Enhancing Disaster Situational Awareness via Automated Summary Dissemination of Social Media Content

Shanshan Zhang, Amitangshu Pal, Krishna Kant and Slobodan Vucetic
Computer and Information Sciences, Temple University, Philadelphia, PA 19122
E-mail:{zhang.shanshan,amitangshu.pal,kkant,vucetic}@temple.edu

***Abstract*—The paper proposes a situational awareness service, named `StayTuned` that collects information from social media, extracts relevant messages, and broadcasts them to the subscribers through wireless emergency alert system. `StayTuned` uses automated *filtering* and *summarization* of messages and updates subscribers with real-time situational summaries. Extensive experiments were conducted using twitter data collected during the Sandy hurricane to evaluate performance of the automated message extraction.**

***Index Terms*—Emergency update application, twitter, crisis management, situational awareness, filtering, summarization.**

I. INTRODUCTION

During a natural or man-made emergency or a disaster, information useful for situational awareness typically arrives from multiple sources, including content generated by people within the impacted areas and data collected by emergency responders. As an emergency unfolds, it is likely that communication infrastructure becomes damaged and heavily strained and that the demand for sending and receiving information vastly exceeds the available capacity. One of the major challenges in disaster response is maintaining flow of information, such that valuable information generated within a disaster area could be disseminated and that the impacted population could stay informed. This paper proposes a system aimed at maintaining flow of information during emergencies and disasters.

While there are multiple information sources, in this paper we focus on collection and dissemination of Twitter messages generated during a disaster. As evidenced by several past disasters, Twitter has established itself as the disaster communication vehicle of choice due to its modest networking requirements, ease of use, and brevity. For example, soon after the 2011 Japanese earthquake, a volume of disaster related tweets exceeded 5,500 per second. Twitter has also been instrumental during a wide variety of emergencies and disasters, such as fires [1], floods in 2009 [2], and terrorist attacks. Although there have been many recent studies on the use of social networks for disaster response, a uniqueness of our work is its focus on resource-aware dissemination of information. In this direction, we propose a *situational awareness service* named `StayTuned` that efficiently processes disaster related information and disseminates it in a compressed form to the impacted population. A user interface of the `StayTuned` is shown in Figure 1.

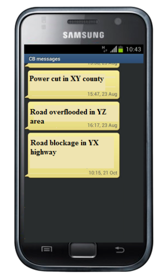


Figure 1: `StayTuned` UI

Although it is easy to envision a fully automated system for processing and dissemination of tweets and real-time optimization of the available communication resources, such a vision is still difficult to realize using the current state of the art. Therefore, we consider a human-in-the-loop system. We assume that a large representative subset of disaster related tweets is available to the emergency authorities who could then select a subset of tweets to be broadcast to the impacted population. The cellular subscribers receive the selected messages using `StayTuned`. Important and representative tweets are selected in two phases. In the first phase, the disaster-related messages are selected by a classification algorithm trained using labeled and unlabeled data from the current and previous related emergencies. Whenever a large group of people are reporting about events within a limited area, it is to be expected that generated messages have significant level of redundancy. To reduce redundancy, during the *summarization* phase a small subset of non-redundant and informative messages are selected by a linear programming algorithm. Extensive simulations using real tweets show that such filtering and summarization process can extract a small fixed subset of messages which are highly representative of the generated stream of disaster-related messages.

The outline of this paper is as follows. Section II proposes the network architecture and protocols for `StayTuned` system implementation. Section III introduces the proposed filtering and summarization framework for extracting a representative subset of relevant messages from the available message pool. Section IV-V describes the experimental setup and shows results obtained by using the tweets collected during the Sandy Hurricane. Discussion of related work is provided in section VI. Finally, the conclusion is provided in section VII.

II. NETWORK ARCHITECTURE FOR `STAYTUNED`

Public Warning System (PWS): In an event of an emergency, the PWS authorities need to provide warnings and update information to the people in an efficient manner. To quickly and effectively alert and warn the public about serious emergencies, Wireless Emergency Alerts (WEAs) is made available through the Integrated Public Alert and Warning System (IPAWS) infrastructure. All major cellular providers

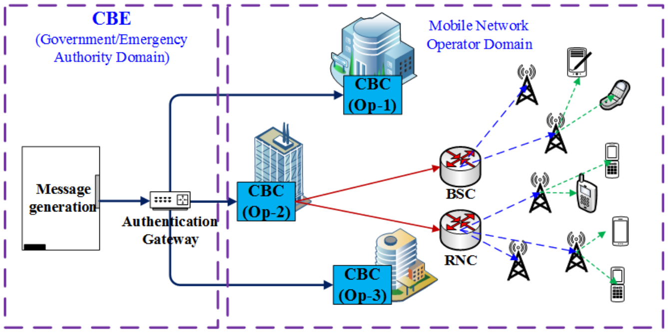


Figure 2: An illustration of a CBS PWS architecture

in USA and many smaller ones currently implemented WEA. WEA is currently supported by nearly all Android/IOS handsets. We assume that `StayTuned` is integrated with the type-2 WEA alert system, which gives updates to the subscribers in case of an imminent threat such as a storm or tornado. As the message exchanges for `StayTuned` take place through the control channel, voice communication is the least affected.

SMS vs CBS: The next question is to decide what messaging system needs to be used for `StayTuned`. There are two alternative messaging system for Public Warning System (PWS), the Short Messaging Service (SMS) and the Cell Broadcast Service (CBS). SMS is primarily used for personal one-to-one messaging solution. However, for a bulk messaging system such as PWS such one-to-one messaging system necessitates that warning messages are sent individually to all users within a certain area. During a disaster situation, the network might be overcrowded [3] and using one-to-one bulk messaging would lead to further network congestion. In addition, the network operators would need to retrieve which phones are present in the affected area and then send individual SMS.

SMS messages are also prone to delivery failure. In [4] the authors observed that in normal operating conditions the SMS message delivery failure rate is as high as 5.1%. The authors observed that under stressful condition such as "flash-crowd" events that occurred in New Year's Eve of 2005, the SMS failure rate varied from 20%-70% due to the eightfold increase in traffic [4]. Such overloaded signaling channels also prevent set-up for the voice calls.

CBS, on the other hand, is a one-to-many broadcast based messaging system. A CBS message consists of 88 octets; 6 of which are used to define the message characteristics, while the remaining 82 are used for carrying payload. A broadcast message can be sent within a few seconds to all the phones in a targeted area, without causing much network congestion. Another key advantage of CBS is its security. Unlike SMS, CBS can only be sent by some authorized organization that has ability to send such messages. Because of these reasons, we choose CBS for our `StayTuned` awareness message exchanges. CBS is also included in current 3GPP 2G, 3G and LTE standards [3].

Network architecture for disseminating CBS: The basic implementation of a CBS, illustrated in Figure 2, consists of cell-broadcast centers (CBCs) and at least one cell-broadcast entity (CBE). CBE resides under a trustworthy or government authority domain. CBE is the message creator that compiles the messages, specifies the locations where they need to be broadcasted and passes them to the CBC operators. The operators then send the cell broadcast messages to the respective targeted areas.

In majority of cases, the CBEs are further divided into two units. The first is a message generation unit that accumulates emergency information from different sources and generates messages that are to be sent. These messages are then sent to the authentication gateway unit that authenticates the sender and validates legitimacy of the messages before sending them to the CBC service providers. To ensure the security and integrity of the messages, the payloads may also be encrypted [3].

In our `StayTuned` system, the messages are collected from different sources, e.g., from different social media. During disasters, the number of collected messages can be overwhelming and redundant. A viable option is to filter and extract a subset of relevant messages from the incoming message stream before forwarding them to the CBCs.

III. STAYTUNED MESSAGE EXTRACTION FRAMEWORK

We propose a two-step message extraction framework for `StayTuned` service as illustrated in Figure 3. The first component is a trainable *filter* which removes irrelevant messages by using a model trained from the current and previous similar disasters. The second component of the message extraction framework is an adjustable *summarizer*, which selects representative messages pertinent to the disaster while removing the redundancies.

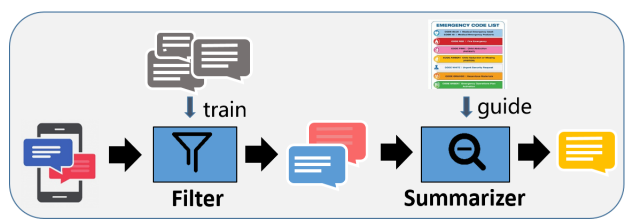


Figure 3: Message extraction framework.

A. Message filtering

During disasters, only disaster-related messages are of interest for `StayTuned`. The goal of message filtering is to discard less relevant messages while keeping the important ones. We formulate the message filtering problem as a text classification task. The problem setup is explained as follows.

Let us assume there is a set of p labeled messages $\mathbf{M}_{train}$ = $\{(m_1, y_1), (m_2, y_2), ..., (m_p, y_p)\}$, where each message m_i is converted into a feature vector $\boldsymbol{x_i} \in \mathbb{R}^n$, and y_i is a binary label indicating whether the message is disaster-related. The objective is to train a model $y = f(\boldsymbol{x})$ whose output can be used to classify messages. We will assume that the output is real-valued and that larger values indicate a stronger likelihood of being disaster-related.

To learn an accurate filter, a large number of labeled messages is preferred. However, during the emerging disaster, the number of labeled messages is likely very small. In our previous work [5], we addressed the small labeled training set problem with the help of a large amount of unlabeled messages from previous disasters, $\boldsymbol{M}_{unl} = \{m_1, m_2, .., m_q\}$, where $q >> p$, in which labels y_i are unavailable. We used *word2vec* [6] algorithm on unlabeled messages to learn how to represent each word as a real valued vector in R^d. *word2vec* is able to represent related words as similar vectors. Given the vector representations, we then used k-means algorithm to group words into k clusters. We list words from several disaster-related clusters in Table I as an illustration. Finally, we represented each message as a bag-of-clusters, which is a count vector of length k. The i-th element of the bag-of-clusters equals the number of words from the message that belong to the i-th cluster. We observed that such representation of messages allows training of accurate classifiers even from very small training sets with a few dozen of labeled messages.

TABLE I
CRISIS-RELATED CONCEPTS LEARNED.

Concept	Words
Flooding	weather, storm, flood, storms, flooding, floods, thunderstorms, #storm, #floods, #flood, #flooding
Power	#mdl, #bier, #stages, #pk, #lb, #sticktogether, #newnorma, #droughts, #powerout
Time	day, time, night, weekend, minutes, min, mins, wknd, offseason, midday, arvo, midnite, w/end, wend,
Verbs	thinking, waiting, telling, hoping, wishing, praying, searching, seeking, needing, demanding

B. Message Summarization

The disaster related messages are likely to be redundant. For example, it is common that multiple people report on important events such as flooding or food shortage. Thus there is an opportunity to further reduce the number of messages by summarization. Let us suppose we are given a set of N messages. We formulate the problem as a subset selection problem following work in [7], [8], where the objective is to select a subset of n messages that maximize the ***informativeness***. In this work, we also maximize ***semantic completeness***, to be defined later in the text.

Informativeness score: Content words can be provided by domain experts, describing disaster situations. In our experiments, we simply use *nouns*, *proper nouns*, *verbs* and *numbers* as content words. Let us denote the set of V unique content words as $\mathcal{CW} = \{cw_1, cw_2, ..., cw_V\}$. Emergency responders at CBE can define a weight for each content word cw_j as $W(cw_j)$. Then, informativeness of message i can be calculated as the sum of the weights of its content words: $I(m_i) = \sum_j^{v_i} W(cw_j)$.

Semantic completeness: While informativeness score is a useful objective criterion for summarization, it is not sufficient to reduce redundancy in messages. For example, suppose the following three messages are given the same informativeness score:

TABLE II
NOTATION USED IN THE SUMMARIZATION STAGE.

Notation	Description
x_i, y_j	Binary values with 1 if the message i or the content word j is selected and 0 otherwise
$I(m_i)$	Informativeness score of message i
$W(cw_j)$	Weight of content word j
L	Maximum number of messages can be selected
$\mathcal{T}(cw_j)$	Indexes of messages of which cw_j is a member
$\mathcal{W}(m_i)$	Indexes of words in message m_i
$\mathcal{C}_k$	Indexes of messages in semantic cluster k
K	Total number of semantic clusters
N	Total number of messages in the pool

(i) `Water has entered my house.`
(ii) `Streets are flooded and not one rain drop came down so far.`
(iii) `Attention everyone I lost power.`

Instead of picking (i) and (ii) out of them, we should select (ii) and (iii), since they refer to two semantic groups and cover two aspects of disaster.

Thus, we propose a preprocessing step that clusters messages. We first represent the N messages as a matrix $R^{N \times V}$, with value at cell (i, j) being the count of word j in message i. Since the count matrix is sparse, we use the *SVD* algorithm to reduce the number of columns to d, $d \ll V$. Then, we apply k-means algorithm to group all messages into K clusters.

The resulting message selection optimization algorithm is described in equation(1). The notation is explained in Table II.

$$\begin{aligned}
\underset{x_i, y_j}{\text{Maximize}} \quad & \sum_{i=1}^{N} I(m_i) x_i + \sum_{j=1}^{V} W(cw_j) y_j \\
\text{subject to} \quad & \sum_{i=1}^{N} x_i < L \\
& \sum_{i \in \mathcal{T}(cw_j)} x_i \geq y_j \\
& \sum_{j \in \mathcal{W}(m_i)} y_j \geq |\mathcal{W}(m_i)| x_i \\
& \sum_{i \in \mathcal{C}_k} x_i \leq \tau \sum_{i=1}^{N} x_i, \text{ for } k \text{ in } 1, 2, ..., K.
\end{aligned} \tag{1}$$

The above optimization maximizes the informativeness and the semantic coverage of the selected messages. The first constraint specifies the budget, which is the number of messages the summarizer wants to extract. The second and third constraints ensure validity of the selected messages, i.e., all of the content words in a selected message are selected and a selected content word must reside in a selected message. The final set of K constraints defines a maximum number of messages that could be selected from any given cluster, which is controlled by parameter $\tau \in (0, 1]$. The formulated problem is solved using GUROBI Optimizer [9].

IV. Experimental Setup

A. Data Sets

We evaluated performance of filtering and summarization in `StayTuned` using real world data. We collected 4.7 million tweets posted during 2012 Sandy Hurricane, which affected the U.S. Northeast from 10/28/2012 to 11/03/2012. We focused on a subset of tweets generated from 4 large metro areas in the region: New York, NY (NYC); Philadelphia, PA (Phila.); Boston, MA (Boston) and Washington, DC (DC).

B. Simulated Settings

Typically, a CBE will divide the targeted area into small *cells* (e.g., school districts). Since we did not have this information handy, we divided each of the 4 cities into a grid of equally sized cells with size $D_{fix} \times D_{fix}$. Note that in this setting the traffic load of each cell is highly imbalanced, because of varying population densities in typical big cities. Figure 4 shows the approximate population density of New York City metro area by considering the number of unique users who tweeted. We assume $D_{fix} = 500$ for Figure 4. The dense regions are denoted by dark color. The figure shows that the population is highly concentrated on the island of Manhattan.

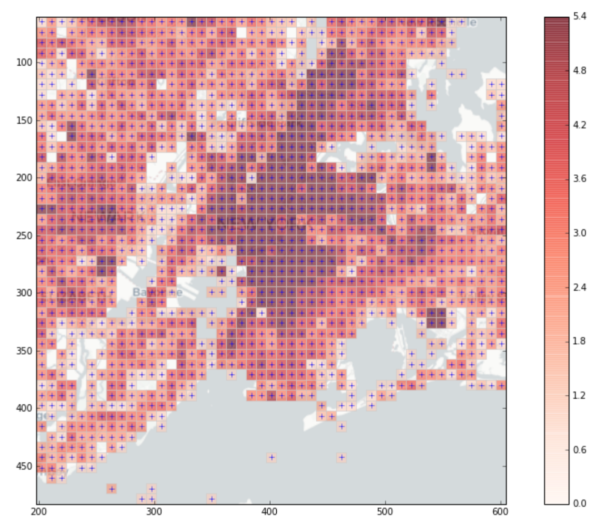
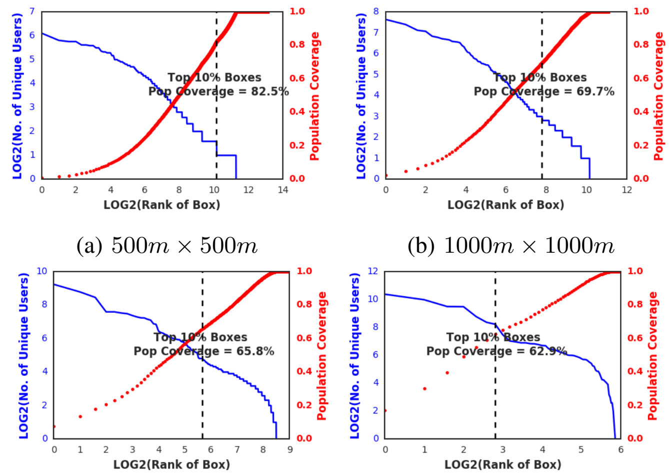


Figure 4: The Manhattan area is partitioned into cells with size $D_{fix} = 500m$. The tweet densities are shown in different color scale.

Dividing a targeted area into small regions results in a large number of cells. For example, since or New York region is of size $48km \times 48km$, we get around 10,000 cells of size $500m \times 500m$. We discovered that the densest 10% of cells cover $\sim 70\%$ of the population. The long tail effect is illustrated in Figure 5. Since for the less dense cells the network congestion is not a significant issue, we only evaluated our filtering and summarization algorithms on the densest 10% of the cells.

Filter: To train the filter we used a labeled dataset M_{train}, obtained from [10]. It contains 3,505 tweets labeled as 1 if they are related to Sandy Hurricane and 0 otherwise. To create the word clusters we downloaded word vectors pre-trained by *wort2vec* algorithm on 400 million unlabeled tweets from the Internet. Then, we created 2,000 clusters using k-means algorithm. We then trained the filter using logistic regression.



(a) $500m \times 500m$

(b) $1000m \times 1000m$

(c) $2000m \times 2000m$

(d) $5000m \times 5000m$

Figure 5: The population density plot for different D_{fix}. X-axis: $\log_2$(rank) of cell by density; Left Y-axis: number of unique users; Right Y-axis: fraction of unique users.

TABLE III
NUMBER (PERCENTAGE %) OF MESSAGES AFTER FILTERING.

	NYC	Phila.	Boston	DC
$\lvert\mathcal{M}\rvert$	$444,016$	$121,122$	$86,731$	$100,508$
$\lvert\mathcal{S}_1\rvert(p(\mathcal{S}_1))$	$16,201(3.65)$	$2,304\ (1.9)$	$1,294(1.5)$	$1,563(1.6)$

Summarizer: The key to the summarizer is the definition of the content words and their weights. In reality, the content words and their weights can be passed from authorities or domain experts to better guide the summarization. For simplicity, we extract content words automatically from all tweets in an area as follows:

1) **Content words:** We first use the tweet POS (part-of-speech) tagger from CMU [11] to tag all of the tweets from a city. Then *nouns*, *proper nouns*, *numbers* and *verbs* with more than 5 appearances are selected as content words.
2) **Weights of content words:** For our experiments, we scored content word cw as: $W(cw) = \log_2(f_{cw} + 1) * \log_2 \frac{N}{|\mathcal{T}(cw)|}$, where f_{cw} is the total count of cw and $\mathcal{T}$ is the number of tweets containing cw.

Another important decision for the summarizer is the message clustering procedure, which helps us achieve a better variety of selected messages. In our experiment we clustered messages in a cell into 10 clusters and removed the smallest 2 clusters.

V. Experimental Results

We evaluated the performance of our proposed message extraction framework for `StayTuned` both quantitatively and

TABLE IV
NUMBER (PERCENTAGE %) OF MESSAGES KEPT AFTER SUMMARIZATION.

	$D_{fix} = 500m$				$D_{fix} = 1000m$			
	$L = 10$	$L = 20$	$L = 50$	$L = 100$	$L = 10$	$L = 20$	$L = 50$	$L = 100$
NYC	2990(0.67)	5277(1.19)	7811(1.76)	8656(1.95)	1250(0.28)	2500(0.56)	5244(1.18)	7149(1.61)
Philla.	689(0.57)	753(0.62)	753(0.62)	753(0.62)	454(0.37)	32(0.03)	664(0.55)	664(0.55)
Boston	392(0.45)	441(0.51)	441(0.51)	441(0.51)	190(0.22)	278(0.32)	278(0.32)	278(0.32)
DC	466(0.46)	510(0.51)	510(0.51)	510(0.51)	224(0.22)	359(0.36)	359(0.36)	359(0.36)

	$D_{fix} = 2000m$				$D_{fix} = 5000m$			
	$L = 10$	$L = 20$	$L = 50$	$L = 100$	$L = 10$	$L = 20$	$L = 50$	$L = 100$
NYC	370(0.08)	740(0.17)	1850(0.42)	3664(0.83)	100(0.02)	200(0.05)	500(0.11)	1000(0.23)
Philla.	128(0.11)	232(0.19)	304(0.25)	304(0.25)	80(0.07)	136(0.11)	208(0.17)	208(0.17)
Boston	80(0.09)	152(0.18)	208(0.24)	208(0.24)	75(0.09)	131(0.15)	211(0.24)	259(0.30)
DC	80(0.08)	152(0.15)	208(0.21)	208(0.21)	48(0.05)	72(0.07)	104(0.10)	112(0.11)

TABLE V
CRISIS RELATED CATEGORIES [12], [13].

Categories	**Topics**
C_1	Injured or dead people
C_2	Missing, trapped, or found people
C_3	Displaced people and evacuations
C_4	Infrastructure and utilities damage
C_5	Donation needs/volunteering services
C_6	Caution and advice
C_7	Sympathy and emotional support
C_8	Other useful information

qualitatively.

A. Quantitative Analysis

We would like to check the percentage of messages kept after filtering and after summarization. Suppose $\mathcal{M}$ is the set of all tweets for a specific city and $\mathcal{S}_1$, $\mathcal{S}_2$ are the remaining tweets after filtering and after summarization respectively, the percentage of messages kept in $\mathcal{S}_i$ is then defined as $p(\mathcal{S}_i) = \frac{|\mathcal{S}_i|}{|\mathcal{M}|} \times 100\%, i \in \{1, 2\}$.

The numbers (percentages) of messages kept after filter are reported in Table III. The fraction of crisis-related tweets is twice larger in NYC compared to other cities, which is expected because NYC was more heavily impacted by Sandy Hurricane.

Then, we checked whether our summarizer works as expected when we vary the cell size D_{fix} and the message budget L. We explored the following limits on the number of messages per cell, $L = [10, 20, 50, 100]$. The cell size D_{fix} ranged from $[500m, 1000m, 2000m, 5000m]$. We fixed $\tau = 0.125$, which encourages about the same number of messages to be selected from each cluster. The numbers and percentages are reported in Table IV. From Table IV we can observe that:

1) As we increase the cell size, we will keep a smaller fraction of messages.
2) As we increase the size of budget, we will keep more messages.

B. Qualitative Analysis

Besides the number of extracted messages, we also seek a good coverage of crisis-related topics.

One standard and widely accepted crisis-related ontology proposed in [12], [13] contains 8 categories (C_1-C_8) as shown in Table V. We measured the distribution of message categories of the filtered tweets and the final summarizations. To get the number of tweets in each category, we took a labeled data set provided by [14], which contains $2,013$ tweets from Odile Hurricane in 2014, and trained a multi-class classifier using the described method for the filter component. Then the classifier was used to predict the category of tweets from Sandy Hurricane.

We first get the set of tweets after filtering $\mathcal{S}_1$ for each city from Table III. We also get a set of tweets after summarization $\mathcal{S}_2$ for each city using $L = 50$, $D_{fix} = 2000m$ and $\tau = 0.125$ from Table IV. Therefore, the percentage of category k in $\mathcal{S}_i$ is defined as: $p_{\mathcal{S}_i}(k) = \frac{n_k}{|\mathcal{S}_i|} \times 100\%, i \in \{1, 2\}$, where n_k is the number of tweets in $\mathcal{S}_i$ falling in category k. Table VI shows how the distribution of message categories changes from $\mathcal{S}_1$ to $\mathcal{S}_2$.

One observation is that the eight categories are not evenly distributed. A small proportion of tweets is classified into categories 1, 2, 3 and 6. It might be because the death or injury of people were rare events. Fair amount of tweets was related to infrastructure damage, donations, and words of sympathy. Many tweets were classified as other useful information.

Comparing the distribution of message categories in $\mathcal{S}_1$ and that in $\mathcal{S}_2$, we do not see dramatic differences. However, the percentage of tweets of category 8 is reduced by $5\% - 10\%$, while the percentages of other categories are lifted slightly. The results illustrate that the summarizer reduced the imbalance slightly.

Finally, we set $L = 5$, $D_{fix} = 1000m$, $\tau = 1$ (by which we do not force the clustering constraints) for all cities and apply our message extraction framework on the most dense cell in each city. The most important 5 messages from the 4 cities are shown in Table VII. The two numbers in the parentheses denote the number of selected tweets and the total number of tweets, respectively. During Sandy Hurricane, the Manhattan area was considerably affected. Therefore, damage and disruptions, such as flight delay, power outage, were commonly reported. Users from the other three cities were worried about power outage and were sending prayers to the people along the eastern shore, but there is much less reporting

TABLE VI
DISTRIBUTION OF MESSAGE CATEGORIES AFTER FILTERING ($\mathcal{S}_1$) & AFTER SUMMARIZATION ($\mathcal{S}_2$).

		NYC	Phila.	Boston	DC
$\mathcal{S}_1$	C_1	0.00	0.00	0.00	0.00
	C_2	0.09	0.04	0.00	0.00
	C_3	0.01	0.00	0.00	0.00
	C_4	4.44	3.04	3.86	3.58
	C_5	3.11	2.95	3.01	3.26
	C_6	0.01	0.00	0.00	0.00
	C_7	4.30	7.29	7.57	5.37
	C_8	84.67	83.29	81.07	84.52
$\mathcal{S}_2$	C_1	0.00	0.00	0.00	0.00
	C_2	0.15	0.00	0.00	0.00
	C_3	0.00	0.00	0.00	0.00
	C_4	4.46	2.96	7.21	5.29
	C_5	4.61	4.61	4.33	6.25
	C_6	0.00	0.00	0.00	0.00
	C_7	4.76	12.83	8.17	7.21
	C_8	80.95	73.03	75.48	79.33

of the damage and disturbances.

VI. RELATED WORK

Smartphone for emergency applications: The increasing penetration of smartphones now-a-days makes them a platform for various emergency applications such as health monitoring, mobile sensing, or building ad-hoc networks in crisis scenarios. Several network testbeds and protocols have been developed for search and rescue in disasters, including `VigiNet` [15], `AlarmNet` [16], `Code Blue` [17], `DieselNet` [18], and `DistressNet` [19]. `TeamPhone` [20] uses a self-rescue system for the trapped victims during disaster using WiFi-Direct communication. `E-DARWIN` [21], [22] proposes a WiFi-Tethering based disaster recovery network formation using wireless smartphones that are largely available in the disaster-hit areas. `Argus` [23] builds a 3D mapping of disaster scene using cooperative data collection from by-standers/drones present around the targeted zone to facilitate rescue operations. Unlike others, the proposed `StayTuned` provides a service for distributing filtered and summarized updates available from different information sources to the subscribers. This unique aspect of this paper is integrate the use of existing cellular infrastructure and analytics of big data obtained from various social media to create an agile emergency update service that is expected to provide significant additional value in future emergencies and disasters.

Twitter for situational awareness: Recent years have seen an increased interest by research community in studying how to exploit Twitter data for situational awareness in the emergency and disaster contexts [24], [25]. Event detection is arguably the most active research topic, where the objective is to detect new events from a real-time twitter stream. A typical approach for event detection is to define one or a few keywords (e.g., earthquake) of interest and to track online if there are temporal bursts of the keywords' use in tweets [26]. Extensions of this approach include general-purpose detection systems that track a large number of keywords [27], phrases [28] or emergence of clusters of similar tweets [29].

TABLE VII
THE 5 SELECTED MESSAGES FROM EACH CITY.

NYC (5/365)
`Flight out on Monday is cancelled, looks like flying today is not an option. Got a hotel reservation here in NY until Thur. // New Jersey will be under water for weeks after this storm and Chris Christie is "going to watch the jets for a couple of hours" // 57th Street closed off. Times Square like a ghost town. Sandy has sent "The City That Never Sleeps" into La La Land. // Wow! RT BREAKING -- Reuters witness: 19 Con Edison workers trapped by rising floodwaters at Manhattan power station.? // Gov Cuomo- subway system damaged heavily. A 30 year vet said hes never seen such damage. East river flooded into ground zero #sandy // lower manhattan power outage Midtown Manhattan`
Phila. (5/56)
`Happy Sunday! I'm in Philadelphia for a trade show, Hurricane Sandy is approaching.I wonder how bad this storm of the century will be? ????? // Ummm pool party at Joani's??? Let's play games by candle light whether we have power or not. // It's been a crazy day on the east coast... Some of the footage of the shore is insane!! I hope you're on high ground! // Turned off gps and data, turned on wifi to conserve battery in case the power goes out tonight. The first real precaution I've taken. #Sandy // 2 minutes of #sandy from CC Philly Monday night #6abcSandy http://t.co/HE2q37jy`
Boston (5/49)
`"@SpeedoUSA: Sending our thoughts to those on the East Coast affected by #sandy #frankenstorm" // yea my dad says he hasn't seen wind this bad since hurricane bob in 91' plus it got worse at high tide // The only thing worse than Hurricane Sandy is watching Piers Morgan talk about. #sandy // Local news decided to stop weather broadcast in favor of Dancing w/the Stars, you know that signals the end of the storm here in Beantown // So no power outage in the back bay == karma from last year's senseless 3 day outage. I guess we're even now.`
DC (5/42)
`Braving the storm while the supermarket is still open. If I don't tweet for a while, send help! #SandyDC // @bethanyshondark The hype was right! Hoping I don't lose power in DC... // @ckbarrett Still raining and freezing, and most everything is shut down. Alex made it home safely this morning from his night shift. #sandy // Offices are fine, no water damage, electricity and Internet... What else could we possibly need (@ Clearly Innovative) http://t.co/GJqp21cl // There are very few times I wish I was allowed to donate blood. Right now is one of those times. #sandy`

More recently, researchers started paying more attention to the spatial aspect of events [30]. For example, [31] considers burstiness of term "earthquake" in both time and space to detect spatial clusters of tweets with that word as candidates for the earthquake event. The unsupervised approach for event detection can be further enhanced by adding a classifier that is trained on previous events to recognize which bursty clusters are events and which are not [32].

Summarizing and visualization of disaster-related tweets help human responders to quickly grasp the vast amounts of generated information. Representative systems are `Senseplace2` [33], a visual analytics system that allows an operator to enter a query (in a form of a term or a hashtag), look at the map to observe where is the keyword common, click on a specific location, and view individual ranked tweets from the selected location, and `Twitinfo` [34], a tool that allows an operator to browse a large collection of tweets using a timeline-based display, drill down to subevents, and explore via geolocation, sentiment, and popular URLs.

VII. Conclusion

In this paper, we envision situation awareness mobile service via automated data collection from the smart phone users and human directed communication such as phone calls and social media. Acknowledging that Twitter has established itself as the premier human communication mechanism during disasters and the wealth of publicly accessible disaster-related twitter data, we considered integration of only the twitter-based information for the purposes of situational awareness. The key objective was to understand how to utilize twitter data for updating the subscribers regarding the important updates by leveraging the previous works on event detection and situational awareness from Twitter data. As such, our proposed work automatically selects, orders and summarizes Twitter data at a large scale, which can be deployed in real-time during disasters via the message broadcasting service `StayTuned`.

Acknowledgments

This research was supported by JST-NSF joint funding, Strategic International Collaborative Research Program, SICORP on Japan side, and CNS-1461932 award from NSF.

References

[1] L. Palen, K. Starbird, S. Vieweg, and A. Hughes, "Twitter-based information distribution during the 2009 red river valley flood threat," *Bulletin of the American Society for Information Science and Technology*, vol. 36, no. 5, pp. 13–17.

[2] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: What twitter may contribute to situational awareness," in *CHI*, 2010, pp. 1079–1088.

[3] "Mobile network public warning systems and the rise of cell-broadcast," https://www.gsma.com/mobilefordevelopment/programme/mobile-for-humanitarian-innovation/mobile-network-public-warning-systems-and-the-rise-of-cell-broadcast/.

[4] X. Meng, P. Zerfos, V. Samanta, S. H. Y. Wong, and S. Lu, "Analysis of the reliability of a nationwide short message service." in *IEEE INFOCOM*, 2007, pp. 1811–1819.

[5] S. Zhang and S. Vucetic, "Semi-supervised discovery of informative tweets during the emerging disasters," *arXiv preprint arXiv:1610.03750*, 2016.

[6] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[7] K. Rudra, S. Banerjee, N. Ganguly, P. Goyal, M. Imran, and P. Mitra, "Summarizing situational and topical information during crises," *arXiv preprint arXiv:1610.01561*, 2016.

[8] ——, "Summarizing situational tweets in crisis scenario," in *ACM HT*. ACM, 2016, pp. 137–147.

[9] G. O. Inc, "Gurobi optimizer reference manual," 2014, http://www.gurobi.com. [Online]. Available: http://www.gurobi.com

[10] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "Crisislex: A lexicon for collecting and filtering microblogged communications in crises." in *ICWSM*, 2014.

[11] K. Gimpel, N. Schneider, B. O'Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, and N. A. Smith, "Part-of-speech tagging for twitter: Annotation, features, and experiments," in *ACL*. Association for Computational Linguistics, 2011, pp. 42–47.

[12] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys (CSUR)*, vol. 47, no. 4, p. 67, 2015.

[13] M. Imran, P. Mitra, and C. Castillo, "Twitter as a lifeline: Human-annotated twitter corpora for nlp of crisis-related messages," *arXiv preprint arXiv:1605.05894*, 2016.

[14] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, "Aidr: Artificial intelligence for disaster response," in *ACM WWW*, 2014, pp. 159–162.

[15] T. He, S. Krishnamurthy, L. Luo, T. Yan, L. Gu, R. Stoleru, G. Zhou, Q. Cao, P. Vicaire, J. A. Stankovic, T. F. Abdelzaher, J. Hui, and B. Krogh, "Vigilnet: An integrated sensor network system for energy-efficient surveillance," *ACM Transactions on Sensor Networks*, vol. 2, no. 1, pp. 1–38, 2006.

[16] A. Wood, J. Stankovic, G. Virone, L. Selavo, Z. He, Q. Cao, T. Doan, Y. Wu, L. Fang, and R. Stoleru, "Context-aware wireless sensor networks for assisted living and residential monitoring," *IEEE Network*, vol. 22, no. 4, pp. 26–33, 2008.

[17] K. Lorincz, D. J. Malan, T. R. Fulford-Jones, A. Nawoj, A. Clavel, V. Shnayder, G. Mainland, M. Welsh, and S. Moulton, "Sensor networks for emergency response: Challenges and opportunities," *IEEE Pervasive Computing*, vol. 3, no. 4, pp. 16–23, 2004.

[18] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "Maxprop: Routing for vehicle-based disruption-tolerant networks." in *Proc. of IEEE INFOCOM*, 2006.

[19] S. George, W. Zhou, H. Chenji, M. Won, Y. O. Lee, A. Pazarloglou, R. Stoleru, and P. Barooah, "Distressnet: a wireless ad hoc and sensor network architecture for situation management in disaster response," *IEEE Communications Magazine*, vol. 48, no. 3, pp. 128–136, 2010.

[20] Z. Lu, G. Cao, and T. L. Porta, "Teamphone: Networking smartphones for disaster recovery," *IEEE Transactions on Mobile Computing*, vol. 16, no. 12, pp. 3554–3567, 2017.

[21] A. Pal and K. Kant, "E-darwin2: A smartphone based disaster recovery network using wifi tethering," in *IEEE CCNC*, 2018, pp. 1–5.

[22] A. Pal, M. Raj, K. Kant, and S. Das, "A smartphone based network architecture for post-disaster operations using wifi tethering," in submission. [Online]. Available: https://astro.temple.edu/~tuf79640/TETC_paper-v4.pdf

[23] V. Sadhu, G. Salles-Loustau, D. Pompili, S. A. Zonouz, and V. Sritapan, "Argus: Smartphone-enabled human cooperation for disaster situational awareness via MARL," in *PerCom Workshops*, 2017, pp. 79–81.

[24] Y. Wu, K. Kant, S. Zhang, A. Pal, and J. Wang, "Disaster network evolution using dynamic clustering of twitter data," in *IEEE ICDCS Workshops*, 2017, pp. 348–353.

[25] Y. Wu, A. Pal, K. Kant, and J. Wang, "Incremental spatial clustering for spatial big crowd data in evolving disaster scenario," in submission. [Online]. Available: https://astro.temple.edu/~tuf79640/ISC_Decay_Model-2.pdf

[26] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: Real-time event detection by social sensors," in *ACM WWW*, 2010, pp. 851–860.

[27] M. Mathioudakis and N. Koudas, "Twittermonitor: trend detection over the twitter stream," in *ACM SIGMOD*, 2010, pp. 1155–1158.

[28] C. Li, A. Sun, and A. Datta, "Twevent: segment-based event detection from tweets." in *CIKM*, 2012, pp. 155–164.

[29] S. Petrović, M. Osborne, and V. Lavrenko, "Streaming first story detection with application to twitter," in *Human Language Technologies*, 2010, pp. 181–189.

[30] X. Dong, D. Mavroeidis, F. Calabrese, and P. Frossard, "Multiscale event detection in social media," *Data Min. Knowl. Discov.*, vol. 29, no. 5, pp. 1374–1405, 2015.

[31] D. C. B. P. S. Earle and M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world," *Annals of Geophysics*, vol. 54, no. 6, pp. 708–715, 2012.

[32] K. Xie, C. Xia, N. Grinberg, R. Schwartz, and M. Naaman, "Robust detection of hyper-local events from geotagged social media data," in *ACM MDMKDD*, 2013.

[33] A. M. MacEachren, A. R. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford, "Senseplace2: Geotwitter analytics support for situational awareness." 2011, pp. 181–190.

[34] A. Marcus, M. Bernstein, O. Badar, D. Karger, S. Madden, R. Miller, and O. Arenz, "Twitinfo: Aggregating and visualizing microblogs for event exploration," in *Annual conference on Human factors in computing systems*, 2011, pp. 227–236.