PhaseVis[1]: What, When, Where, and Who in Visualizing the Four Phases of Emergency Management Through the Lens of Social Media

Seungwon Yang[1], Haeyong Chung[1], Xiao Lin[1], Sunshin Lee[1], Liangzhe Chen[1], Andrew Wood[1], Andrea L. Kavanaugh[1], Steven D. Sheetz[2], Donald J. Shoemaker[3], and Edward A. Fox[1]

[1]Department of Computer Science, [2]Department of Accounting and Information Systems, [3]Department of Sociology, Virginia Tech, Blacksburg, VA 24061 USA

+1-540-231-6931

{seungwon, chungh, linxiao, sslee777, liangzhe, andywood, kavan, sheetz, shoemake, fox}@vt.edu

ABSTRACT

The Four Phase Model of Emergency Management has been widely used in developing emergency/disaster response plans. However, the model has received criticism contrasting the clear phase distinctions in the model with the complex and overlapping nature of phases indicated by empirical evidence. To investigate how phases actually occur, we designed PhaseVis based on visualization principles, and applied it to Hurricane Isaac tweet data. We trained three classification algorithms using the four phases as categories. The 10-fold cross-validation showed that Multi-class SVM performed the best in Precision (0.8) and Naïve Bayes Multinomial performed the best in F-1 score (0.782). The tweet volume in each category was visualized as a ThemeRiver™, which shows the 'What' aspect. Other aspects - 'When', 'Where', and 'Who' - are also integrated. The classification evaluation and a sample use case indicate that PhaseVis has potential utility in disasters, aiding those investigating a large disaster tweet dataset.

Keywords

Four Phases of Emergency Management, disaster, Twitter, ThemeRiver™

THE FOUR PHASE MODEL OF EMERGENCY MANAGEMENT

The Four Phases Model of Emergency Management (i.e., Mitigation, Preparedness, Response, and Recovery) (Baird, 2010) has been widely used in developing the response plans by emergency organizations such as Federal Emergency Management Agency (FEMA, 1996; Schwab, Topping, Eadie, Deyle, and Smith, 1998). Although some sources (FEMA, 2006; NFPA, 2007) add 'Prevention' as a fifth phase, we focus on four phases, largely since all of those phases appear in all five sources for this study. Having a clear understanding of the current disaster phase assists disaster managers in making good decisions regarding which actions to perform and how best to prepare for the tasks of the next phase. It also helps disaster researchers gain insights after the fact regarding the details of progression of disaster events and response. In spite of the model's potential benefits, critiques of the four phase model come from both academic researchers and field practitioners (Neal, 1997). They argue that the division between phases in the model is arbitrary and the emergency management activities often take place in an ad hoc fashion in practice. This concern may originate from the complex nature of disasters and the overlapping nature of their phases in real situations, which does not fit well with the distinct phase descriptions in the model (Neal, 1997). This may confuse the phase identification and thus adversely affect decision-making for actions in critical situations.

RELATED STUDIES

The use of micro-blogs such as Twitter as a medium to effectively communicate and share information before, during, and after disastrous events has become increasingly popular. For example, tweet messages provide situational awareness and report issues during disaster situations (Vieweg, Hughes, Starbird, and Palen, 2010;

[1] http://spare05.dlib.vt.edu/~ctrvis/phasevis/

Vieweg, 2012). As shown during the Japanese Earthquake and Tsunami in early 2011, micro-blogs became the communication lifeline for victims and families when cell phone communication became almost useless due to heavy cell network traffic (Gaudin, 2011). Hurricane Isaac, which struck the Caribbean and Southern USA in August and September 2012, was no exception.

Some prior research exists in the area of visual analytics tools to understand emergency situations via Twitter for the purposes of crisis management. Cameron et al. developed a web-based tool for helping to monitor and detect Twitter messages for crisis coordination and situational awareness (Cameron, Power, Robinson, and Yin, 2012). MacEachren et al. presented SensePlace2, a geographical visual analytics tool, which leverages Tweets for crisis management (MacEachren, Jaiswal, Robinson, Pezanowski, Savelyev, Mitra, Zhang, and Blanford, 2011). For the VAST 2011 Mini-Challenge 1 (Grinstein, Cook, Havig, Liggett, Nebesh, Whiting, Whitley, and Konecni, 2011), many new visual analytics tools were presented to identify an epidemic outbreak, its means of transmission, and its impact on the region in a fictitious city from about a million geo-tagged Tweets. All of these tools support ways to explore important crisis information from a large tweet collection through accessing temporal, spatial, and content-related information and overlaying the information on the map. Our work also shares inspiration with such prior visual analytics for crisis management from Twitter, but few of them employ the four phases of emergency management with tweets, nor tightly integrate them into interoperable multiple visualizations.

METHODOLOGY

Data Selection and Cleaning

We collected a total of 140, 571 tweets about Hurricane Isaac for about a month (8/23-9/26, 2012) for this study. From the resulting data set, we selected 5,675 tweets, which included major disaster organization and agency names such as FEMA, Red Cross, and Salvation Army. We considered both 'Red Cross (one space between words)' and 'RedCross (no space)'. The same applied for Salvation Army. An assumption for selecting these tweets was that they should contain either the actions conducted by those representative disaster organizations or observations about those actions by the general public or both. From this target set of 5,675 tweets, we selected 1,457 tweets (i.e., FEMA (388), Red Cross (1,052), and Salvation Army (52)), which are not retweets (RTs), to use as training data considering that RTs could decrease the ability of the classifier to accurately distinguish unique tweets.

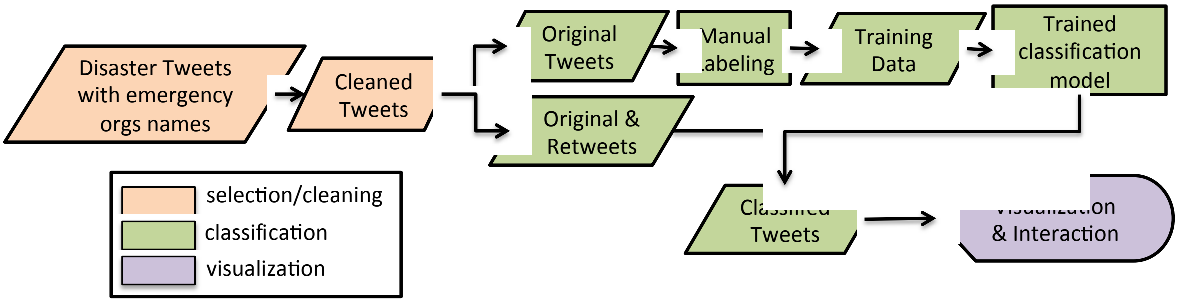

Figure 1. A data flow diagram: selection, classification, and visualization.

Figure 1 shows an overview of the procedures for the selection of relevant tweets, classification using a training set, and visualization of classified tweets. Since tweets often embed shortened links to webpages, and those webpages might provide essential information relevant to the tweet content, we concatenated the webpage title of the first link with the tweet to form a richer tweet text in order to strengthen the distinguishability of the classifier. As a pre-processing step, we removed English stopwords (e.g., and, or, the, a, for, Monday, March, hurricane, isaac, etc.), special characters (e.g., *, &, %, $, etc.), and URLs from tweets, and applied the Porter stemmer (http://tartarus.org/martin/PorterStemmer/) provided in the WEKA (www.cs.waikato.ac.nz/ml/weka).

Classification

Each tweet in our training set was manually categorized as one of '1:Response', '2:Recovery', '3:Mitigation', '4:Preparedness', and '5:Other' based on the activities described in the content. For example, a tweet "Red Cross provided 20,000+ overnight shelter stays and served 400,000+ meals/snacks to help those impacted by #Hurricane #Isaac. THANK YOU!!" is categorized as '1: Response' phase. We excluded 331 tweets categorized as '5:Other'. In total, 1,121 tweets were used as our training set.

We trained three classification algorithms -- multi-class SVM, multinomial Naïve Bayes, and Random Forest --

each with different settings for baseline, term frequency (TF), inverse document frequency (IDF), and normalization. The words in concatenated tweet text and title of embedded resource were the classification features. We used the WEKA implementation for multinomial Naïve Bayes and Random Forest. For Multi-class SVM, we used Joachims Multiclass SVM (svmlight.joachims.org/svm_multiclass.html), which is a speed-optimized implementation (Tsochantaridis, Hofmann, Joachims, and Altun, 2004) of a multi-class SVM formulation described in (Crammer and Singer, 2002). These three algorithms are known for their good performance in classifying textual data. We labeled our target set with the trained Multi-class SVM, which performed the best in precision, and used the results for our visualization.

Visualization

The design goal of our visualization is to help gain insight into vast numbers of tweets efficiently to further the assessment and understanding of emergency situations. However, the main challenges of directly employing micro-blogging data for visualization are twofold: The large volume of tweets and the complex combination of different types of information structures (e.g., spatial, temporal, relational). These challenges may increase the complexity and uncertainty involved in managing this large dataset. To address such challenges, we developed a multi-view tweet visualization interface, *PhaseVis*.

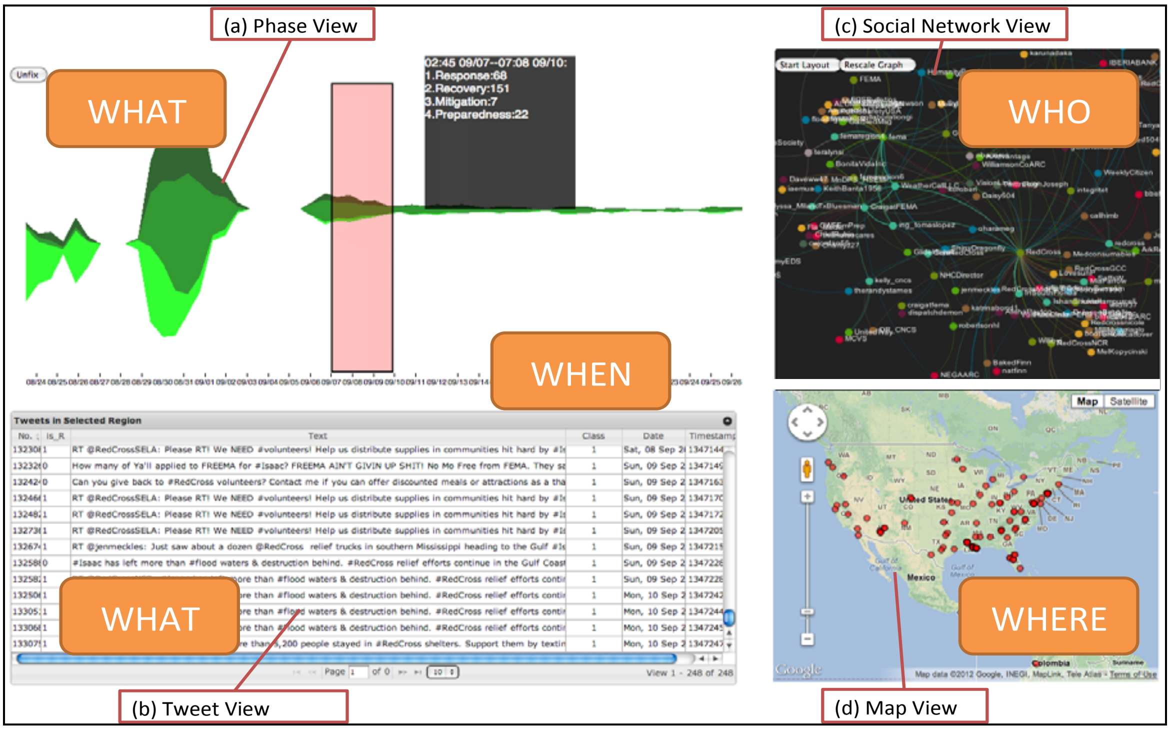

Figure 2. The PhaseVis integrating (a) Phase View, (b) Tweet View, (c) Social Network View, and (d) Map View.

It integrates autonomous data analysis approaches using classification with visual methods. Different information structures in a tweet are separated into multiple views (Baldonado, Woodruff, and Kuchinsky, 2000), being mapped into more optimal visual representations:

- The Phase View (Figure 2(a)) presents the four phases classified by the multi-class SVM through ThemeRiver™ visualization (Havre, Hetzler, Whitney, and Nowell, 2002). The tweet volume in each phase is represented by a different stream using a different color. This view also provides the user interactivity to control/query the other views (Buja, McDonald, Michalak, Stuetzle, 1991). A user can select a time range using a selection bar along the timeline by a mouse.
- The Tweet View (Figure 2(b)) focuses primarily on tweet content exploration. It displays relevant tweets, which are matched by the query or selected by different views in a scrolling list. For example, selecting a time range on the Phase View results in listing tweets in that range on the Tweet View. Also, clicking a user name in the Social Network View will list her tweets within the already selected time range.
- The Social Network View (Figure 2(c)) represents relations between users and their mentions in a

network graph during the time period selected in the Phase View.

- The Map View (Figure 2 (d)) shows the geospatial distribution of users and mentions, whose tweets are from within the time range in the Phase View, using small red circles. User IDs from the selected tweets are identified, and locations are extracted from their publicly available Twitter profile page in advance.

These interconnected views represent the 'What' (phase, tweet content), 'When' (time period), 'Where' (user location), and 'Who' (user network) respectively of the tweet dataset to help understanding the event details (Figure 2). For instance, if a specific user node is selected on one view, the related information of this user is also highlighted on the other views facilitating finding spatio-temporal answers.

EVALUATION AND A USE CASE

Cross-Validation for Trained Classifier

10-fold stratified cross-validation results for the Naïve Bayes Multinomial, Multi-class SVM, and Random Forest algorithms are shown in Table 1. For all cases, term-frequency weighting, normalization of the text length, and stemming were applied. The Multi-class SVM has the highest precision score among the three, but its F-1 measure is slightly lower than that of Naïve Bayes Multinomial due to its lower recall. Applying the inverse document frequency (IDF) decreases the performances of all three algorithms. This probably results from the use of very short documents (i.e., cleaned tweet text + cleaned title of embedded link) to compute IDF. When the training set is converted into a term-document matrix, most of the terms rarely appear in multiple documents, resulting generally in very low document frequency (DF) values of terms. When most terms gain high IDF values (i.e., inverse of low DF), the distinguishing power of the classifier may be diminished, resulting in lower performance. Considering that only about a half of our training examples were concatenated with the title of an embedded link (if it exists), normalization was a necessary step to reduce this effect and to increase performance. The use of the Porter stemmer contributed to a slightly increased performance as well.

Table 1. Cross-validation for three classification algorithms used. The highest numbers are in bold case.

Algorithm	Precision	Recall	F-1	Correctly Classified
Multi-class SVM	0.8	0.745	0.77	80.82%
Naïve Bayes Multinomial	0.79	0.779	0.782	77.88%
Random Forest	0.762	0.763	0.754	76.27%

A Sample Use Case

A researcher in the Disaster Informatics field wants to know what was happening during certain time periods of Hurricane Isaac from the associated tweet data collection. So, she first selects a time interval from the Phase View (Figure 2(a)). For the selected interval, a gray box immediately displays details of date and time range in textual format, for example, "05:25 08/24—06:02 08:26." The box also shows that Preparedness was the dominant phase followed by Response and Recovery. She then looks at the Map View (Figure 2 (d)), which shows that users are tweeting from the Caribbean, Cuba, and Dominican Republic to southeastern states of the US. She gains an insight that the high volume of Preparedness tweets might have been posted by people in the southeastern US states, who were preparing for the coming hurricane, while Response and Recovery tweets mostly came from those countries in the Caribbean, which were probably responding to and starting recovery actions then. In the Social Network View (Figure 2(c)), she examines the connections among Twitter IDs. The three IDs with the largest number of connections were 'FEMA', 'RedCross', and interestingly 'HumaneSociety'. She reads some tweets posted by HumaneSociety from the Tweet View in Figure 2 (b). Most of them were about the RedCross' opening of shelters, some of which are pet-friendly.

CONCLUSION AND FUTURE WORKS

Starting from a question, how the four phases of emergency management appear in real data, we collected tweets about Hurricane Isaac, classified the data using a machine-learning approach, and designed a multi-view integrated visualization tool, PhaseVis, based on visualization principles. Analysis results are visualized as a ThemeRiver™ as well as related views to show different aspects of the dataset. This paper contributes to knowledge regarding designing an integrated visualization tool for supporting researchers in Disaster Informatics, helping them gain insights from a large tweet data.

We acknowledge some of the limitations in our study. Using only three emergency organization names might be too restrictive considering that many other local, state, and federal level organizations and even foreign aid workers participate in various phases of disasters. As our future works, we plan to identify tweets that include many more organization names in our dataset, and to apply classification and visualization. Specifically, we aim to analyze a tweet archive of the Hurricane Sandy disaster. For this, we collected more than 2 million tweets. Improvements for PhaseVis interface also are planned. Sentiment analysis will be part of the Phase View that shows streams of emergency phases. The other improvement is to have server-side programming to deal with big data, which is not unusual for massive scale disasters. It is our hope that PhaseVis aids in gaining insights from a large tweet data as well as in refining the current disaster phase model in the long run by illustrating phases as they actually occur.

ACKNOWLEDGMENTS

We thank NSF for supporting the CTRnet project (IIS-0916733), of which this work is part. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

1. Baird, M. (2010) The "Phases" of Emergency Management - Background Paper. Retrieved Nov 3, 2012, from http://www.memphis.edu/cait/pdfs/Phases_of_Emergency_Mngt_FINAL.pdf.
2. Baldonado, M., Woodruff, A. and Kuchinsky, A. (2000) Guidelines for Using Multiple Views in Information Visualization, *Proceedings of Advanced Visual Interfaces*, Palermo, Italy.
3. Buja, A., McDonald, J., Michalak, J., Stuetzle, W. (1991) Interactive Data Visualization using Focusing and Linking, *Proceedings of IEEE Visualization 1991*, Los Alamitos, CA.
4. Cameron, M., Power, R., Robinson, B. and Yin, J. (2012) Emergency situation awareness from twitter for crisis management, *Proceedings of the 21st international conference companion on World Wide Web,* Lyon, France.
5. Crammer, K. and Singer, Y. (2002) On the algorithmic implementation of multiclass kernel-based vector machines, *The Journal of Machine Learning Research*, 2, 265-292.
6. Federal Emergency Management Agency (FEMA). (1996) Guide for All-Hazard Emergency Operations Planning. Retrieved Nov 3, 2012, from http://www.fema.gov/pdf/plan/slg101.pdf.
7. Federal Emergency Management Agency (FEMA). (2006) Principles of Emergency Management, Independent Study, IS230, Washington.
8. Gaudin, S. (2011) Twitter, Facebook become lifelines after Japan quake. Retrieved in Nov. 20, 2012, from http://www.computerworld.com/s/article/9214213/Twitter_Facebook_become_lifelines_after_Japan_quake.
9. Grinstein, G., Cook, K., Havig, P., Liggett, K., Nebesh, B., Whiting, M., Whitley, K. and Konecni, S. (2011) VAST 2011 Challenge: Cyber Security and Epidemic, *Proc. of IEEE VAST*, Providence, RI.
10. Havre, S., Hetzler, E., Whitney, P. and Nowell, L. (2002) ThemeRiver: Visualizing Thematic Changes in Large Document Collections, *IEEE Transaction on Visualization and Computer Graphics,* 8, 1, 9-20.
11. MacEachren, A., Jaiswal, A., Robinson, A., Pezanowski, S., Savelyev, A., Mitra, P., Zhang, X. and Blanford, J. (2011) SensePlace2: GeoTwitter analytics support for situational awareness, *Proceedings of IEEE VAST*, Providence, RI.
12. National Fire Protection Association (NFPA). (2007) Standard on Disaster/Emergency Management and Business Continuity Programs (2007 Edition), Quincy, Massachusetts: National Fire Protection Association.
13. Neal, D. M. (1997) Reconsidering the Phases of Disaster. *International Journal of Mass Emergencies and Disasters,*15, 2, 239-264.
14. Schwab, J., Topping, K., Eadie, C., Deyle, R. and Smith, R. (1998) Planning for Post-Disaster Recovery and Reconstruction, PAS Report 483/484 (FEMA 421), American Planning Association.

15. Tsochantaridis, I., Hofmann, T., Joachims, T. and Altun, Y. (2004) Support vector machine learning for interdependent and structured output spaces, *Proceedings of the twenty-first international conference on Machine learning*, Banff, Alberta, Canada.
16. Vieweg, S., Hughes, A., Starbird, K. and Palen, L. (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. *Proceedings of the CHI 2010,* Atlanta, GA.
17. Vieweg, S. (2012) Situational Awareness in Mass Emergency: A Behavioral and Linguistic Analysis of Microblogged Communications, Retrieved in Nov. 20, 2012, from http://works.bepress.com/vieweg/15.