

Thematically Analysing Social Network Content During Disasters Through the Lens of the Disaster Management Lifecycle

Sophie Parsons
Electronics and Computer Science,
University of Southampton, UK
sp13g10@southampton.ac.uk

Peter M. Atkinson
Geography and Environment
University of Southampton, UK
p.m.atkinson@southampton.ac.uk

Elena Simperl and Mark Weal
Electronics and Computer Science,
University of Southampton, UK
{e.simperl,mjw7}@soton.ac.uk

ABSTRACT

Social Networks such as Twitter are often used for disseminating and collecting information during natural disasters. The potential for its use in Disaster Management has been acknowledged. However, more nuanced understanding of the communications that take place on social networks are required to more effectively integrate this information into the processes within disaster management. The type and value of information shared should be assessed, determining the benefits and issues, with credibility and reliability as known concerns. Mapping the tweets in relation to the modelled stages of a disaster can be a useful evaluation for determining the benefits/drawbacks of using data from social networks, such as Twitter, in disaster management.

A thematic analysis of tweets' content, language and tone during the UK Storms and Floods 2013/14 was conducted. Manual scripting was used to determine the official sequence of events, and classify the stages of the disaster into the phases of the Disaster Management Lifecycle, to produce a timeline. Twenty-five topics discussed on Twitter emerged, and three key types of tweets, based on the language and tone, were identified. The timeline represents the events of the disaster, according to the Met Office reports, classed into B. Faulkner's Disaster Management Lifecycle framework. Context is provided when observing the analysed tweets against the timeline. This illustrates a potential basis and benefit for mapping tweets into the Disaster Management Lifecycle phases. Comparing the number of tweets submitted in each month with the timeline, suggests users tweet more as an event heightens and persists. Furthermore, users generally express greater emotion and urgency in their tweets.

This paper concludes that the thematic analysis of content on social networks, such as Twitter, can be useful in gaining additional perspectives for disaster management. It demonstrates that mapping tweets into the phases of a Disaster Management Lifecycle model can have benefits in the recovery phase, not just in the response phase, to potentially improve future policies and activities.

Categories and Subject Descriptors

K.4.0. [COMPUTERS AND SOCIETY]: General

General Terms

Management, Human Factors, Theory

Keywords

Disaster Management, Thematic Analysis, Disaster Management Lifecycle, Situation Awareness, Twitter, Social Networks, Natural Disaster.

1. INTRODUCTION

Disaster management involves handling, organising and controlling the events and activities in preparation and response to disasters [Alexander 2005]. A disaster can be defined as "a singular, large scale, high impact extreme event" [Cutter 2003]. They often result in severe repercussions, requiring unordinary response [Alexander 2005]. Natural disasters are provoked by geological, hydro-meteorological or biological factors [Asmatullah and Himayatullah 2008]. Examples include earthquakes, cyclones and floods.

During a natural disaster, people increasingly use microblogging and social networking sites, to share and retrieve information in real-time; an imperative requirement of communication during a disaster [Vieweg 2010; Kongthon et al. 2012]. Microblogging is a type of online communication that enables users to post short entries quickly. Twitter, a social network, is a popular form of microblogging [Lie et al. 2014].

This paper analyses Twitter data in relation to the phases of the Disaster Management Lifecycle (DML) through an exemplar case study of flooding in the UK in 2013/14. It focuses on analysing the content, language and tone of tweets posted during a disaster. The study classifies the Met Office event reports from the period under examination into the phases of the DML, to produce a timeline. It considers whether mapping the analysed tweets to this timeline provides new insights that might be beneficial to Disaster Management.

Relevant literature is discussed in section 2. In section 3 the methodology used is described, including a description of the UK storms and floods 2013/14 case study. This is followed by results in section 4 and a discussion in section 5. Finally the paper concludes in section 6.

2. BACKGROUND

Disaster Management is improved by utilising strategies and frameworks, such as the Disaster Management Lifecycle (DML). The DML models the stages of a disaster, and highlights the operations that take place within each phase [Faulkner 2001].

Although some activities are unique to their particular stage, it is not a series of events; phases can overlap. Furthermore, the timescales of each phase vary depending on the type and severity of the disaster [Asmatullah and Himayatullah 2008]. A generic representation of the DML encompasses three predominant phases; Preparedness, Response and Recovery. However, it is commonly adapted. Faulkner [2001] describes a DML framework containing six phases; Pre Event (action taken to prevent or mitigate potential disasters), Prodromal (when it is apparent a disaster is imminent), Emergency (initial effects of the disaster is felt), Intermediate (short-term needs have been addressed; the main focus is to restore to normal), Long Term (activities that take longer to resolve) and Resolution (routine restored and reviewing takes place).

Communication is essential throughout the DML, particularly during the Response Phase. Yet, Lie et al. [2014] state that the public tends to feel dissatisfied with the information from official sources. Past case studies illustrate the public's preference of accessing information regarding a disaster through the use of the Web. Following the World Trade Centre and Pentagon Attacks in 2001, the number of Web searches amplified; during the Indian Ocean tsunami of 2004, the use of photo-repository sites dramatically increased; and Social Networks were the main source for obtaining news updates during the Virginia Tech Shooting 2007 [Starbird and Palen 2010]. Authors, Bruns and Stieglitz [2012], identified that users generally share key emergency response information, and are less interested in posting comments or statements leading up to a potential event. This demonstrates the value of information from Microblogs during the response of a disaster [Terpstra et al. 2012]. It can also help to shape future innovations, policy and practice [Palen et al., 2011]. This has inspired the development of the research domain Crisis Informatics. Combining the study of Disasters and Computing, it aims to discover how together, they can assist in mitigating disasters' effects and impacts [Palen et al. 2007].

Vieweg et al. [2010] highlight the potential of gaining situational awareness features for qualitatively analysing tweets submitted during a disaster. Situation Awareness is defined as "tactical, actionable information that can aid people in making decisions, advise others on how to obtain specific information from various sources, or offer immediate post-impact help to those affected by the mass emergency" [Imran et al. 2013]. However, credibility and reliability are concerns in using social networks in Disaster Management. Esoteric language and grammar, message length and locale-specific references can cause difficulty in ensuring useful and relevant information is aggregated [Kieryev et al. 2009].

Thus, there is a need for gaining a theoretical understanding of the communications on microblogs and social networks to minimise the problem [Vieweg 2010]. It is also imperative to understand the nature of information-sharing activities, prior to incorporating the use of microblogging and social networking sites as a standard in Disaster Management [Palen et al. 2011].

3. METHODOLOGY

Thematic analysis and manual scripting were the methodologies employed to conduct this study. This required Twitter data and the official Met Office reports of the event. Several steps were taken to obtain an adequate sample of the Twitter data and acquire the Met Office reports and execute the analysis.

3.1 Case Study

The case study selected was the recent flooding disaster in the UK. October 2013, begun with St Jude's Storm¹. In December 2013 a further storm occurred, which was followed by numerous bouts of stormy weather; this was the stormiest weather the UK had seen since 1969. Following the New Year 2014, there was further damage and flooding. Without having fully recovered from the previous storms, January 2014 was the wettest January for over 100 years. With a short respite at the end of January, further storms were experienced during February 2014. Continuous rainfall caused rivers to burst their banks, hundreds of people left without power, railways destroyed, infrastructure damaged, and many people forced to evacuate their homes. As conditions worsened, recovery actions were prohibited. Late February finally saw the crisis subside. Major recovery and resolution actions were required to restore the UK back to normality [Met Office 2014].

3.2 Obtaining Twitter Data

Twitter sourced the data used for part of this study. 'Tweets'; messages posted on Twitter, consist of a variety of attributes including the message content, a date/time stamp and a user ID (representing the author of that message). This study solely focuses on the message content of the tweet and its date/time stamp.

To collect a sample relevant to the case study and with minimal noise, the term "flooding" and hashtag "#floods" were used to gather tweets from Twitter. It should be noted that the term and hashtag used are by no means the only ones representing the entire conversations about the UK floods, however, the scale of the project required that the number of tweets be limited.

The tweets' date/time stamp was also constrained to the three months; December 2013; January 2014; and February 2014. This was a critical period of the crisis and represents the time before, during and immediately following the disaster.

With an approximate 8.8 million tweets, the sample required further reduction to enable the manual coding process to be executed. To accomplish this, a script was created using the programming language PyMongo. The script consisted of a regular expression, searching for tweets containing a specified set of case insensitive characters, 'flood' and either 'uk or england', in the message content. This provided a final sample size of 638 tweets. However, after an initial parsing by eye, 34 tweets were classed as 'off-topic,' due to their irrelevance to the case study, for example:

"I'm flooding all my social networks this weekend and spring break like fuk it"

3.3 Thematic Analysis

Thematic Analysis, a qualitative methodology, was employed to explore peoples' engagement with the disaster on Twitter. It takes a bottom-up approach, involving the formation of three levels; 'Basic Themes', 'Organised Themes' and a 'Global Theme' [Thomas 2003]. Due to the purposive nature of this study, it was quickly apparent the global theme would be UK Storms and Floods 2013/14.

¹ BBC NEWS. (2014) 10 key moments of the UK winter storms. <http://www.bbc.co.uk/news/uk->

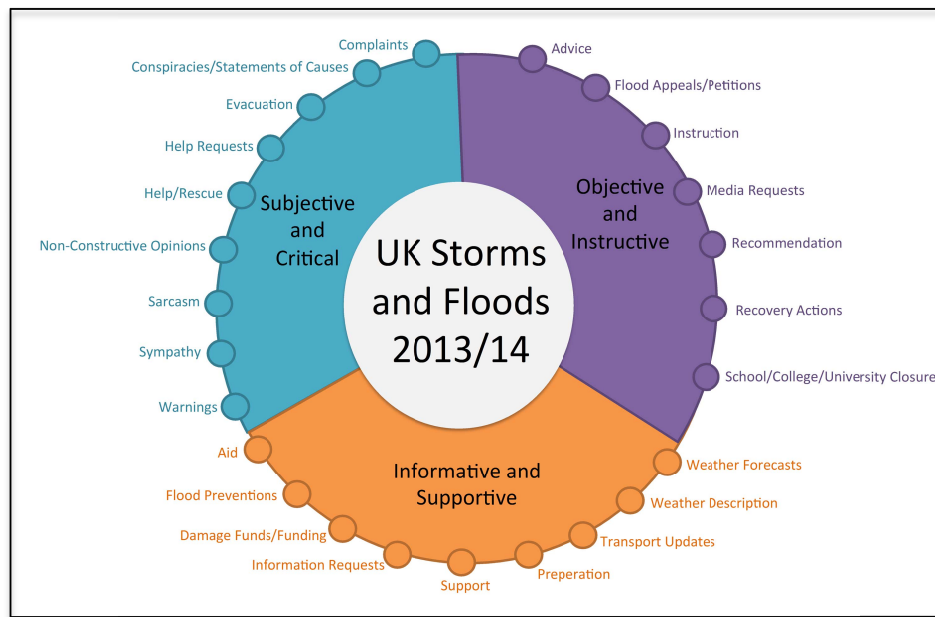


Figure 1. A Visualisation of the Basic, Organised and Global Themes.

Using a sample of tweets, keywords facilitated the identification of the basic themes. During this process, a variety of questions were asked such as “What are the main keywords of the tweet?”, “How was the tweet conveyed?” and “What is the main piece of information to take away from the tweet?” This helped to understand the information each tweet was attempting to communicate, and thus classify it into the category deemed most appropriate. It should be noted that in this study, a ‘retweet’ (a reposting of another user’s tweet), was treated independently. It is also important to state that for this study a tweet was only grouped into one category. Using this methodology, an identification of the situation awareness features communicated was obtained.

Finally, the categories were grouped into organising themes, corresponding to the type of tweets made in terms of language and tone. According to Verma [2011], identifying the type of tweets in this way can be useful for determining the value of the information.

3.4 Timeline Construction

Manual Scripting was used to form an understanding and visualisation of the sequence of events based upon official sources. The Met Office Report ‘UK Storms and Floods 2013/14’ [Met Office 2014], the Met Office Archives and BBC News links, were analysed. Through selecting the dates and surrounding information from the three sources, the significant stages of the disaster were identified. Using the details of the dates extracted from the reports, the dates were categorised into the phases of Faulkner’s [2001] DML.

4. RESULTS

Having conducted a Thematic Analysis and constructed an event timeline, the results illustrate the content, language and tone of tweets posted during the event, the timeline of the event classified into the phases of the DML, and the emerging patterns between the analysed tweets and the timeline.

4.1 Content Analysis

In the sample of tweets obtained, February had the highest percentage of tweets; 65%. Although the first impacts of the disaster were felt during December, it received the lowest percentage of tweets. However, there was only 1% difference between December and January.

Each tweet in the sample was independently analysed, identifying the messages they were attempting to communicate. Examples of tweets include:

*“FLOOD WARNING and this evenings tide times here:
<http://t.co/9NjtZTSMF> #cornwallstorm #cornwall #scilly
#flooding #ukstorm #coast”*

*“If you’ve been affected by #flooding, you can call the
@EnvAgency Floodline 24/7 for advice on 03459881188
#ukstorm #floodaware”*

*“RT @OliversSherbs: People. The roads are rapidly flooding
around here. If you are not at home please try to get there asap
#ukstorm”*

After several passes through of the data, the basic, organised and global themes emerged, shown in Figure 1. The 25 basic themes represent the users’ topics of interest during the disaster. These categories also illustrate the types of situation awareness features that were available during the UK Storms and Floods 2013/14. The organised themes were determined, based upon both the language and tone of the tweets.

Here, ‘Subjective and Critical’ type tweets are defined as personal, and/or provided urgent information/statements, such as an expression of sympathy or a statement of imminent flood warnings. The group ‘Informative and Supportive’ represent tweets providing useful information regarding the disaster, portrayed with a supportive tone. The ‘Objective and Instructive’ class corresponds to any tweet that is specific to the disaster, but

does not present any personal attachment, and/or provides instructive or helpful information succinctly.

The total number of tweets in each category varied. This is shown visually in Figure 2, where the size of the category text is indicative of the quantity of Tweets in that category. The results demonstrate that information regarding ‘Weather Descriptions’ and ‘Flood Appeals/Petitions’ were the most popular topics discussed. Users conversed little regarding ‘School/College/University Closures’ and ‘Help Requests.’

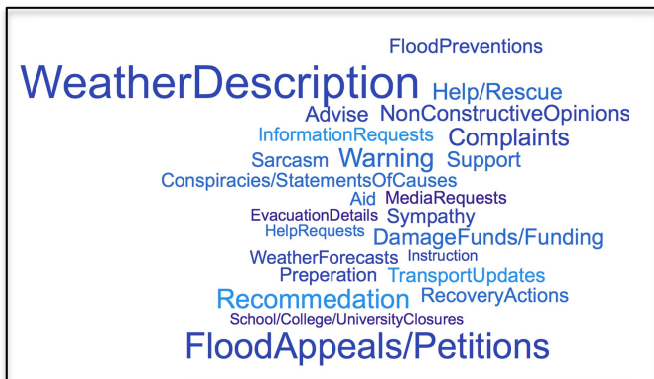


Figure 2. Visualisation of the number of tweets in each category.

4.2 The Events Timeline

Dates explicitly stated were extracted from the Met Office sources, forming the timeline of the UK Storms and Floods 2013/14, see Figure 3. It shows that six specific dates were referred to for the month of December. January had a considerable increase in the number of dates mentioned: eleven, and for February, twelve dates were recorded. The timeline also maps the categories of the DML model that each date was classified in (the different block colours in the visualisation each represent a category of the model; see the key for details).

During the three months, the dates were only classified into three of the six phases of the model: Prodromal, Emergency and Intermediate. However, the disaster was predominantly classed in the Emergency phase.

According to the timeline, the disaster ran through several rotations of parts of Faulkner’s DML model. Generally the dates alternated between the Prodromal phase and the Emergency

phase, however, the crisis also moved back to the Prodromal phase in the beginning of February after an Intermediate phase at the end of January. Only the months of January and February saw stages of the crisis classified into the Intermediate phase.

This timeline aids in the contextualising of the use of Twitter during the disaster.

4.3 Tweet Types Across Time

Figure 4 illustrates the use of the different tweet types across the three months. The ‘Subjective and Critical’ class dominated the type of tweets made in December. However, January and February saw the categories beginning to level out.

By February, the percentage of ‘Subjective and Critical’ tweets dramatically dropped, with ‘Informative and Supportive’ tweets being the most popular type, although ‘Objective and Instructive’ type tweets were never the most popular class of tweets, as the event continued, the percentage of tweets in this category increased.

5. DISCUSSION

The outcomes of the thematic analysis and timeline mapping highlight three insights regarding the use of Twitter data in Disaster Management.

5.1 Twitter Use Variation across the Lifecycle

The sample of tweets used in this study implies the interest of submitting tweets regarding the event heightens as the disaster persists and magnifies. Additionally, the language and tone alters significantly across the three months. Considering this, and the predominant categorisation of dates falling in the Prodromal or Emergency phases, it could be suggested that a link is apparent between the types of tweets submitted, and the phase of the disaster.

5.2 Users’ Motivation for Using Twitter

Users tend to express emotion and a sense of urgency in their tweets during a disaster. The results illustrated a difference in the number of categories assigned to the different organised themes. ‘Subjective and Critical’ and ‘Informative and Supportive’ encompasses nine basic themes, whereas the ‘Objective and Instructive’ type of tweets only contained seven basic themes. This suggests Twitter is predominantly used for providing urgent information and expressing support during a disaster.

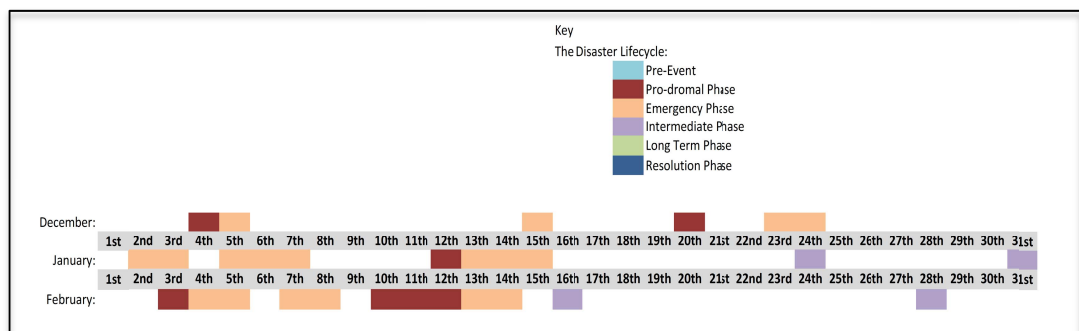


Figure 3. Timeline of the UK Storms and Floods 2013/14, Classified in the Appropriate DML Phases.

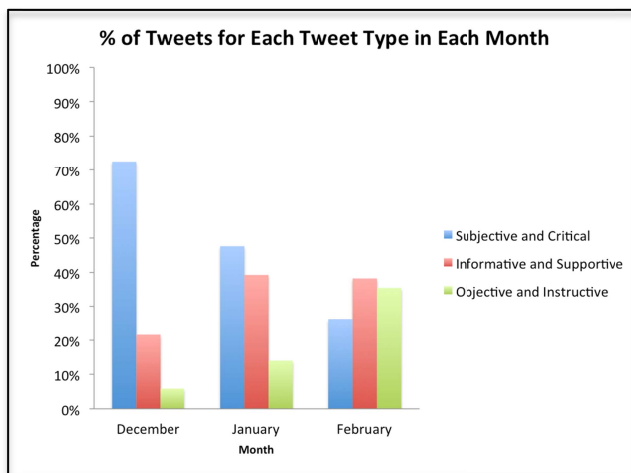


Figure 4. Percentage of Tweets for Each Tweet Type in Each Month.

5.3 Benefits for Disaster Management

It can be beneficial to incorporate an analysis of social networks in disaster management at more than one stage of the Lifecycle. The thematic analysis has demonstrated the potential for extracting situation awareness features, such as evacuations, and identifying unhelpful categories of tweets, for example conspiracies/statements of causes, submitted during an event.

Discovering key categories such as weather descriptions, in real-time, can be vital information during the Response phase of the DML. Descriptions regarding the severity of the event can help highlight areas to focus on, and assist with assigning activities and resources.

Analysing the weightings of the categories can be valuable during the review of the event, typically carried out in the Recovery phase. Several tweets were flood appeals and petitions. This indicates that Twitter could reveal areas requiring action to better prepare for possible future events. Furthermore, it highlights categories users conversed little on, for example School/College/University Closures. Thus, it could be concluded that Twitter emphasises areas of Response and Recovery requiring more attention to enhance Disaster Management.

6. CONCLUSION

This type of study can be beneficial for future works in Crisis Informatics; qualitative analysis can improve data interpretation. To conclude, this study has discussed a qualitative analysis of content from the social networking site, Twitter, illustrating the benefits of its use in disaster management, and the potential use of mapping tweets according to the defined phases of the DML.

Situation awareness features shared on Twitter were identified. This is useful for the Response phase of the DML, and could potentially support the functional requirements for automatic classification systems.

Much of the literature focuses on the use of social networks in the Response phase of a disaster, however, this study has also demonstrated the potential benefits of reviewing social networks in the Recovery phase. The results suggest analysing tweets could highlight areas to focus on to enhance future policies and practices. Finally, it has demonstrated an insight to people's changing behaviour and reactions across the timeline of an event.

This could be valuable to disaster managers as it provides further context on the tweets posted.

7. ACKNOWLEDGMENTS

Thank you to Ramine Tinati (Research Staff at University of Southampton) for assisting with the collection of the Twitter data.

8. REFERENCES

- [1] Alexander, D. 2005. Towards the development of a standard in emergency planning. *Disaster Prevention and Management*. 14, 2. 158–175.
- [2] Asmatullah, V., & Himayatullah, K. 2008. Disaster Management Cycle - A Theoretical Approach. *Management and Marketing – Craiova*. 1, 43–50.
- [3] Bruns, A., and Stieglitz, S. 2012. Quantitative Approaches to Comparing Communication Patterns on Twitter. *Journal of Technology in Human Services*. 30, 3-4. 60-185.
- [4] Cutter, S. L. 2003. Review Article GI Science. *Disasters, and Emergency Management*. 7,4. 439–445.
- [5] Faulkner, B. 2001. Towards a Framework for Tourism Management Disaster. *Tourism Management*. 22,2. 135-147
- [6] Imran., M. et al. 2013. Extracting Information Nuggets from Disaster-Related Messages in Social Media. *Proceedings of the 10th International ISCRAM Conference – Baden- Baden, (Germany, May. 2013)*, 1-10.
- [7] Kireyev, K., Palen, L., and Anderson, K. 2009. Applications of Topics Models to Analysis of Disaster-Related Twitter Data. *NIPS Workshop on Applications for Topic Models: Text and Beyond, December 2009*. [Online] Available from: http://www.umiacs.umd.edu/~jbg/nips_tm_workshop/15.pdf
- [8] Kongthon, A. et al. 2012. The Role of Twitter During a Natural Disaster: Case Study of 2011 Thai Flood 2012. *Proceedings of PICMET '12: Technology Management for Emerging Technologies*. 2227 - 2232.
- [9] Li, J., Vishwanath, A., and Rao, R. 2014. Retweeting the Fukushima Nuclear Radiation Disaster. *Communications of the ACM*. 57,1. 78-85
- [10] Met Office. 2014. The Recent Storms and Floods in the UK. *The Centre for Ecology & Hydrology, Natural Environment Research Council*. 129. [Online] Available From: http://www.metoffice.gov.uk/media/pdf/n/i/Recent_Storms_Briefing_Final_07023.pdf
- [11] Palen, L. et al. 2007. Crisis Informatics: Studying Crisis in a Networked World, *Social Science Computer Review*. 27. 467-480.
- [12] Palen, L., Vieweg, S., and Anderson, K. 2011. Supporting “Everyday Analysts” in Safety- and Time-Critical Situations. *The Information Society*. 27,1. 52-62.
- [13] Starbird, K., and Palen, L. 2010. Pass it On? Retweeting in Mass Emergency. *Proceedings of the 7th International ISCRAM Conference*. p.1-10
- [14] Terpstra, T. et al. 2012. Towards a Realtime Twitter analysis during crises for operational crisis management. *Proceedings of the 9th International ISCRAM Conference (April. 2012)*, 1-9.

- [15] Thomas, D. 2003. A General Inductive Approach for Qualitative Data Analysis. *American Journal of Evaluation*. (June, 2006), 27.2, 237-246.
- [16] Verma, S., et al. 2011. Natural Language Processing to the Rescue? Extracting “Situational Awareness” Tweets During Mass Emergency. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. 385-392.
- [17] Vieweg, S. 2010. Microblogged Contributions to the Emergency Arena: Discovery, Interpretation and Implications. *Computer Supported Collaborative Work*. 515-516.
- [18] Vieweg, S. et al. 2010. Microblogging During Two Natural Hazards Events: What Twitter May Contribute to Situational Awareness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1079-1088.