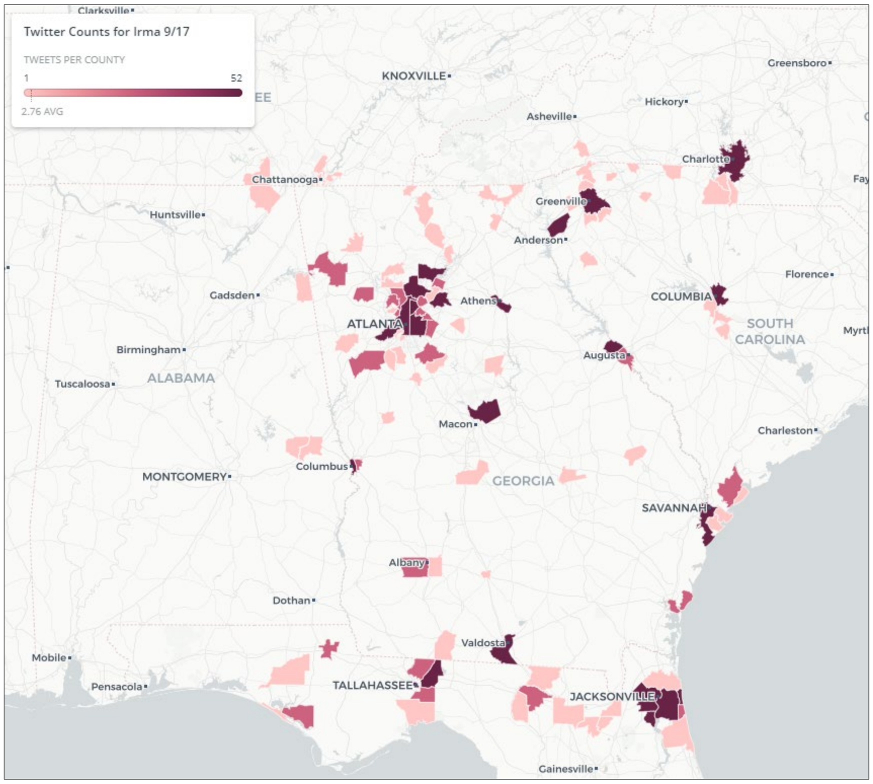


Figure 4-5. Distribution of Tweets during Hurricane Irma in the state of Georgia (Sept. 10-17, 2017).

This map provides us with two distinct conclusions. The first is that we observe very similar counties providing the most number of Tweets, i.e. most of the counties within the Greater Metropolitan Area of Atlanta consistently produce Tweets that are caught within the automated filters. Secondarily, we see far more areas outside of Atlanta producing data for Tropical Storm Irma than for the ice/icy events. Part of this is because of the increased activity on the coast. For the ice/icy events, much of southern Georgia did not experience persistent ice. Tropical Storm Irma, however, impacted the coast with more strength, and had a much broader geographical impact on the state. There were additionally more preparatory Tweets anticipating hazards that were posted prior to flooding events than there were for the ice/icy events.

In reviewing and comparing the relevance of the posted Tweets in both cases, we determined that there is a higher percentage of Tweets depicting an actual roadway hazard outside of the Greater Metropolitan Atlanta Area than there are within the city limits. This is not to say that there is less information available through Twitter within Atlanta, simply that there are more irrelevant Tweets posted alongside the relevant ones. This could indicate a need for a separate filtration technique for Tweets posted within the city itself.

Following our review of the data relevance, we turned towards the distribution of the Twitter topics. The topics of these Tweets were distributed are presented in **Figure 4-6**.

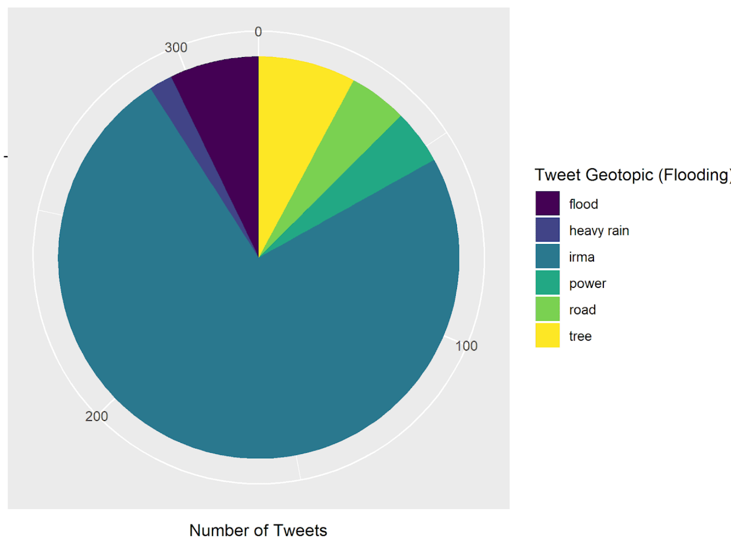
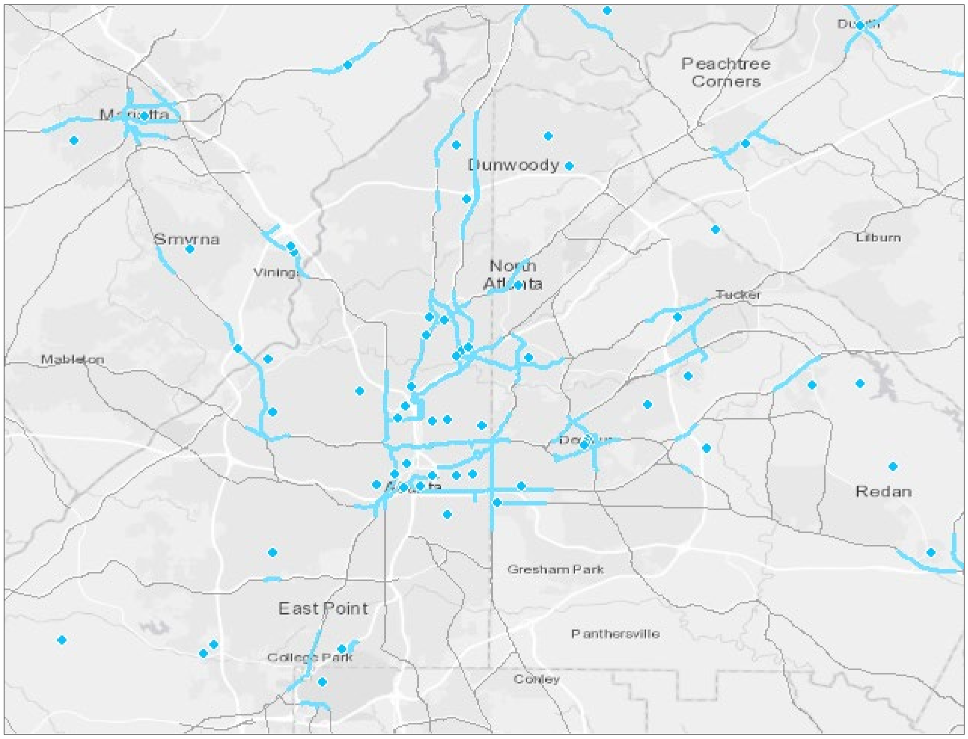


Figure 4-6. Distribution of Tweets by geo-topic during Hurricane Irma in the state of Georgia (Sept. 10-17, 2017).

The prevalence of "Irma" Tweets indicates the need for additional filtration methods when utilizing the name of a storm, as many of the posts did not indicate a specific and relevant hazard. In fact, the vast majority of relevant Tweets containing the word "Irma" were accompanied by pictures that showed some form of damage caused by the storm. The text itself was less informative than the accompanying pictures, indicating that additional work utilizing a form of image-recognition technology or a human-in-the-loop would be necessary to derive the maximum amount of information from the data filtered through our process. The most common secondary keyword was road, followed by tree. The most common pairing, “road” and “tree”, resulted from posts about trees fallen across the road or across power lines.

Similar to the work performed to identify GDOT-relevant information performed for the ice/icy event Tweets, we determined the miles of state highways located within 0.5 miles of a Tweet. As such, the state highway sections located within less than 0.5 miles of one of the filtered Tweets are depicted in blue in **Figure 4-7**. We additionally derived the code necessary to receive a list of the road names located within a specific boundary distance in order to facilitate the work performed late in this report to specify a visualization system.

The highlighted roadway sections are equivalent to approximately 315 miles of state highways. In terms of the proximity of the Tweets to state highways, we determined that, of 167 Tweets within the state of Georgia (out of 276 within the state of Georgia and the 30-mile buffer), 108 were within 0.5 miles of state highways; 90 were within that range for major roads, and 47 were within that range for expressways. If that search radius is expanded, we find that 135 Tweets were within 1 mile of state highways; 107 were within 1 mile of major roads, and 69 were within 1 mile of an expressway.


Peachtree
Corners
Dunwoody
Smyrna
Vinings
North
Atlanta
Lilburn
Tucker
Mableton
Redan
Gresham Park
East Point
Panthersville
College Park
Conley

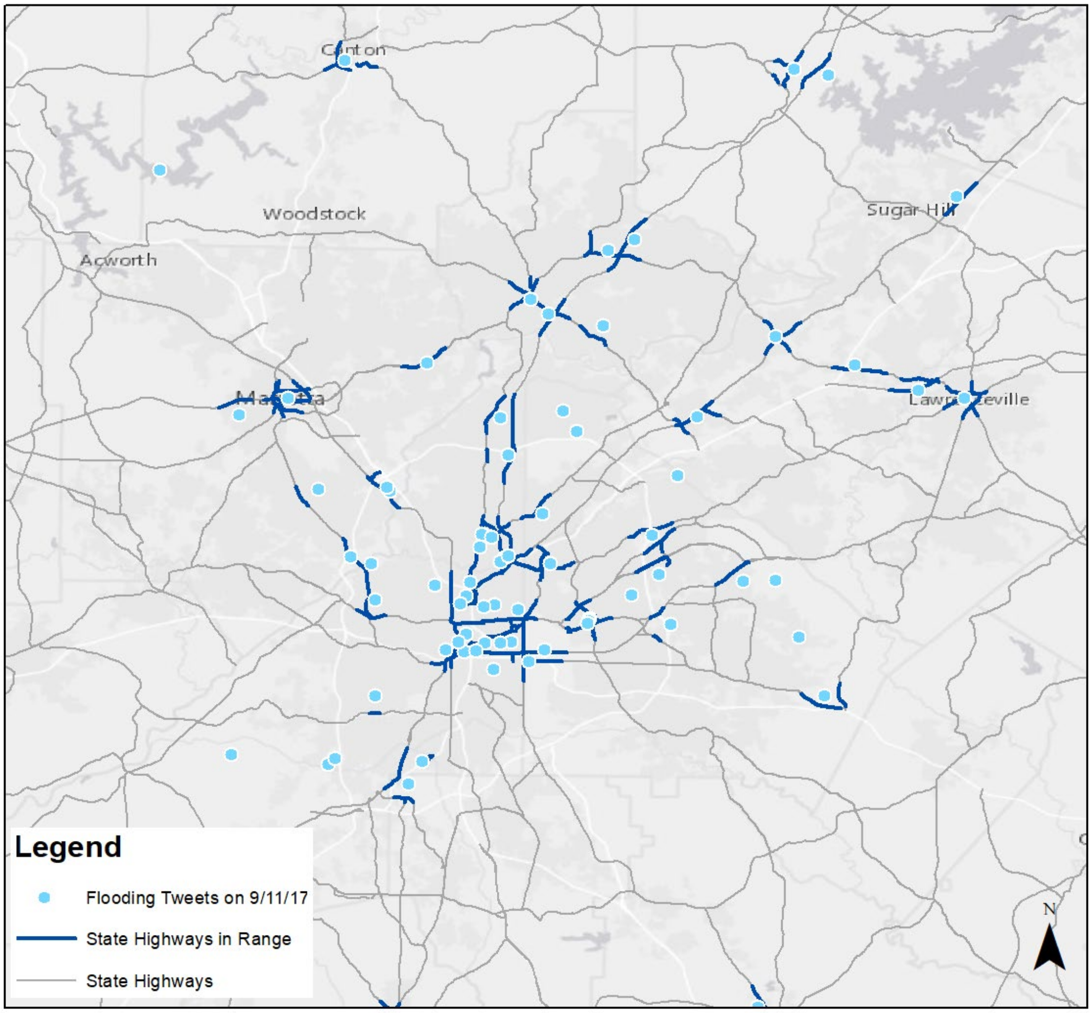


Figure 4-7. Distribution of a total of 315 miles state highways within ½ mile proximity of Tweets during Hurricane Irma's landfall in the state of Georgia (Sept. 11, 2017).

The proximal highway section length for Irma is more than triple the amount of mileage located within range of the ice event Tweets. This correlates strongly with the number of Tweets observed for each event, which is a strong indicator that the "covered" roadway amounts are more strongly influenced by the number of people affected by an event and thus their propensity to Tweet about it than simply the number of Twitter users in any given area, especially when using topic-filtration. This additionally promotes the utilization of additional filters within the city of

Atlanta, for although more roads are located proximal to the increased numbers of Tweets, there may be pockets of roads that are not covered by people with the propensity to Tweet information relevant to GDOT.

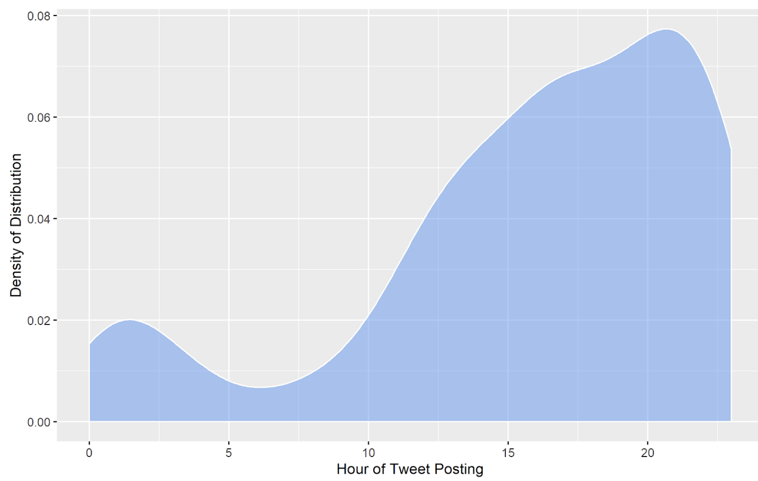


Figure 4-8. Hour of Tweet posting distribution during Hurricane Irma in the state of Georgia (Sept. 10-17, 2017).

Finally, doing a similar temporal analysis to the one performed for the winter storm, we see a similar temporal distribution of Twitter data for flooding events as we saw for ice events; however, the decreased information available at night is more pronounced. The distribution is presented in **Figure 4-8**.

4.4 Interactive Visualization

As part of our work towards ultimately specifying an interactive visualization and tracking system, we developed an interactive online map for GDOT personnel to use in providing us feedback. The interactive map contains the data from both of the above test cases and test days and can be accessed here: https://maphub.net/rsamuels3/GDOT_Test_Cases. The Ice and Flood cases can be selected on the website by toggling the eye icon adjacent to the Ice or Irma layers to

On, and the individual Tweets are depicted by differently colored dots. The colors correspond to the primary keyword associated with each Tweet. The platform allows visual inspection of Tweets by left-clicking on the individual points. After selecting a point, a pop-up will appear depicting the text of the Tweet, the time it was posted, and the associated keywords for reference. An image of the interactive map is depicted in **Figure 4-9**.

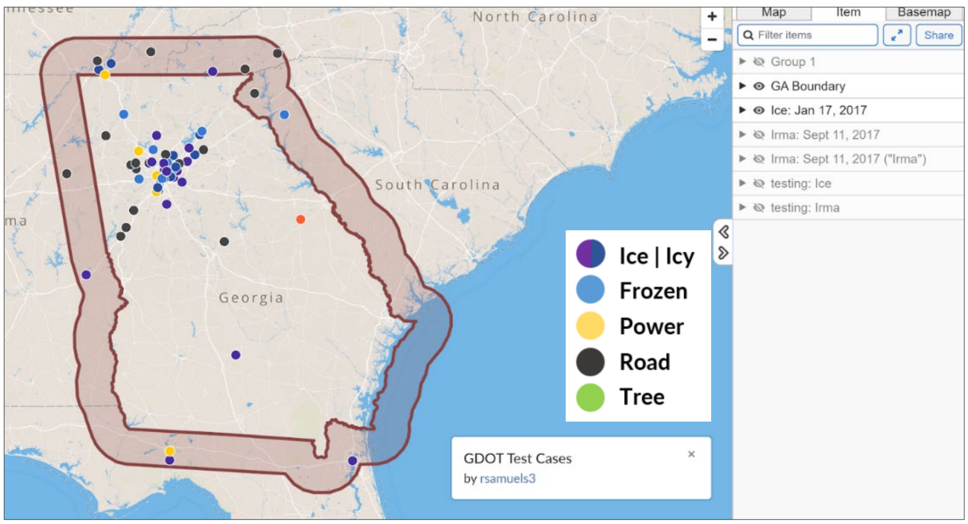


Figure 4-9. Interactive map of social media activities during 2017-18 ice events and Hurricane Irma in the state of Georgia. The colors of the dots correspond with the primary keyword associated with each Tweet. The single orange dot on the center eastern part of the state can be clicked to reveal the legend shown in this figure at the right.

Examples of how the Tweet text is portrayed within the interactive map are portrayed below in **Figure 4-10** and **4-11**. The first is an example Tweet from the ice event and the second is an example Tweet posted during Hurricane Irma.

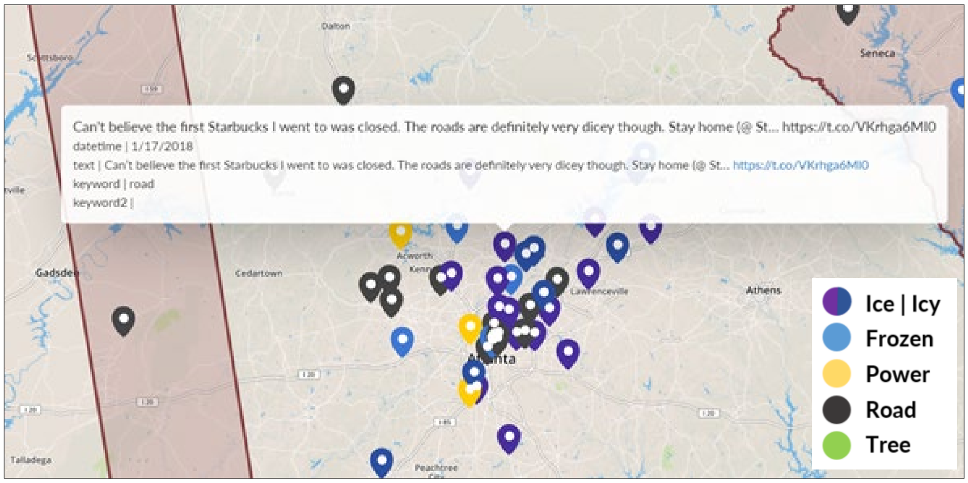


Figure 4-10. Example of Twitter activity during 2017-18 ice events in the state of Georgia. The title of each Tweet is the most relevant keyword from the text of the Tweet; the other included attributes are the date, the full text, and the primary and, if relevant, secondary keywords.

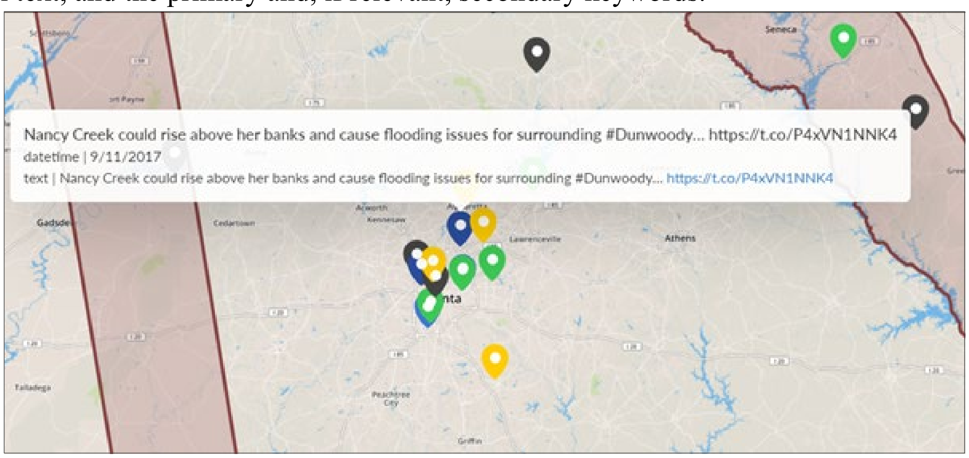


Figure 4-11. Example of Twitter activity during Hurricane Irma in the state of Georgia. The title of each Tweet is the most relevant keyword from the text of the Tweet; the other included attributes are the date, the full text, and the primary and, if relevant, secondary keywords.

Following the creation of this interactive map containing the filtered Twitter data and a meeting with the GDOT Technical Implementation Team, we shifted our focus towards seeking additional insights from the collected data and the development of additional filtration and tagging methods. These methods were mainly centered on analyzing the relative importance of the filtered data for GDOT use and are described in the following chapter.

5 GEO-TOPIC RANKING

On obtaining the relevant Tweets above, we turned towards ranking the relative importance of geo-topics as outlined in our proposal. In understanding the relative importance of the geo-topics, we analyzed their spatiotemporal clustering, the semantic topics, and how the crisis dynamically evolves over time. The focus of the analytics utilized in this research were the spatial proximity of the Tweets, the sentiment recorded within the Tweet text, and the temporal evolution and availability of the data. These analytics were applied to the two text cases, and the results for each case study are detailed below.

5.1 Ice Events in Atlanta

Following the identification that the Tweets appeared to be more clustered in the city of Atlanta than in the remainder of the state, it was necessary to determine the extent of this spatial clustering in metropolitan areas. We utilized the Nearest Neighbor Ratio to determine the extent of the likelihood of the clustering of the Twitter data. For this analysis, a more negative z-score indicates a higher degree of clustering. The p-value indicates the certainty with which the z-score was determined. The results of the analysis are depicted in **Figure 5-1**.

The results indicate that the Tweets are clustered extremely non-randomly, and that, as expected, the clusters of Tweets correlate strongly with urban centers. We additionally determined that the average distance between Tweets was 3.1 miles, although the distribution of these distances was heavily skewed to the right and strongly varied between city centers. Next, in terms of geo-topic detection, it was necessary to understand whether these clusters were also influenced by the type of event that was described in the Tweet. We utilized this analysis to understand whether or not clusters of Tweets could be expected to appear around a single definitive event that would produce the same topics within a short geographic distance. The Moran's Index results for the ice event Tweets, utilizing the weight of primary keyword and the weight of the sentiment attribute, are shown in **Figure 5-2** (de Jong et al. 1984).

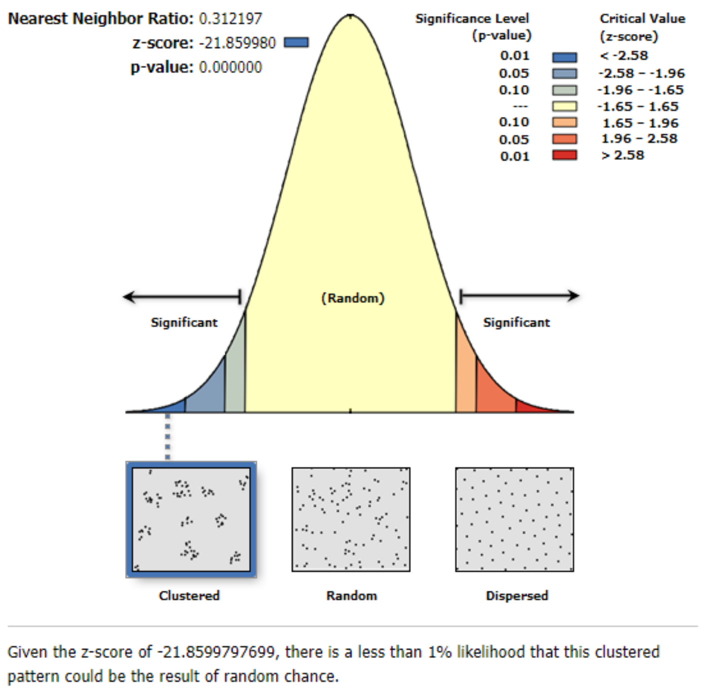


Figure 5-1. Nearest Neighbor Ratio test results for ice events in Atlanta, GA: negative z-score indicates a statistically significant high degree of clustering.

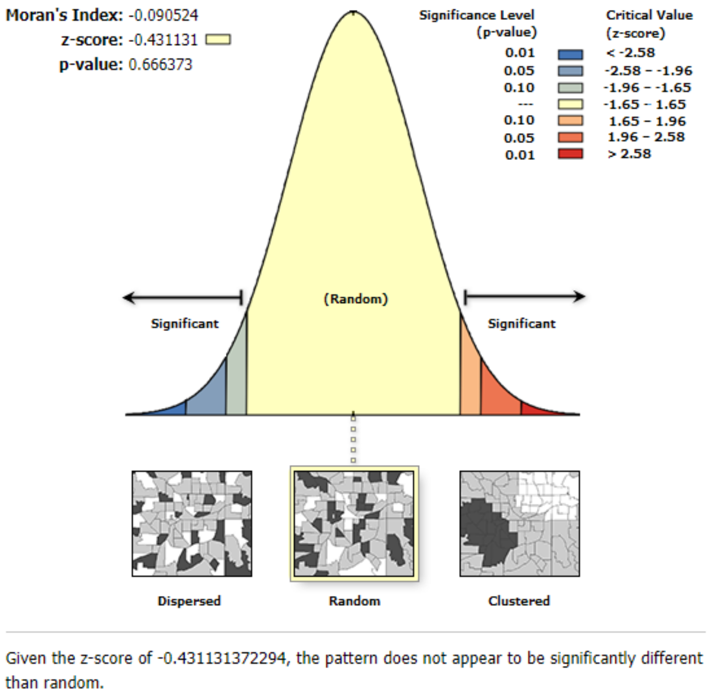


Figure 5-2. Moran's I test results for ice events in Atlanta, GA weighted using sentiment. Tweets are not clustered and randomly distributed in space.

In this case, the analysis determined that Tweets are not clustered in space with respect to individual keywords. This likely indicates that each Twitter user is reporting an event that is unique to them, or uniquely seen by them. A review of the pictures taken by users in relatively close geographic distances confirms that the subject matter of the event details varies, even within close proximity. The second confounding factor for this analysis is that similar events—a road covered in ice, a power outage—were occurring broadly across the state. Because people in the city of Atlanta were posting about the same topics as people in Douglasville and in Athens, the analysis sees the topics covering a broad geographic range and therefore does not indicate the topics as clustered.

For this research, this result may indicate that geo-topic clusters of Tweets are unlikely to be able to indicate a midpoint of an event occurrence, and that the Tweets would best be utilized as individual points of information on a broad scale. Following the spatial clustering analysis, the next aspect of geo-topic ranking that we explored was a sentiment analysis of the Tweets. We utilized AFINN (Nielsen 2011), a lexicon-based sentiment analysis library for sentiment analysis of Twitter microblogs including 2477 words and valence value range of (-5 sad to +5 happy), to automatically determine the sentiment of each individual Tweet. We additionally augmented this analysis using the Natural Language Toolkit (NLTK) sentiment analysis (Bird et al. 2009), which produces both what NLTK terms a polarity (i.e., a sentiment) and a subjectivity metric. The distribution of these Tweets for the ice event test case is depicted in **Figure 5-3**.

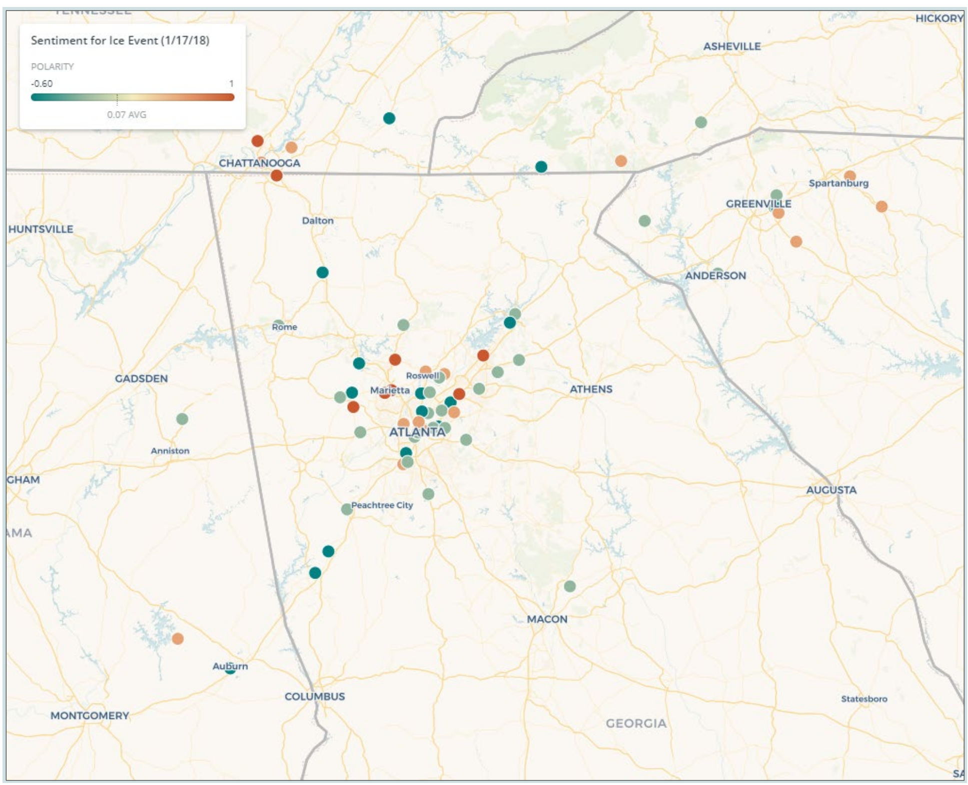


Figure 5-3. Sentiment of Tweets during ice events in Atlanta, GA.

One of the reasons for the usage of two different sentiment analyses was that we needed the ability to check the accuracy of both. After manually labelling the sentiment of a range of Tweets and comparing the results for both the AFINN and NLTK metrics, we ultimately utilized the NLTK for analyses moving forward. An example of a correctly-classified negative Tweet is depicted as **Figure 5-4**.

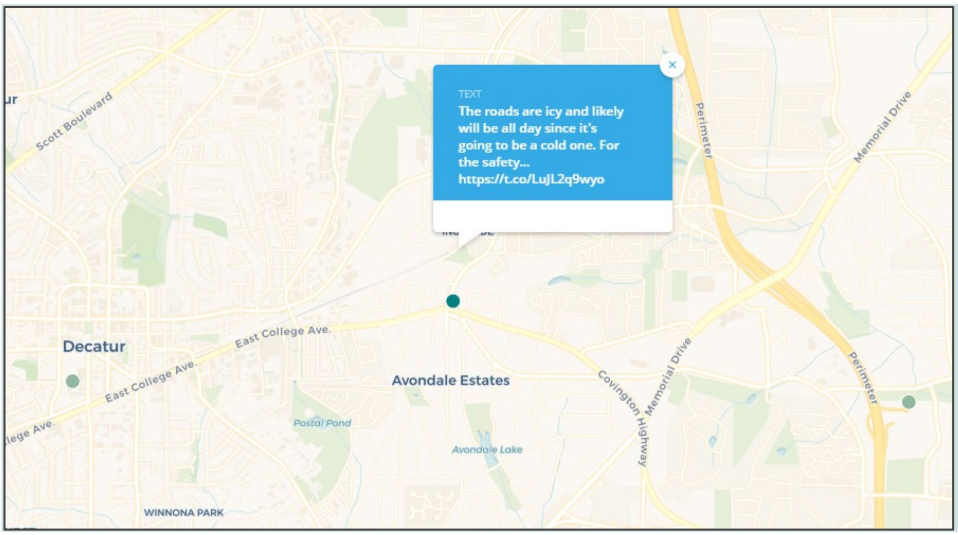


Figure 5-4. Example Tweets identified with negative sentiment during ice events in Atlanta, GA.

Additionally, the ranking of geotopics and Tweets is dependent on the combination of both topic-relevance and sentiment. The distribution of sentiments for each of the identified keyword topics is depicted in **Figure 5-5**.

The results of this analysis inform the usage and utilization of the selected keywords and topics of concern. The peaks of negative sentiment displayed by “frozen”, “ice”, and “icy” occur in Tweets focused on dangers. Therefore, the combination of those keywords and negative sentiment indicate a strongly relevant Tweet. Words such as “road” appeared approximately equally in Tweets containing a happy sentiment about not needing to go to work or school and those concerned about conditions. However, warnings about icy roads contained positive words such as “stay safe!”, which increased the sentiment of the Tweets. As such, sentiment may not be a good differentiator for Tweets focused on the roads, and additional filtration systems or human judgment may be necessary for the interpretation of those Tweets. The prevalence of positive sentiment with the word “frozen” indicates that it may not be useful as a keyword in identifying relevant information about ice events.

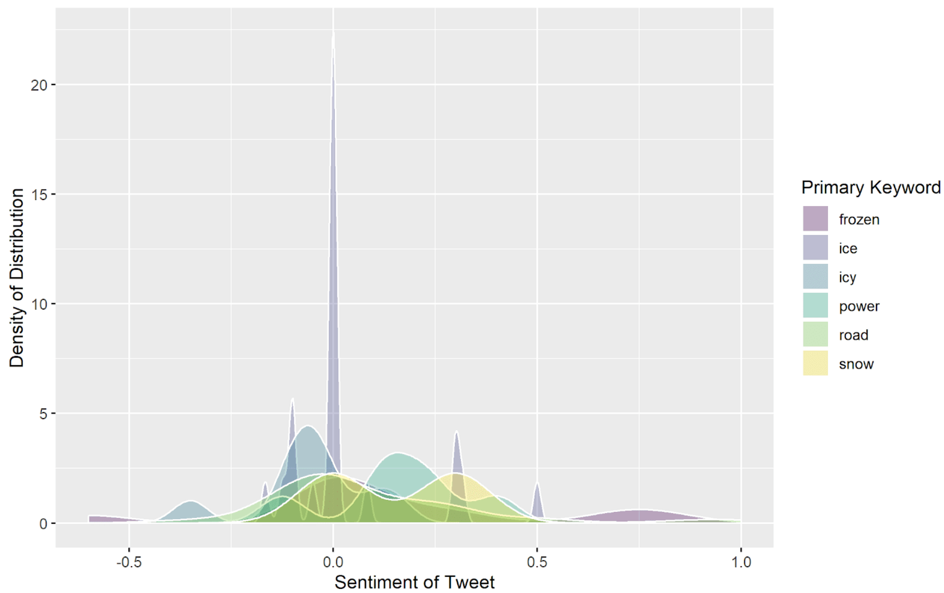


Figure 5-5. Tweet sentiment distribution by geo-topic during ice events in Atlanta, GA.

Finally, the temporal distributions of the Tweets were analyzed with respect to when different keywords were posted. Those distributions are shown below in **Figure 5-6**.

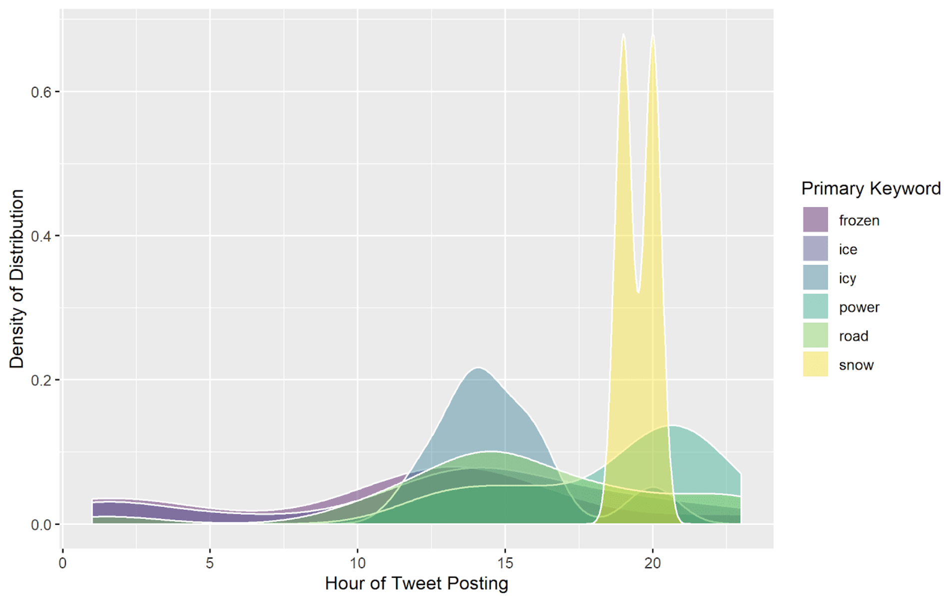


Figure 5-6. Hour of Tweet posting distribution by geo-topic during ice events in Atlanta, GA.

The spikes in Tweets about snow later in the day coincide with when snow was falling. This may be useful as a real-time indicator of weather patterns from sensors in addition to those utilized by the United States Geological Survey (USGS) and the National Oceanic and Atmospheric Association (NOAA). The peaks in the late afternoon correspond to the peaks observed in general, non-event Twitter usage; however, the excessive peak in 'icy' make indicate a time in which more people were considering driving, and the late-evening peak in 'power' is a likely indicator of when people lost power. As such, these temporal bursts may be additionally useful in predicting when more people are on icy roads and thus when more road-related accidents may occur or when widespread power outages are occurring.

5.2 Flood Events in Atlanta

The above analyses were repeated for the flood events generated by Hurricane Irma. The spatial clustering of these Tweets was greater than that for the ice events (z-score of -32.1 with a significant p-value). This is mostly likely because we observed additional clusters in new metropolitan areas, such as Savannah, that did not have Tweets about ice or snow. This confirms our previous hypothesis that the increased number of Tweets was likely due to an increased number of affected citizens, and that the geographic clustering of Tweets is most likely due to urban centers than to clusters of topics.

As before, we determined the sentiment of the Tweets using a combination of AFINN and NLTK. A map of the distribution of the sentiments of the Irma-related Tweets and an example of a negative Tweet are shown in **Figures 5-7** and **5-8** respectively.

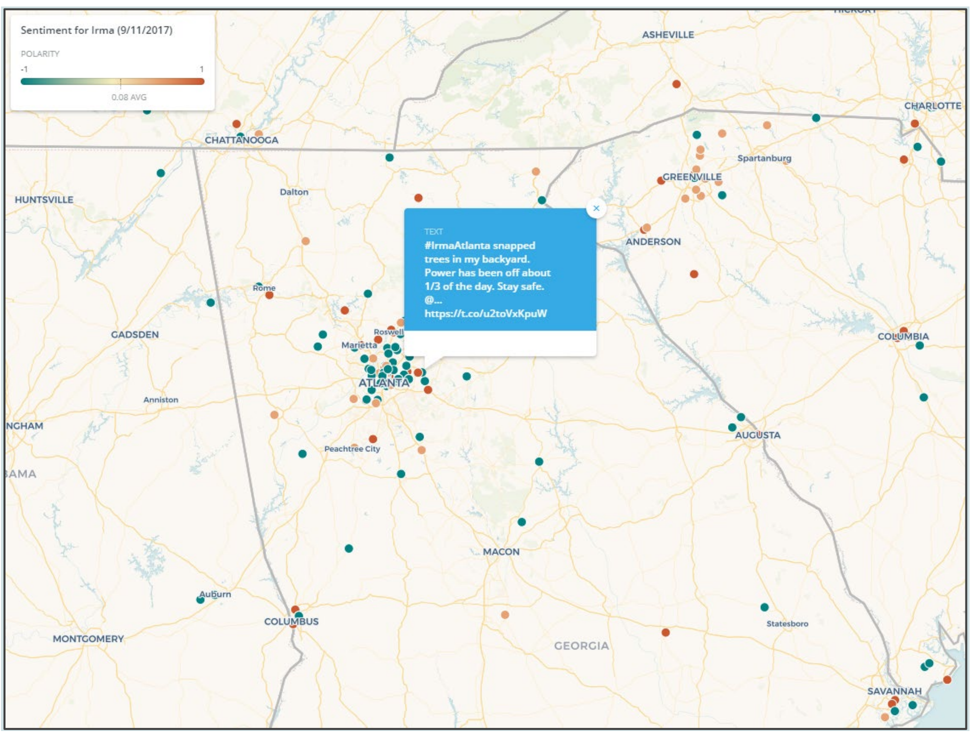


Figure 5-7. Sentiment of Tweets during Hurricane Irma in the state of Georgia.

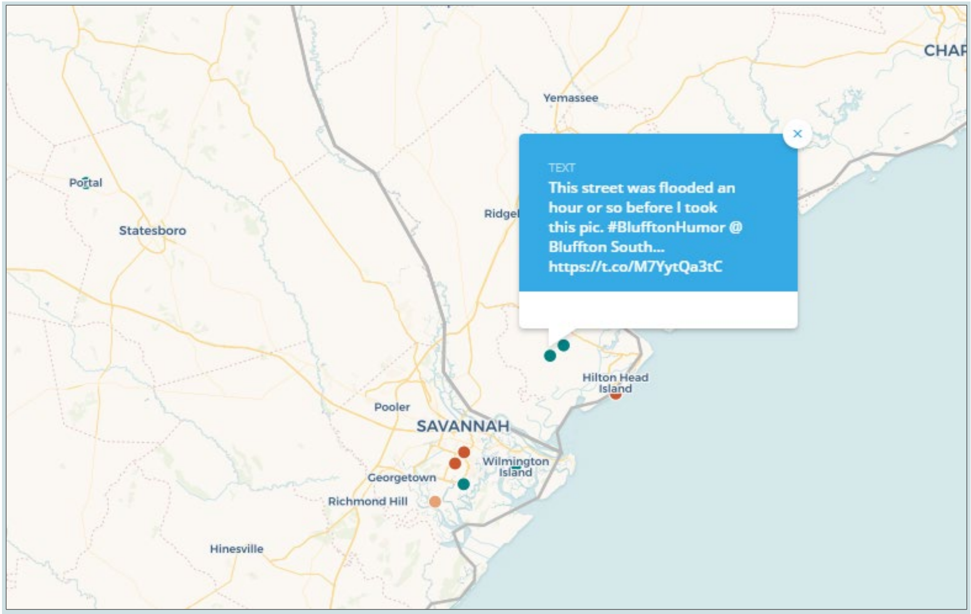


Figure 5-8. Example Tweets identified with negative sentiment during Hurricane Irma in the state of Georgia.

We performed an additional Nearest Neighbor Ratio analysis on the Irma Tweets seeking to identify whether the utilization of both topic and sentiment as weights would improve the ability of the algorithm to identify related clusters; however, the p-value for this analysis was not less than 0.05, indicating that the data was not significantly clustered.

The sentiment distributions by keyword were analyzed for Hurricane Irma as well and are presented in **Figure 5-9**. This was particularly necessary because of the prevalence of the keyword "Irma" identified in the previous chapter. Identifying useful topics of concern and whether sentiment would be a good additional filter for relevant data was key.

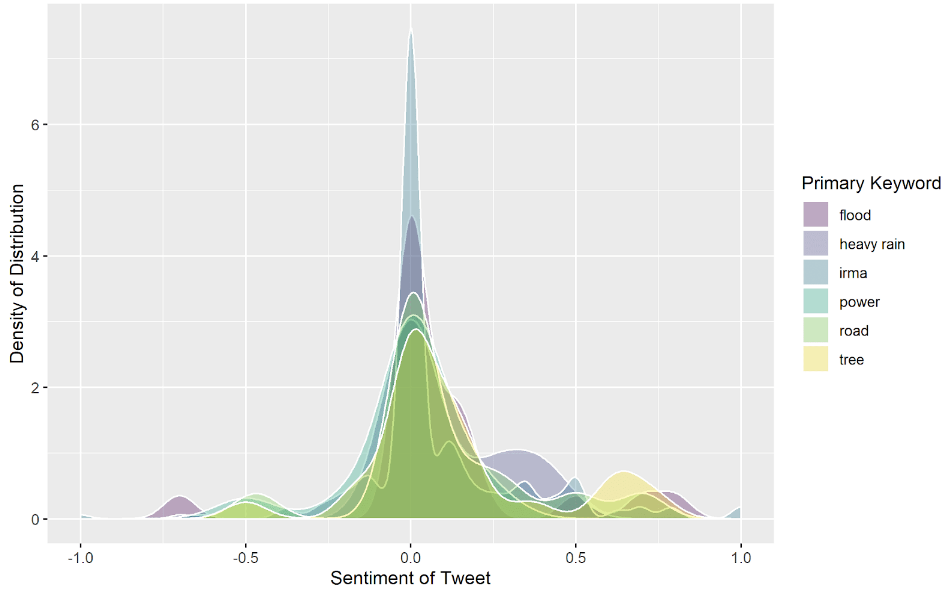


Figure 5-9. Tweet sentiment distribution by geo-topic during Hurricane Irma in the state of Georgia.

Interestingly, there is a strong peak at the "0.0" mark for sentiment for almost all of the Tweets that was not noticeable in the ice events dataset. The greatest of these peaks is for Irma. This is likely because many more of the Irma-related Tweets were descriptive of the storm or an event. The subjectivity of these "0.0"-sentiment Tweets was also very low. However, simply because the sentiment of these descriptive Tweets was not negative does indicate that the descriptions are not useful. Positive sentiment may be an indicator of the irrelevance of the Tweet for topics with an excess of positive sentiment. From these results, thresholds of sentiment may be useful in determining which Tweets should be subject to additional filtration methods or sent to GDOT personnel for additional review. Negative peaks primarily occur in the topics of "flood", "tree", "road", and "power", confirming those topics as ones of particular concern within the analysis.

In terms of the temporal distributions of the posting of keywords, the distributions follow relatively the same pattern as each other and are presented in **Figure 5-10**.

Figure 5-10. Hour of Tweet posting distribution by geo-topic during Hurricane Irma in the state of Georgia.

This pattern is one that is typically observed on non-event days as well. This adherence of posting distributions to typical patterns could be due to the long anticipation of the event and the relatively slow build-up of event occurrences such as flooding or heavy rain. It should be noted that the expected dip in Twitter activity during the night is not matched by the "power" distribution, which might indicate that bursts in people discussing power outages could be relied upon for additional hours and might even be able to indicate when power is restored.

5.3 Topic and Topic Evolution

Finally, we compared the evolution of the topic distributions and sentiment distributions over time and between events. For the topic distributions, we noticed a distinct drop-off of relevant Tweets following the events themselves. Specifically, the number of Tweets that passed through the filters were essentially halved following each event day. Within these halved Tweets, there was an additional decrease in the number of Tweets that were deemed irrelevant to GDOT purpose.

For example, following the ice and snow events on January 17th in which only 3-5% of Tweets were irrelevant, 46% of Tweets on January 18th were irrelevant. This indicates that different sets of filters may be necessary for different stages of an emergency (i.e., preparedness, response, mitigation, and recovery) to identify more relevant Tweets after the immediate danger has passed. This could include additional stop words (i.e. "Thank God" or "survived!"), or changing the focus of the filtration to topics that could persist, such as falling trees. Additionally, the average sentiment there was a decrease in sentiment on the day of landfall for the Irma-related events, as compared to the day before. Sentiments prior to landfall were wary; sentiments after were mostly hopeful. This indicates that sentiment analysis could be additionally useful for disaster-phase specific filtration.

The lessened Tweets and increased average sentiment on the day before Tropical Storm Irma made landfall, September 10th, and the day after landfall, September 12th, are depicted in **Figures 5-11** and **5-12** respectively.

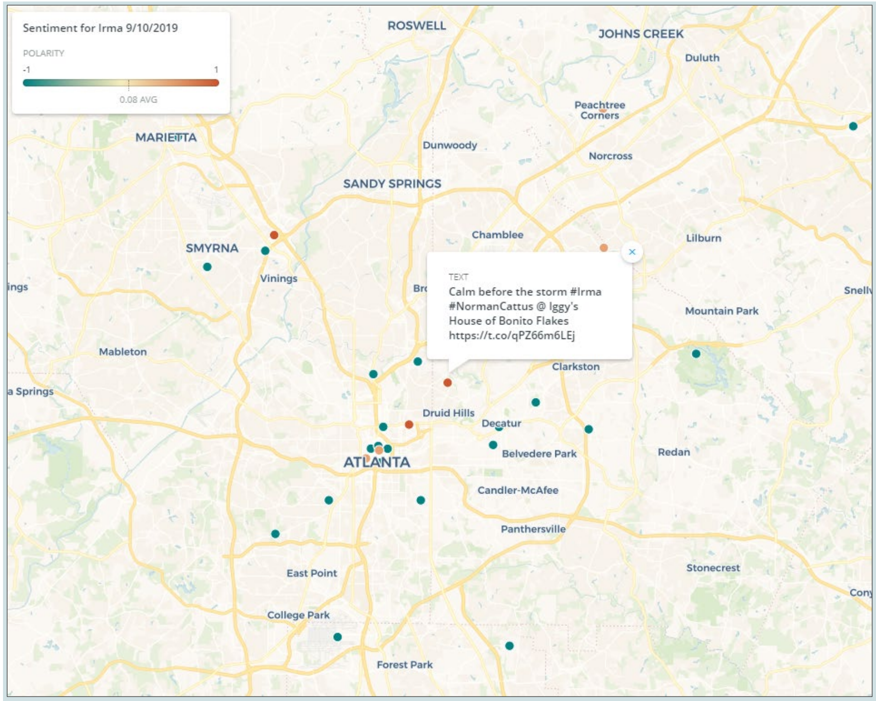


Figure 5-11. Sentiment of Tweets on the day before Tropical Storm Irma's landfall (September 10th) in the state of Georgia.

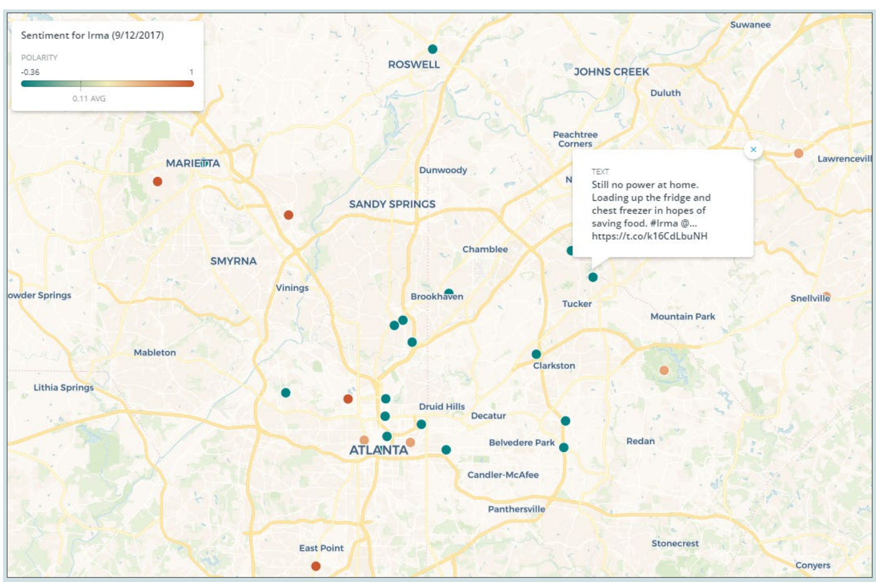


Figure 5-12. Sentiment of Tweets on the day after Tropical Storm Irma's landfall (September 12th) in the state of Georgia.

Finally, in comparing between the two case studies, the distribution of sentiments for both events are depicted in **Figure 5-13**.

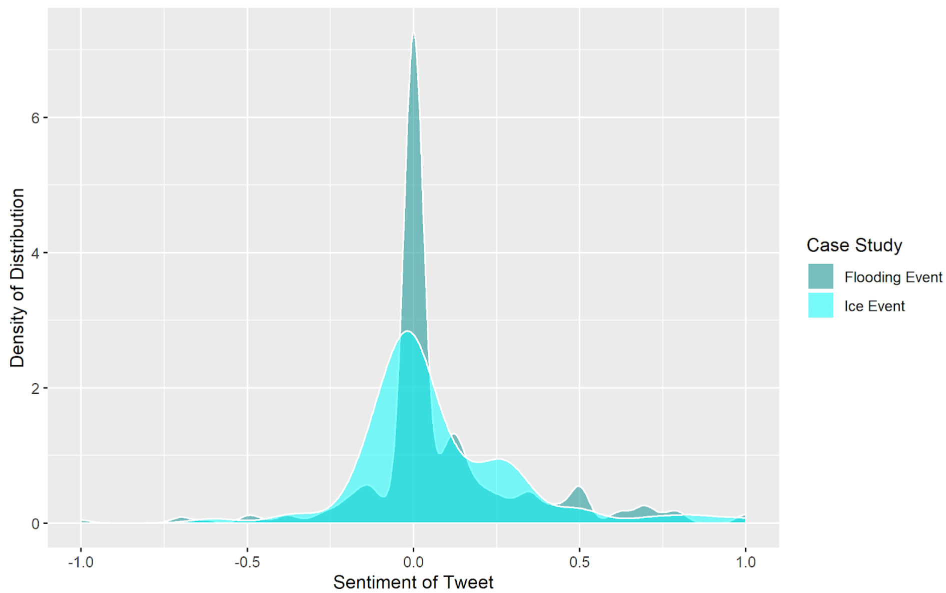


Figure 5-13. Tweet sentiment distribution during ice events and Hurricane Irma in the state of Georgia.

As shown above, the sentiments associated with the flooding event are more symmetrically distributed than the sentiments associated with the ice event, which are more distributed, bimodal, and slightly skewed to the right. This seems to indicate that sentiment may be a more useful filtration method for events that occur suddenly and have a more direct and severe impact on individuals than a large, anticipated event that broadly and often indirectly affects a large number of people.

5.4 Results of the Case Studies

In our analysis of the Twitter data from the two completed test cases, we determined that Twitter texts alone are not sufficient to outpace the time interval between an incident occurring and a Waze alert appearing in the GDOT system. Following this insight, we have made two determinations for the project:

First, that Twitter data can best be used to augment the existing Waze alert process. Eight alerts must be logged before a Waze incident is created; Twitter data can therefore function as an additional source of alerts, thereby increasing the alert dataset's size and reducing the potential time taken to produce an alert. As the GDOT Technical Implementation Team stressed the importance of getting an alert closer to the time at which the event occurred, as, we were told, this will add value to their alerts determination. Additionally, the incorporation of geotagged photos from the Twitter stream will aid personnel by providing visual context for the information in the incidents in addition to aiding in the identification of where an incident is located so that they can more quickly identify either the state road camera facing the incident, or, in more rural areas with fewer cameras, where the incident is actually located. Lastly, Twitter posts will produce a direct user contact that could be accessed by the GDOT Communications Officer.

Second, that more reliable, relevant, and reliably relevant data could be produced through a social media campaign encouraging users to take and post pictures of incidents on Twitter. Potential future work would include developing a hashtag specific to this campaign and/or relevant incidents and performing a pilot study of the idea utilizing GDOT employees.

A combination of the above two paths forward would, we believe based on the data presented above, be most beneficial for GDOT's response capabilities.

6 SPECIFYING AN INTERACTIVE TRACKING AND VISUALIZATION SYSTEM

6.1 Framework Development: GDOT Needs

We met with members of the GDOT Technical Implementation team to understand the needs that could be met through social media. In so doing, we toured various centers (Transportation Management Center (TMC), State Operations Center (SOC), and Web Emergency Operations Center (WebEOC)) and met with the teams who manage the emergency operations at GDOT including: Transportation Systems Management and Operations (TSMO), the Intelligent Transportation Systems (ITS) group, as well as the Georgia Emergency Management Agency (GEMA) to discuss the best path towards incorporating our information into the existing response framework at GDOT. Our highest priority in understanding these systems was to find the best way to integrate multiple social media and official data streams into a system that performs a risk assessment of potential hazards.

In addition to understanding the integration of the systems, we also sought to understand GDOT's existing social media data filtration practices. We reviewed the National Weather Service (NWS) data that is used, the requirements for a verified instance on Waze (8 clicks), and how the EMA directs posts from 911. We also reviewed potential data inputs and the potential problems/barriers with data interoperability. We additionally sought feedback on our design and research progression from three members of the Social Media Communications team. These insights were primarily used in the specification of a visualization system, and they are further discussed in the test case framework research outlined below.

6.2 Exploration of Information Platforms

With respect to the visualization platform, we investigated the potential for the incorporation of external social media visualization and analytical platforms. These included EchoSec, DataCapable, and Esri's in-house ArcGIS response platform.

In Phase 1, we had sought to review alternative data sources/streams (1.1.1) and had identified Waze as a possible data source to integrate. Especially upon reviewing the somewhat limited capacity for custom analytics available on the 3rd party platforms mentioned above and following our insights from the test cases completed in Phase 1 and 2, we decided to proceed by developing a framework for incorporating Twitter data into the existing Advanced Transportation Management System (ATMS) work flow. Based on the information obtained by the potential end users above, we decided to specify a system that would generate an alert on the WebEOC map utilized by the SOC operators in a manner similar to the Waze data.

We have proceeded in generating a test case demo framework. Our work on a test case demo consisted of three primary avenues of investigation: *1. Improve usability of Twitter data, 2. Define Twitter data and Waze convergence,* and *3. Access GDOT's existing data streams*. That work is detailed in the following section.

6.3 Test Case Framework

6.3.1 Improve Usability of Twitter Data

There were a number of factors noted above that seemed to indicate the need for a "human-in-the-loop" with respect to the filtration of the Twitter data. Although the algorithms and methods that were utilized above have been able to filter the hundreds of thousands of incoming Tweets down to a few hundred, the data indicates that the ultimate recognition of whether the information produced should be used to motivate the distribution of GDOT personnel and resources. In speaking with the GDOT Social Media Coordinators, those personnel agreed with that determination, and offered insights into their ability to vet Tweets and social media information.

Ivan Lichtenstein, the Media Liaison for Georgia 511, described his involvement with social media as primarily unilateral with social media users. His primary use involved distributing information about events and providing information about alternative routes. He additionally provided information about the social media engagement of GDOT, which showed that the number of people who viewed GDOT Twitter posts was double that of their Facebook (54,000 versus 27,000) even though the relative number of engagements for Twitter was lower (1,000 to 4,900).

This interaction and data laid the groundwork for an innovative procedure to request additional information from Twitter users regarding disaster events in order to promote the generation of more geotagged and relevant data for incorporation into the system. The heightened number of views for Twitter indicates that more people would see a posted request for information on Twitter than on Facebook. The decreased number of engagements on Twitter is likely a reflection on how little information can be included in a 280-character text, but it also suggests that people might be interacting with the given information on their own posts instead of as replies to GDOT's posts. If this is the case, more original posts and community information spread would be beneficial to our system and deserves additional investigation.

We additionally discussed a proposed framework with the Social Media Coordinators Breawna Kirkpatrick and Susan Rodman. Ms. Kirkpatrick indicated that she engaged in more bilateral communication, discussing complaints and concerns with individuals, and claimed that an incorporation of Twitter messages into the filtration system would be beneficial to her. She agreed that a human-in-the-loop would be necessary as a final step in filtering usable data into the system, and that she agreed with the concept of an automated system sending a summary of data to a GDOT employee who would ultimately decide on whether or not the event should be logged in the GDOT event system. Finally, she requested the incorporation of three additional confidence metrics into the system: analyses of "Is this about the designated topic?", "Is this a problem relevant to GDOT?", and "Is this in the designated location?"

6.3.2 Define Twitter Data and Waze Convergence

We utilized the test cases to define the necessary Twitter information for the development of a joint Waze and Twitter SQL database. This involved researching the data format conversion, methods of expanding the dataset of geolocated data (SensePlace2, generating lists of trusted users, and literature on identifying hyperlocal incidents), and methods for developing an internal "confidence score" for Twitter data to match a similar metric on Waze data (bot determination, rumor identification). Following this, we created a framework for the processing of Twitter data and the convergence of Twitter and Waze data.

This required the conversion of our stored JSON Twitter data to match the format of the incoming Waze data, and to convert the incoming Waze XML to JSON to jointly store the data in a single SQL database. The proposed framework for the data convergence is described in the following chapter. The specified framework for the data convergence and analysis includes the conclusions derived from the spatial and temporal proximity results, topic filtration results, and the sentiment filtration results for Twitter data. These results and the concerns of the Social Media Coordinators indicated above were used to ultimately match the format of the Waze reports for improved integration into the existing Waze data analysis method currently utilized (albeit by a 3rd party) by GDOT.

6.3.3 Access GDOT's Existing Data Streams

The extent of the scope of our initial proposal was limited to the specification of the interactive visualization system; however, based on our analyses, future research should test the data convergence of the Twitter data and the Waze data in a real-time data streaming setting. In order to complete such a test it will be important to address any data privacy concerns and/or technical issues. Completing this was beyond the scope of this project.

7 DESIGNING THE INTERACTIVE TRACKING AND VISUALIZATION SYSTEM

In addition to specifying the interactive tracking and visualization system outlined above, we developed specifications for how the Twitter data would be incorporated into the Waze data filtration system and be vetted by a GDOT employee. This system is designed to convert Twitter data into Waze data, and thus we have designated it as a “convergence system”. The specification for how the convergence system would process incoming data is described in this chapter of the report in the images below following **Figure 7-1**.

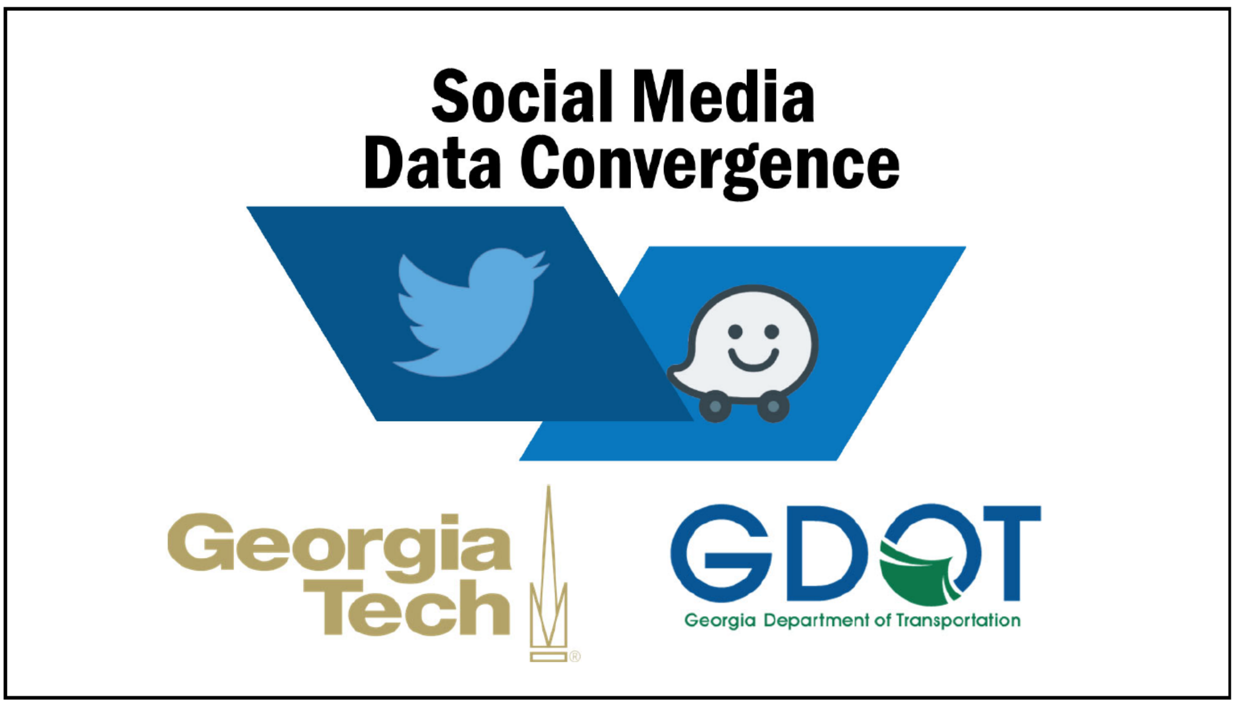


Figure 7-1. Introduction to the depiction of the proposed social media convergence system.

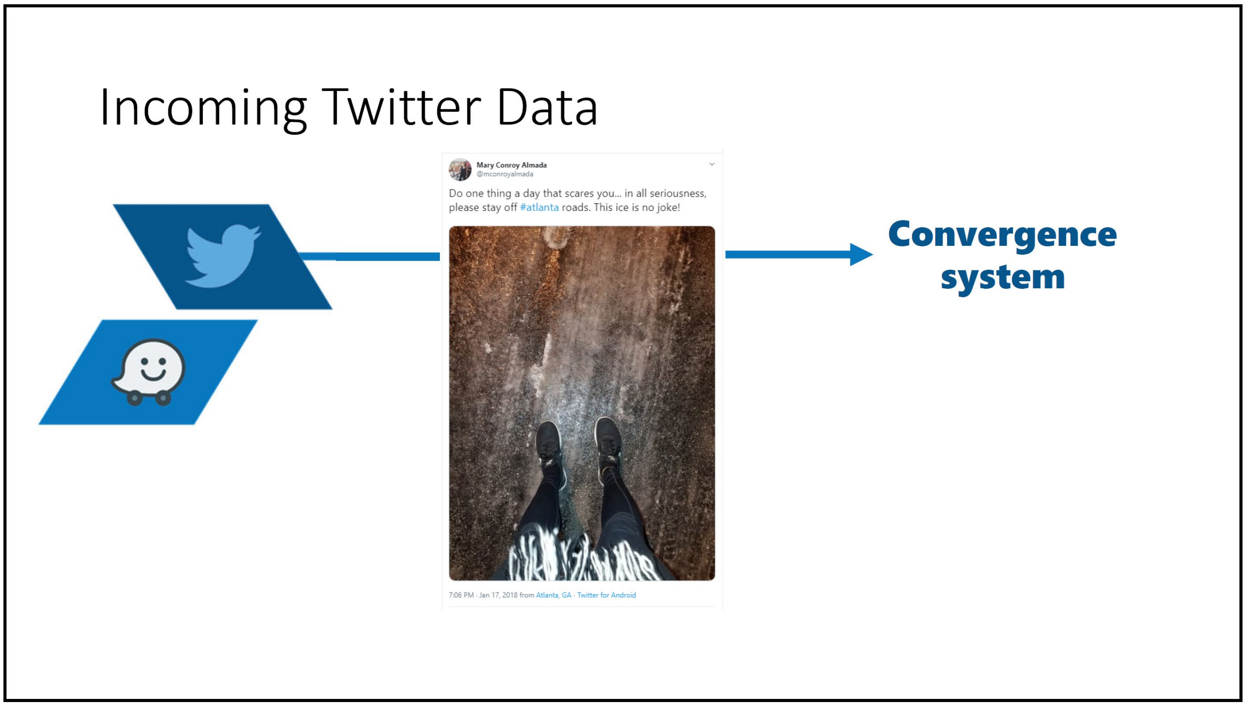


Figure 7-2. Depiction of a potential Tweet that would be caught by the data filtration methods and directed into the data convergence system.

One of the determinations made through our communications with GDOT personnel was the necessity of seamlessly integrating the Twitter data with the existing Waze data structure, creating a style of notification that would enable easy data comprehension, and decreasing the time necessary for GDOT personnel to receive a relevant, informative emergency alert. In line with those goals, we developed a method of transforming the incoming Twitter data such that it would match the Waze data format prior to being sent through the ATMS and ultimately to 511. **Figure 7-2** shows an example Tweet, one that was posted during a winter storm, and the associated data that would be analyzed by the system.

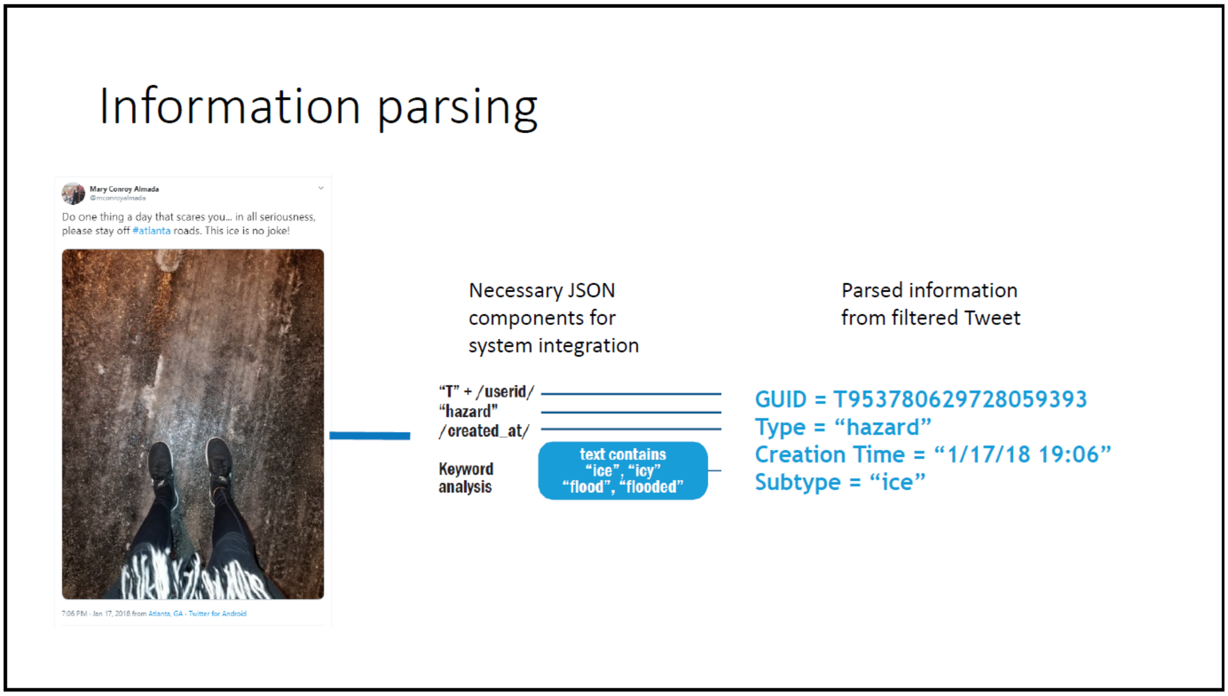


Figure 7-3. Depiction of how the necessary JSON components for system integration would be derived from the Tweet information.

As described earlier in the report, we transitioned our flat data storage from pickle to JSON in order to facilitate storage in a SQL database. As Waze can also be accessed in a JSON format and is stored in a SQL database by GDOT, the convergence of these two data streams into a single unified database from which to perform analyses is a matter of translating the Twitter data into the series of attributes utilized by Waze. The "Information parsing" stage of the design process depicted in **Figure 7-3** shows how the information from the incoming Tweets are parsed into the JSON components necessary for integration.

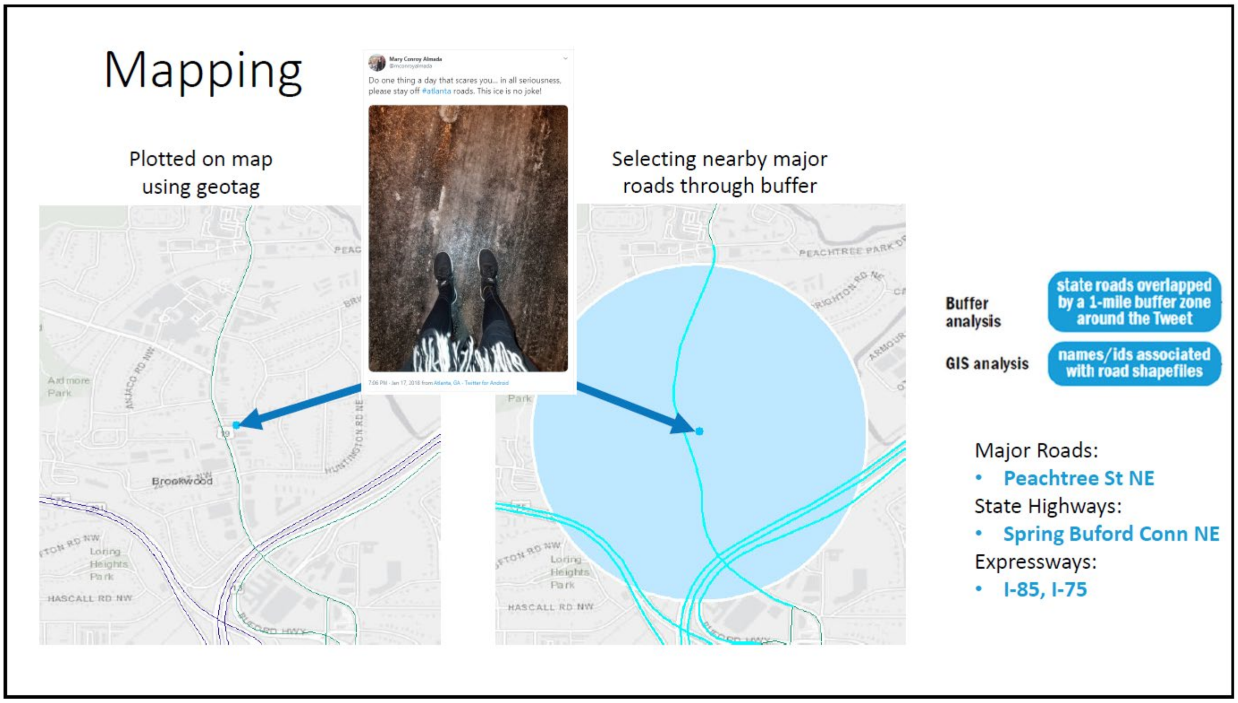


Figure 7-4. Depiction of how the necessary geospatial components for integration with Waze data would be derived from the geolocated Twitter data and GIS analyses.

As the geographic location of the Twitter data is of key importance in determining where the noted event is happening but the location of the data is not often directly on a road under GDOT's jurisdiction, a circular buffer is created around the Twitter data point. The depicted buffer in **Figure 7-4** is 1 mile in diameter (0.5 miles in radius) and can be adjusted within the system. Any roadway section that overlaps this buffer is highlighted and its name is placed into the system using GIS.

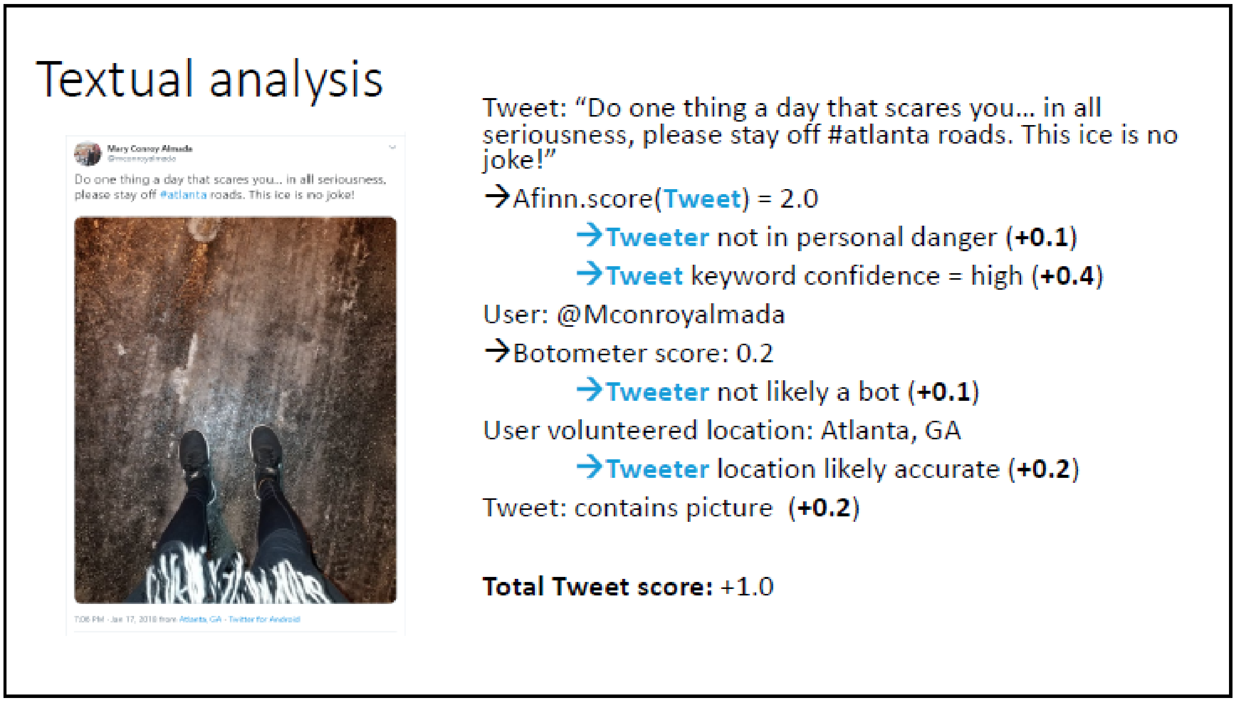


Figure 7-5. Depiction of how the information value metric, or "Tweet score", is determined through sentiment analysis, analysis of the user's Twitter history, location confidence, and the potential for additional value in the form of an image.

The next step of filtration is the value metric for the Twitter data and is shown in **Figure 7-5**. The GDOT Technical Implementation Team advised us that each Waze data point that enters the system is only utilized if it has a confidence score of at least 8 out of 10. Additionally, 8 Waze data points are necessary to send an alert to the GDOT system, which is one of the reasons for the increased time between an event occurring and the GDOT team being notified. In the proposed system, the Twitter data will function as additional points of value towards reaching the total 8 points. It will also provide additional value by including pictures and a point of contact for GDOT if necessary. However, as noted, some of the Tweets identified in the analysis are not as valuable as others. To address this concern, a framework was developed to assign a value of information metric to the Tweet through the sentiment of the Tweet (shown in the graphic as being determined through AFINN); the likelihood of the user being a bot; the likelihood of the user being in the geolocated position based on their home location; and whether or not the Tweet contains a picture. The thresholds for sentiment, bot likelihood, and the worth of each point can be adjusted within the system to better serve GDOT's needs and the ultimate findings of the process.

An overview of the entirety of the analysis process is presented in **Figure 7-6**, which shows the two-pronged analysis steps of mapping the point to determine its geographical relevance and the text parsing steps to determine its keyword and topic relevance. The combination of the attribute information and the mapping is sent to the visualization of the event.

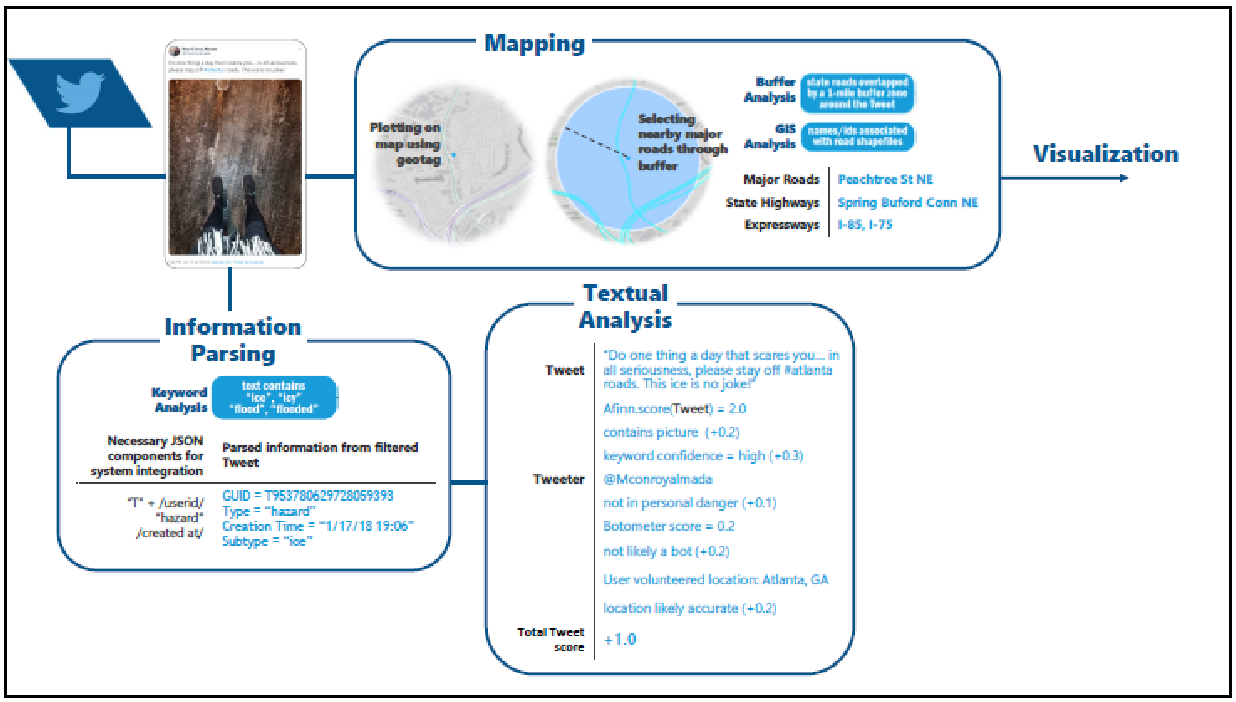


Figure 7-6. Depiction of an overview of how the different components of an incoming, filtered Tweet would be converted into the Waze format for joint data analyses.

As can be seen our example scenario, the value of the information assigned to the incoming Tweet is +1.0. There are seven Waze points within a 0.5-mile radius range of the Tweet with the same associated type ("hazard") and subtype ("ice"). The confluence of these eight data points sends an alert to a social media coordinator with the location of each data point on a map, the text of the Tweet, the included picture, and a way of contacting the user who posted with Tweet for more information. From this data, the social media coordinator can choose to dismiss the alert, adjust the details of the alert, or confirm the alert as an incident to be logged into the WebEOC system. The ultimate visualization of the event would include both information from Waze and information from Twitter for cross-reference. A summary of what the ultimate visualization would be on a map for a GDOT employee is shown in **Figure 7-7**.

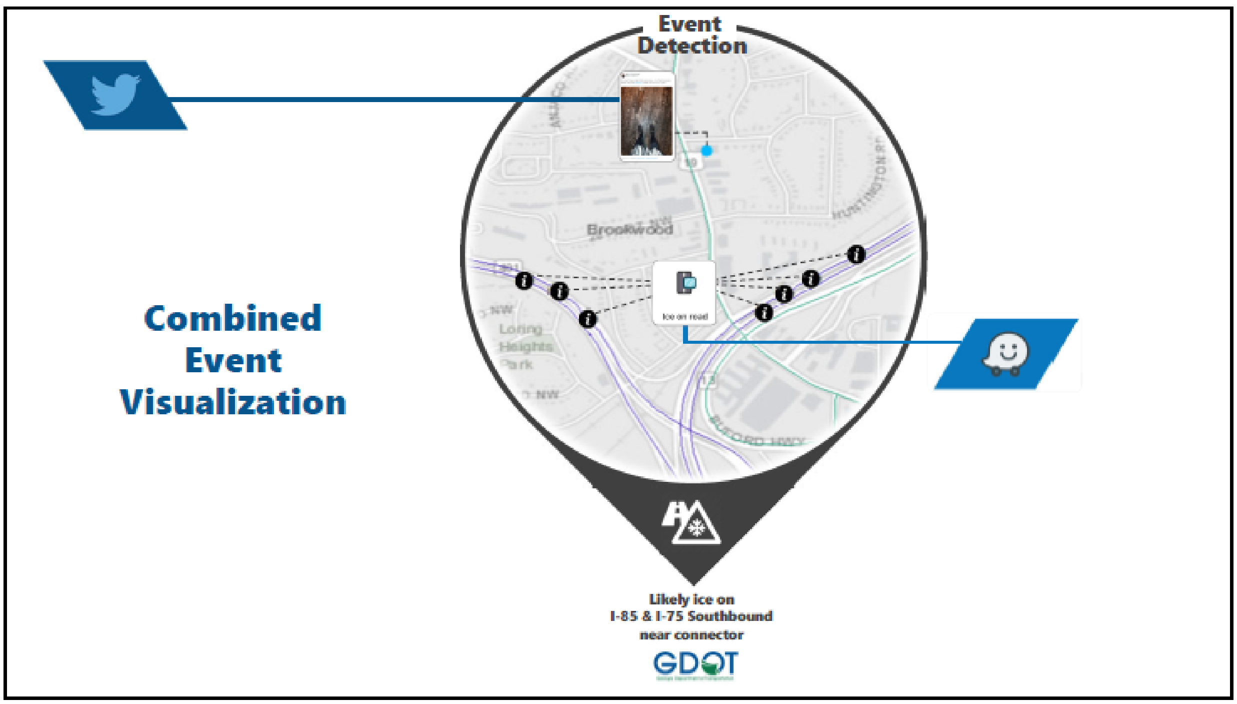


Figure 7-7. Depiction of how the combined Twitter and Waze data points would appear on a map for assessment by GDOT personnel.

8 CONCLUSIONS

The central goal of this research project was to identify the potential utility of social media data as a source of information for GDOT during extreme events. To identify that utility, the research team developed and tested two case studies, one for a winter storm event and one for an episode of extreme amounts of flooding. Within each of those case studies, the goal was to identify the relevant information contained within posted social media data, evaluate the relevant importance of that data and what can be determined through Natural Language Processing (NLP) techniques, and specify a tracking and visualization method for GDOT that could be developed as a potential secondary stage of the project. The chosen case studies that were analyzed within this report were 1) the winter storm that impacted north Georgia on January 16th and 17th, 2018, and 2) the flooding that occurred across the state of Georgia as a result of the impact of Tropical Storm Irma from September 10th to the 17th, 2017.

Utilizing the two case studies, we developed a method of filtering geolocated Tweets from the tens of thousands of Tweets that are produced within the state of Georgia every day. Each Tweet was assessed based on a list of concerns developed between the GDOT Technical Implementation Team and the research team. In terms of filtering the Twitter stream into relevant data, the team first identified the necessity of utilizing a variety of filtration methods. Keywords alone were not sufficient to filter out unwanted or irrelevant Tweets due to the prevalence of hazard words (such as “ice” or “flood” or “power”) in alternate concerns (such as “ICE (Immigration Control and Enforcement)”, emotional comments such as “thank you for the flood of well-wishes!”, or metaphors such as “looks like the power’s out for the Falcons!”). Sets of stop words and semantic analyses such as parts-of-speech analysis were additionally necessary in filtering out the unwanted, irrelevant Tweets. We also concluded that the utilization of keyword pairs involving GDOT-specific concerns (such as a Tweet with the pairs “icy roads” or “frozen streets”) were of pivotal necessity in identifying very relevant Tweets. Thus, a minimum of a primary and secondary set of keywords is necessary in Twitter filtration. In evaluating the Twitter data and feedback from GDOT personnel, we also identified the potential value in the geolocated images associated with incoming Tweets. The images posted of trees across roadways and black ice on top of roads showed specific information on the magnitude and exact location of the events in question. From this, we concluded that incorporating Twitter images into the alerts the system would produce would provide a substantial amount of value.

Following the evaluation of topic detection, sentiment analysis was evaluated as a method of ranking the value and criticality of posted data. This approach allowed for two distinct evaluations: is someone in danger, and is he or she posting about a dangerous event? Following our analysis of the sentiment of each Tweet incorporated into the case studies, we concluded three things. The first was that the utility of sentiment as a metric was highly dependent on the keyword category to which the Tweet was applied. Sentiment was useful in identifying relevant and irrelevant Tweets for keywords that had a high percentage of irrelevant Tweets that were not caught by the above-outlined filtration methods (such as “power”). Sentiment was most particularly useful in identifying the most relevant Tweets posted within the limits of a major metropolitan area such as Atlanta, which was identified as having a higher percentage of less relevant Tweets than areas with less dense populations. We also analyzed the subjectivity of the Tweets and found low-subjectivity Tweets with pictures to contain the highest quality of information. Following this, we concluded that sentiment and subjectivity are good indicators of information value, although they are not good indicators of whether or not a person is actively in danger due to a specified hazard.

Finally, in designing and specifying a system for GDOT, we concluded based on the availability and quality of data filtered through our data intake processes, that Twitter data would best be used to supplement the incoming UVI data that GDOT is already using. From our conversations with the GDOT Technical Implementation Team, the team's highest priority is reducing the time taken for UVI data to generate an alert that is viewed by GDOT employees. Additional information about the event to assist in locating it and assessing the severity of the event would be second priority. From this and our review of existing UVI analytical platform, our conclusion was that the best way to visualize and track events using Twitter data would be to develop a system that could be incorporated with the existing Waze data stream into the ATMS that GDOT is already using. This would generate the most added value (reducing the time to 't-zero') while minimizing the additional training and software necessary to include the data in GDOT operations.

We additionally concluded, based on conversations with the GDOT Social Media Coordinators, that the social media data alerts generated through the system would need to be reviewed by GDOT personnel before they could be put into the system as an incident. Utilizing these conclusions from our assessment of the quality and components of our assessed Twitter data, our review of the existing GDOT operations structure, and our conversations with both the GDOT Social Media Coordinators and the Technical Implementation Team, we generated a framework for the conversion of Twitter data into the Waze data format and a design for how the finished system would operate.

Finally, we developed recommendations for how best to implement and test the framework and system design that was generated. In doing so, we created solution designs for the two most immediate problems that we identified. First, to test the added value of the system (such as improved event timing) and to determine the optimal thresholds of the filtration methods outlined in Chapter 7, we proposed to build a pilot version of the system using historical data and comparing the results of that pilot test to what the GDOT system produced during that historical period. Second, to improve the availability and quality of the Twitter data, we proposed a social media campaign similar to the "See Something, Say Something" campaign. Further details on these two recommendations are outlined in the following chapter.

9 RECOMMENDATIONS

Following this research, we recommend a two-pronged approach for implementing an social media-informed interactive tracking and visualization system for detecting crises in the state transportation network.

The first is a technical approach for testing the reliability and value of the outlined framework for data convergence and interpretation. As the team was unable to access the Waze data stored by GDOT, the team was unable to numerically evaluate the positive impact of incorporating Twitter data into the existing Waze data stream. The first step towards assessing the potential profit of the framework would be to assess the data overlap between Twitter and Waze, the number of events that would be triggered by incoming Twitter data, and the historic incident occurrences (i.e. salt truck deployments, accidents, etc.) that occurred in the vicinity of the alerts that would have been created. Additionally, the most optimal thresholds of the various filtration methods utilized throughout the report would be determined by training the system on these test cases. These two test cases would likely utilize the data and findings of this report. Following the determination of the added value from Twitter data and the optimum thresholds for filtration, the system would then need to be tested using a different crisis event. Finally, on completion of that work, a system would need to be created to incorporate and parse data in real time. That new system would need to be activated during a real crisis and have its results (i.e. the events it identifies and the information it provides) compared to the events identified through the existing Waze evaluation system. The differences between the two systems and the added value of information provided by the new system would then need to be assessed on its relevance to GDOT personnel.

The second is a social approach for increasing the reliability and relevance of Twitter data. We recommend the design and implementation of a social media campaign. The intent of the social media campaign would be to design a hashtag and system to follow that would assist both social media users in generating content that would be more easily parsed by the system and would in turn be more easily interpreted by social media coordinators. From our research into the topic, people believe that the information that they produce is being monitored by government organizations. If they are using a tool (social media) that is already close at hand to document a hazard that they see, but they are not tagging it in such a way that it can be registered by the social media system or GDOT employees, then that data is lost. Improving social recognition of how to have their data seen through a hashtag campaign or awareness campaign would improve the amount and quality of social media data that is available for GDOT use. This hashtag and data generation framework could be piloted on a small scale (such as involving GDOT personnel) and then expanded following the evaluation of the pilot program.

10 REFERENCES

de Albuquerque, J. P., Herfort, B., Brenning, A., and Zipf, A. (2015). “A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management.” International Journal of Geographical Information Science, 29(4), 667–689.

Amin-Naseri, M., Chakraborty, P., Sharma, A., Gilbert, S. B., and Hong, M. 2018. "Evaluating The Reliability, Coverage, And Added Value Of Crowdsourced Traffic Incident Reports From Waze". Transportation Research Record, 2672(43): 34–43.

Barich, F. (2019). “Social Media Guidebook for Emergency Management.” Transportation Review Board. In progress research.

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. " O'Reilly Media, Inc.".

Chaney, A. J. B., and Blei, D. M. (2012). “Visualizing topic models.” ICWSM 2012 - Proceedings of the 6th International AAAI Conference on Weblogs and Social Media.

Cutter, S. L., and Derakhshan, S. (2018). “Temporal and spatial change in disaster resilience in US counties, 2010–2015*.” Environmental Hazards, Taylor & Francis, 0(0), 1–20.

Guth, D. (2019). “Emergency Management Agencies: Pilot for a Crisis Communication Analysis Assessment Test.” Transportation Review Board. In progress research.

Houston, N. (2006). “Best Practices in Emergency Transportation Operations Preparedness and Response: Results of the FHWA Workshop Series.” (Report No. FHWA-HOP-07-076). Washington, DC: Federal Highway Administration.

Imran, M. ., Castillo, C. ., Lucas, J. ., Meier, P. ., and Rogstadius, J. . (2014a). “Coordinating human and machine intelligence to classify microblog communications in crises.” ISCRAM 2014 Conference Proceedings - 11th International Conference on Information Systems for Crisis Response and Management, (May), 712–721.

Imran, M., Castillo, C., Lucas, J., Meier, P., and Vieweg, S. (2014b). “AIDR: Artificial intelligence for disaster response.” Proceedings of the companion publication of the 23rd international conference on World wide web companion, (October), 159–162.

Imran, M., Mitra, P., and Castillo, C. (2016). “Twitter as a lifeline: Human-annotated Twitter corpora for NLP of crisis-related messages.” Proceedings of the 10th International Conference on Language Resources and Evaluation, LREC 2016, European Language Resources Association (ELRA), 1638–1643.

Jennex, M. E. (2012). “Social Media – Viable for Crisis Response?” International Journal of Information Systems for Crisis Response and Management, 4(2), 53–67.

de Jong, P., Sprenger, C., and van Veen, F. (1984). “On Extreme Values of Moran’s I and Geary’s c.” Geographical Analysis, 16(1).

Kropczynski, J., Grace, R., Coche, J., Obeysekare, E., Bénaben, F., Halse, S., Montarnal, A., and Tapia, A. (2018). “Identifying Actionable Information on Social Media for Emergency Dispatch.” Proceedings of ISCRAM Asia Pacific 2018, 1–11.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. arXiv preprint arXiv:1103.2903.

Nobles, M., Neill, D. B., Lall, R., and Mathes, R. (2019). “Multidimensional Semantic Scan for Pre-Syndromic Disease Surveillance.” Online Journal of Public Health Informatics.

Purohit, H., Hampton, A., Bhatt, S., Shalin, V. L., Sheth, A. P., and Flach, J. M. (2014). “Identifying Seekers and Suppliers in Social Media Communities to Support Crisis Coordination.” Computer Supported Cooperative Work: CSCW: An International Journal, 23(4–6), 513–545.

Reuter, C., and Kaufhold, M.-A. (2017). “Fifteen years of social media in emergencies: A retrospective review and future directions for crisis Informatics.” Journal of Contingencies and Crisis Management, 41–57.

Sack, R., Foreman, C., Forni, S., Glynn, R., Lehrer, A., Linthicum, A., & Perruzzi, A. (2019, May). Social media practices in traffic safety (Report No. DOT HS 812 673). Washington, DC: National Highway Traffic Safety Administration.

Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). “Earthquake shakes Twitter users.” Proceedings of the 19th international conference on World wide web - WWW ’10, ACM Press, New York, New York, USA, 851.

Schubert, E., Weiler, M., and Kriegel, H.-P. (2014). “SigniTrend.”

Sheffi, Y. and Goentzel, J. (2019). “Big Data During Crisis: Lessons from Hurricane Irene.” (Report No. MITR24-10). Cambridge, MA: New England University Transportation Center.

Starbird, K., Maddock, J., Orand, M., Achterman, P., and Mason, R. M. (2014). “Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing.” iConference 2014 Proceedings.

Stirtz, K. (n.d.) "Twitter Ranking: Which States Twitter The Most?". Accessed on March 05, 2020. URL: https://www.allbusiness.com/twitter-ranking-which-states-twitter-the-most-12329567-1.htmlUkkusuri, S. and Ozbay, K. (April, 2016). “Modeling Disaster Operations from an Interdisciplinary Perspective in the New York-New Jersey Area." (UTRC/RF Grant No: 49997-61-24). New York University: University Transportation Research Center.

Wang, Q. and Taylor, J. (2014). "Quantifying Human Mobility Perturbation and Resilience in Hurricane Sandy," PLOS One, 9(11): e112608. doi: 10.1371/journal.pone.0112608.

Wang, Q., and Taylor, J. (2015). "Resilience of Human Mobility under the Influence of Typhoons." Procedia Engineering, Elsevier B.V., 118, 942–949.

Wang, Y. and Taylor, J. (2018). "Coupling Sentiment and Human Mobility in Natural Disasters: A Twitter-Based Study of the 2014 South Napa Earthquake," Natural Hazards, 92(2): 907-925.

Wang, Y., Wang, Q. and Taylor, J. (2017). "Aggregated Responses of Human Mobility to Severe Winter Storms: An Empirical Study," PLOS ONE, 12(12): e0188734. doi.org/10.1371/journal.pone.0188734.

Wang, Y., Agichtein, E., and Benzi, M. (2012). "TM-LDA: Efficient online modeling of latent topic transitions in social media." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Wang, Y., and Taylor, J. E. (2019). "DUET: Data-Driven Approach Based on Latent Dirichlet Allocation Topic Modeling." Journal of Computing in Civil Engineering, 33(3), 04019023.

Xie, W., Zhu, F., Jiang, J., Lim, E. P., and Wang, K. (2016). "TopicSketch: Real-time bursty topic detection from twitter." IEEE Transactions on Knowledge and Data Engineering.

Yin, J., Karimi, S., Lampert, A., Cameron, M., Robinson, B., and Power, R. (2015). "Using social media to enhance emergency situation awareness:" IJCAI International Joint Conference on Artificial Intelligence, 2015-Janua, 4234–4239.

Zhang, C., Zhou, G., Yuan, Q., Zhuang, H., Zheng, Y., Kaplan, L., Wang, S., and Han, J. (2016). "GeoBurst: Real-time local event detection in geo-tagged tweet streams." SIGIR 2016 - Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval.

Zhang, M. and Pan, Q. (2019). "Utilize Crowd-Sourced Data and Machine Learning Technology to Enhance Planning for Transportation Resilience to Flooding." Transportation Review Board. In progress research.