Twitter Use During an Emergency Event: the Case of UT Austin Shooting

Lin Tzy Li[1,4,5], Seungwon Yang[1], Andrea Kavanaugh[1], Edward A. Fox[1],
Steven Sheetz[3], Donald Shoemaker[3], Travis Whalen, and Venkat Srinivasan[1]

[1]Department of Computer Science, Virginia Tech, Blacksburg, VA 24061
[2]Department of Accounting and Information Systems, Virginia Tech, VA 24061
[3]Department of Sociology, Virginia Tech, Blacksburg, VA 24061
[4]Institute of Computing, University of Campinas, SP, Brazil, 13083-852
[5]Telecommunications Res. and Dev. Center, CPqD Foundation, Campinas, SP, Brazil, 13086-902.

lintzyli@ic.unicamp.br, {seungwon, kavan, fox}@vt.edu,
{sheetz, shoemake, tfw115, svenkat}@vt.edu

ABSTRACT

This poster presents one of our efforts developed in the context of Crisis, Tragedy, and Recovery Network (CTRnet) project. One of our derived works from this project is the use of social media by government to respond to emergency events in towns and counties. Monitoring social media information for unusual behavior can help identify these events once we can characterize their patterns. As an example, we analyzed the campus shooting occurred in the University of Texas, Austin, on September 28, 2010. In order to study the pattern of communication and the information communicated using social media on that day, we collected publicly available data from Twitter. Collected tweets were analyzed and visualized using Natural Language Toolkit, word clouds, and graphs. They showed how news and posts related to this event swamped the discussions of other issues.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous;
J.4 [**Social and Behavioral Sciences**]: Sociology

General Terms

Management, Human Factors, Experimentation

Keywords

campus shootings, crisis informatics, microblogging, Natural Language Toolkit, social media, word clouds, Twitter

1. INTRODUCTION

This work is connected with an ongoing NSF project on "CTRnet: Integrated Digital Library Support for Crisis, Tragedy, and Recovery" [1] at Virginia Tech. Building upon prior studies of the tragic shootings on April 16, 2007, we have collected, archived, and analyzed information and communications associated with CTR-related events [1].

---

[1]http://www.ctrnet.net/

As part of the CTRnet services, we can develop an applications that could alert government emergency response teams when crisis/disaster events are detected from tweet streams. The use of a microblog like Twitter is widely studied, for example, in the area of situational awareness [2].

As an initial work to understand tweeting patterns in crisis situation, we analyzed tweets from shooting incident in University of Texas (UT) at Austin on September 28, 2010.

2. THE STUDY CASE AND DATA SET DEVELOPMENT

In order to study how the community of UT reacted to this event and communicated it in Twitter, we collected Twitter posts publicly available from users who follow UTAustin[2], which is an official Twitter screen name of the University of Texas, Austin. From its around 5,245 followers, we were able to collect public posts from 2,857 followers between September 18 and October 16, though some of these followers did not post on a daily basis.

There were three phases in our procedure to prepare and analyze the dataset. In Phase I, we collected information of the followers of UTAustin. In the next phase, we crawled tweets that had a time stamp of Sept. 19, 2010 or later by using the follower information from Phase I. In Phase III, we analyzed the dataset using MySQL queries and the Natural Language Toolkit (NLTK). Specifically, NLTK's feature to find frequently collocated word pairs was helpful. The results were then visualized using multiple word clouds and a bar graph.

3. ANALYSIS

From Figure 1, which shows the number of posts per day, we noted that for the day of the event (Sept. 28), there was a peak of over 15,000 posts, while for the other days the maximum number of posts was around 6,000 to 10,000.

Analyzing the most common words in Twitter posts of Sept. 28, we found that words such as "shooter", "gunman", "shooting", "utshooting", "suspect", and "university" – besides "campus", "UT", "Austin", and "RT" (which stands for retweets) – were the most frequent words. We presented

---

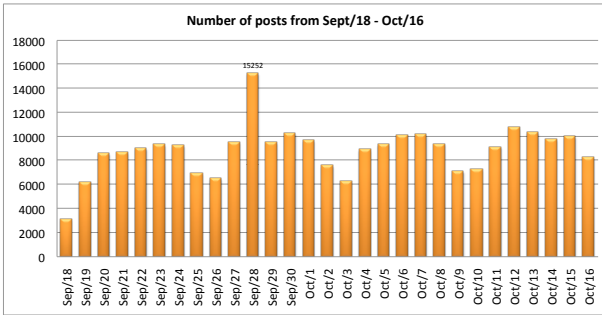[2]http://twitter.com/utaustin

Figure 1: Number of Twitter posts from followers of UT Austin from Sept. 18 to Oct. 16, 2010.

this result in a word cloud (Figure 2), since it will give the user a quick snapshot of frequent words. For more detailed word counts histograms could be displayed along with word clouds.

Figure 2: Word cloud of Sept. 28th Twitter posts from followers of UTAustin.

Comparing the distribution of number of posts over time for Sept. 28 to the day before and after it (Figure 3), the Twitter posts nearly doubled to 900 at 8 AM as compared to other days when by this time it would be below 500 posts. The peak of Sept. 28 was at 9 AM with 2,623 Twitter posts, when other days it would be around 600 by that time, an increase of over 400%.

Users mostly tweet about the shooting event for 7 hours after it happened. At 7 AM there was nothing related to it, but by 8 AM (Figure 4), around when the event happened, the words "UT", "campus", "gunman" and "shooter" were among the most used on tweet posts. The UT shooting dominated the Twitter posting of this community until 3
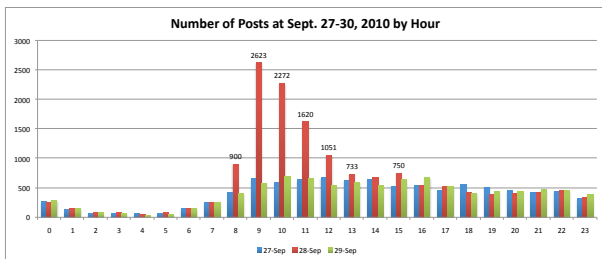

Figure 3: Number of Twitter posts over time from followers of UT Austin from Sept. 27 to Sept. 29, 2010.

Figure 4: Word cloud for Sept. 28, 8AM.

Figure 5: Word cloud for Sept. 28, 3 PM.

PM (Figure 5), when the most visible were shooter's name at UT Austin (Colton Tooley).

The 'collocation' feature in NLTK provides the top 20 frequently appearing word pairs in a data file. After separating tweets from the dataset by hour into different files, we ran the NLTK toolkit on them.

At 7 AM on the day of the incident, no word pairs were marked as related as we can expect. But it began to change radically from 8 AM. People were tweeting about the incidents frequently. Example pairs include 'active shooter', 'shot himself', 'armed suspect', 'Castaneda Library' (The library where the suspect finally went and committed suicide), and 'emergency text'.

From this study case we observed that during crises people used Twitter to share, comment on information about the event. We find that a spike in the number of tweets and changes in the ideas in the tweets signal that an event is occurring or occurred. Our content analysis method in this study was based on the word frequencies; however, other methods such as semantic analysis [3] can further help distinguish events of interest to governments emergency teams, for example. Future work will include other crisis related events and comparison of result, adding also retweet and location analysis.

4. ACKNOWLEDGMENTS

5. REFERENCES

[1] E. A. Fox, C. Andrews, W. Fan, J. Jiao, A. Kassahun, S. Lu, Y. Ma, C. North, N. Ramakrishnan, A. Scarpa, B. H. Friedman, S. D. Sheetz, D. Shoemaker, V. Srinivasan, S. Yang, and L. Boutwell. A Digital Library for Recovery, Research, and Learning from April 16, 2007, at Virginia Tech. *Traumatology*, 14(1):64–84, Mar. 2008.

[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 *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, CHI '10, pages 1079–1088, 2010. ACM.

[3] S. Yang, A. Kavanaugh, N. P. Kozievitch, L. T. Li, V. Srinivasan, S. D. Sheetz, T. Whalen, D. Shoemaker, R. da S. Torres, and E. A. Fox. CTRnet DL for Disaster Information Services. In *Proceedings of the 11th ACM/IEEE annual joint conference on Digital Libraries*, JCDL '11, 2011. ACM.