Identifying Valuable Information from Twitter During Natural Disasters

Brandon Truong, Cornelia Caragea
Computer Science and Engineering
University of North Texas
BrandonTruong@my.unt.edu, ccaragea@unt.edu

Anna Squicciarini, Andrea H. Tapia
Information Sciences and Technology
Pennsylvania State University
asquicciarini@ist.psu.edu, atapia@ist.psu.edu

ABSTRACT
Social media is a vital source of information during any major event, especially natural disasters. However, with the exponential increase in volume of social media data, so comes the increase in conversational data that does not provide valuable information, especially in the context of disaster events, thus, diminishing peoples' ability to find the information that they need in order to organize relief efforts, find help, and potentially save lives. This project focuses on the development of a Bayesian approach to the classification of tweets (posts on Twitter) during Hurricane Sandy in order to distinguish "informational" from "conversational" tweets. We designed an effective set of features and used them as input to Naïve Bayes classifiers. In comparison to a "bag of words" approach, the new feature set provides similar results in the classification of tweets. However, the designed feature set contains only 9 features compared with more than 3000 features for "bag of words." When the feature set is combined with "bag of words", accuracy achieves 85.2914%. If integrated into disaster-related systems, our approach can serve as a boon to any person or organization seeking to extract useful information in the midst of a natural disaster.

Keywords
Informational vs conversational tweets, tweet classification.

INTRODUCTION
Social media is an invaluable source of almost any information. Social media opens up access to an "effective and irrepressible real-time mechanism to broadcast information" (Stefanidis, Crooks, Radizikowski, 2014). Although these data may be very useful, the majority of social media data is conversational, therefore holding no actual weight for somebody actually searching for information. Data from social media are "vast, noisy, distributed, unstructured and dynamic" (Gundecha, Liu 2012). Inherently, a huge research focus is currently on how to make sense of the social media data that are pouring into databases and how to extract important information. This project examines the identification of informative tweets from social media data, particularly during natural disasters when, being informed, is essential to people's safety.

77th ASIS&T Annual Meeting, October 31- November 4, 2014, Seattle, WA, USA.

Twitter, a microblogging site, serves as an immediate form of broadcasting information to the world; it is a place where "people digitally converge during disasters" (Starbird, Muzny, & Palen, 2012). Informative data from sites such as Twitter can either be obtained directly from bystanders of a disaster or "derivative - that is, information in the form of reposts or pointers to information available elsewhere" (Starbird, Muzny, & Palen, 2012). Both misinformation and conversation can cloud the entirety of data coming in through social media during disasters.

Analysis of informative data from tweets will allow further understanding of trends over the course of a disaster. Informative data sources and their "derivative" branches can be mapped out to "identify the nodes (e.g., users) with the most outbound (e.g., sent @replies) or inbound connections (e.g., received retweets)," which may lead to identifying numerous patterns including "betweenness or centrality measures" and "clusters and divisions in the network" (Bruns, Liang, 2012).

Were there a system to filter informational from conversational tweets, relief efforts would have an enormous advantage in deciding what to do and where to focus relief efforts. In combination with geo-location, sentiment analysis, and other social media data mining research, informative social media filtering can lead to more correct decisions leading to fewer casualties or harmed bystanders. In this paper, our goal is to design novel features that can be used as input to machine learning classifiers in order to automatically and accurately identify informational tweets from the rest in a timely fashion.

RELATED WORK
Researchers have demonstrated the power of microblogging on the diffusion of news-related information (Kwak et al., 2010; Java et al., 2007; Oricchio, 2010; Lerman and Ghosh, 2010). Microblogging has been under the lens of researchers with regards to its use in disasters and other high profile events (Hui et al., 2012). However, in times of crises, microblogging can create a lot of noise in which

stakeholders need to sift through to find relevant information. With the use of the Internet, news about a crisis can spread quickly and without any boundaries (Bucher, 2002).

If microblogged data were guaranteed to lead to optimal decision-making, the use of unverified data from unvetted persons would not feel like the insurmountable challenge that has been described in (Tapia et al. 2011). The problem is a classic rhetorical debate of optimizing v. satisficing in decision theory sciences (Odhnoff, 1965; Byron, 2004). One belief is that by optimizing all available data, including data gleaned from social media, disaster responders will be able to make the best possible decisions. However, as Palen et al (2010) contend, crisis responders never have perfect knowledge of any given crisis, as crises, by definition, are scenarios where conditions hinge on extreme instability. In order to have all available data, an event must exist in a bounded universe, however, disasters and other crises are complex by nature, with many moving parts, not at all capable of being finite in disaster situations which evolve, hour by hour. Thus, satisficing, or the "good enough" principle in decision making should also apply to the technology employed (Granger-Happ, 2008). No automated tool is going to account for all the unknown variables of a situation, but the tools employed can maximize the assessment of the variables known and can help establish a certain level of trust.

As research in humanitarian disaster response teams has shown, even with the incorporation of all possible data, optimal decision making is difficult to achieve (Muhren and Walle 2010). Overcoming the barrier of processing information may be the biggest factor in establishing verifiable information. Machine learning, data mining, and natural language processing have made great leaps in extracting, processing and classifying microblogged feeds. For example, Qazvinian et al. (2011) and Yang et al. (2012) have focused on identifying misinformation in microblogs, whereas Castillo et al. (2011) analyzed information credibility in microblogs (i.e., information "offering reasonable grounds for being believed"). However, these works are not done in the context of disaster-related events. Sakaki et al. (2010) have used machine learning techniques to detect earthquakes in Japan, and Mendoza et al. (2010) studied the propagation of rumors and misinformation from the Chilean earthquake using only a small set of cases. Most research in this area has been performed post-hoc, and the most important aspect of any intelligence received, intelligence that is actionable and precisely geo-located, has not yet been achieved and is also complicated by translation and nuance of understanding language (McClendon and Robinson 2012; Munro 2011). To our knowledge, the automated detection of disaster-related, informational data within microblogging platforms has not been researched.

APROACH

The project's purpose was to develop a set of features for use in machine learning algorithms that would be able to accurately distinguish "informational" tweets from "conversational" ones. We defined "informational" tweets as any tweets, which would provide valuable, concrete information to anybody viewing the tweet. "Conversational" tweets therefore were defined as having no concrete information; the information would not be universally useful (within language barriers) to anybody who could read the tweet. Examples of informational and conversational tweets are shown in Table 1. We treated the problem as a binary classification problem, considering the fact that a tweet should be either informative or not. An irrelevant tweet would be classified as not informative since it would not provide information immediately usable to a person. We manually labeled 1086 tweets for the classification task with help from students from our research labs. This set of tweets contains 139 informational and 943 conversational tweets.

Tweet	Classification
This hurricane sandy twitter is so annoying	Conversational
RT @cnnbrk: More than 765000 in 7 states have no electricity with NY and NJ being most affected. #HurricaneSandy http://t.co/XEYNBgW0	Informational

Table 1. Examples of informative and conversational tweets

Next, we discuss the features used in the classification task.

Feature Development

A primary distinguishing factor between informational and conversational tweets is discrepancies that reveal formality. Aspects of formality include correct grammar, lack of slang, lack of swear words, etc. A formal tweet is likely informative, since many credible sources of information will most likely structure their tweets to look as professional as possible. Therefore, the most effective features in classification revolve around the idea of formality. A discussion of the proposed features and the intuition behind them is presented in what follows. The features are listed in descending order based on Information Gain Ranking from the Weka data mining toolkit[1].

URL extraction

URLs are used extensively in tweets to link to external sources that could not ordinarily fit in the length-restricted structure of a tweet. However, URLs found in tweets are shortened in order to accommodate for the length-restriction. Inherently, a few features can be extracted from the URL itself, considering that each shortened URL has a base domain and a randomly generated code appended to that (i.e. "t.co/[code]"), which, when clicked on, will redirect to an actual web page. We developed a method to

[1] http://www.cs.waikato.ac.nz/ml/weka/

bypass this system through having our feature extraction methods request through the shortened URL to the legitimate webpage, returning an analyzable URL. From this URL, we developed a method of analysis to distinguish between informational and conversational tweets. The URLs were analyzed for hyphenated or underscore-separated article names in the path, indications of dates, and the presence of reputable news sources that were extracted from the dataset. Conversational indicators, such as social media pictures with no informative text, were factored in.

Emoticons

In addition, emoticons (e.g., ":)") are highly representative of a conversational tweet. Reputable sources would most likely not include emoticons since they have almost no informational weight. Credible sources of information likely to provide informative tweets will avoid indications of conversation and attempt to retain seriousness in the tweet, which emoticons detract from.

Instructional keywords

Through examination of tweets, certain keywords implying formality were developed. A large number of informative tweets often give instructions to people who see the tweet. For example, reputable news sources or relief efforts often include instructions to "Text" or "Call" a phone number or "Donate" at a URL. These words and related words common in an instructional tweet were grouped together and searched for in the tweets as a feature. The presence of a keyword from this set indicates informative-ness.

Phone numbers

Along with the instructional keywords comes the presence of a phone number, which is common in instructional tweets. In context, generally, only reputable sources would expose a phone number to public social media.

Internet slang

Internet abbreviations and slang are representative of informalities, which are representative of conversations. We developed a dictionary of common slang and variations. A credible source is not as likely to have these abbreviations in their tweets, although they may be included occasionally for brevity.

"RT", and Profanity

In addition, the presence of "RT" (retweet) was found to aid in accuracy. Profanity is also a common indicator of informality. Presence of a curse word would therefore indicate a conversational tweet, as a source attempting to broadcast information would most likely not include profanity in their tweets.

Sentence structure analysis

We used OpenNLP Java Libraries in order to check various grammar rules. Again, informative tweets will most likely contain more formal grammar. Checking for punctuation and complete sentences is therefore a somewhat informative feature. Parsing for sentences through the OpenNLP parser, checking for abrupt sentences and whether or not there are multiple sentences also provide important hints for classification.

In summary, our nine proposed features are: "*has hash tag*," "*abrupt sentence*," "*multiple sentences*," "*informative URL*," "*has phone number*," "*has emoticon*," "*has retweet*," "*has keyword*," "*has curse word*."

RESULTS AND EVALUATION

We used 10-fold cross-validation for evaluation, using Naïve Bayes classifiers as implemented in the Weka toolkit. The results of the designed feature sets were compared with the outcomes of the "bag of word", along with the results of a combined result set of "bag of words" and the designed feature set. We report precision, recall, and F-measure for each class. These measures are widely used in information retrieval tasks. In addition, we also used the area under the Receiver Operating Characteristic curve (AUC).

Results from all found methodologies are depicted in Table 2. As can be seen from the table, the combined "bag of words" and selected feature set resulted in the most accurate prediction model. The designed feature set performs similarly and in some cases better than the "bag of words" approach. However, the number of features used in the BOW approach (1488) is significantly larger compared to our 9 features. The worse performance on the informational class as compared with that of the conversational class could be due to unbalanced data (more conversational than informational).

We achieved models of over 85% accuracy using the combined strategy. Also, the combined strategy achieves an F-measure of 0.514 for the informational tweets as compared with 0.42 and 0.478 for the designed features and the BOW, respectively.

	Our Features	BOW	Combined (features + BOW)
Precision - conversational	0.916	0.928	0.939
Precision - informational	0.405	0.442	0.444
Recall - conversational	0.907	0.903	0.889
Recall - informational	0.435	0.522	0.609
F-Measure - conversational	0.912	0.916	0.913
F-Measure - informational	0.42	0.478	0.514
AUC	0.812	0.86	0.865

Table 2. Shows the resulting data from classification techniques.

CONCLUSION

We designed novel features for use in the classification of tweets during natural disasters in order to develop a system through which informational data may be filtered from the conversations, which are not of much value in the context of searching for immediate information for relief efforts or bystanders to utilize in order to minimize damages.

The results of our experiments show that classifying tweets as “informational” vs. “conversational” can use solely the proposed features if computing resources are concerned, since the computing power required to process data into featured is immensely decreased in comparison to a BOW feature set which contains a substantially larger number of features. However, if computing power and time necessary to process incoming Twitter data are not a concern, a combined feature set of the nine proposed features and BOW-presence approach will maximize overall accuracy.

Through the proposed feature set, data can accurately be filtered and utilized by those in need of information along with those intending to organize relief efforts. By having a database of informative tweets, one can further understand, in real time, the effect that a disaster is having on the affected population. Although the primary aim of the project was focused on the classification of tweets during a natural disaster, the selected features are general enough and could inherently be applied to other fields of research and situations with only little or no modifications.

ACKNOWLEDGMENTS

We are grateful to Kishore Neppalli, Daniel Stieben, Terrence Liu and Sam Stehle for helpful discussions and their help with data annotation. This research was supported in part by NSF awards #1353418 and #1353400. Any opinions, findings, and conclusions expressed here are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Alexander, M. (2000). Confronting Catastrophe: New Perspectives on Natural Disasters.

Bruns, A., & Liang, Y. E. (2012). Tools and methods for capturing Twitter data during natural disasters. First Monday, 17(4).

Bucher, H.-J. (2002). Crisis Communication and the Internet: Risk and Trust in a Global Media. First Monday. http://www.firstmonday.org/issues/issue7_4/bucher/

Byron, M. Satisficing & Maximizing Moral Theorists on Practical Reason. Cambridge, United Kingdom: Cambridge University Press., 2004.

Castillo, C.; Mendoza, M.; & Poblete, B. Information credibility on twitter. In Proc. Of WWW, pages 675–684, 2011.

Gao, H.; Barbier, G.; & Goolsby R. "Harnessing the crowdsourcing power of social media for disaster relief." IEEE Intelligent Systems 26.3 (2011): 10-14.

Granger-Happ, E. The good enough principle - nonprofits. Retrieved from www.fairfieldreview.org/.../The Good Enough Principle..., pages 1–25, 2008.

Gundecha, P; & Liu, H. "Mining social media: A brief introduction."Tutorials in Operations Research 1.4 (2012).

Hui, C.; Tyshchuk, Y.; Wallace, W.; Goldberg, M.; & Magdon-Ismail, M. (2012). Information cascades in social media in response to a crisis: a preliminary model and a case study. WWW ’12 Companion Proceedings of the 21st international conference companion on World Wide Web, pp. 653–656.

Java, A.; Song, X.; Finin, T.; & Tseng, B. (2007). Why we twitter: understanding microblogging usage and communities. Network, 43(1), 56–65.

Kwak, H.; Lee, C.; Park, H.; & Moon, S. (2010). What is Twitter, a social network or a news media? Proceedings of the 19th international conference on World Wide Web Vol. 112, pp. 591–600.

Lerman, K., & Ghosh, R. (2010). Information Contagion: an Empirical Study of the Spread of News on Digg and Twitter Social Networks. Fourth International AAAI Conference on Weblogs and Social Media, 90–97.

Odhnoff, J. On the techniques of optimizing and satisficing. The Swedish Journal of Economics, 67(1):24–39, 1965.

Oricchio, R. (2010). Is Twitter A Social Network? Inc. Magazine. http://www.inc.com/tech-blog/is-twitter-a-social-network.html

Palen, L.; Vieweg, S; & Anderson, K. M. Supporting “everyday analysts” in safety-and time-critical situations. The Information Society, 27(1):52–62, 2011.

Mendoza, M; Poblete, B; & Castillo, C. Twitter under crisis: Can we trust what we rt? In Proceedings of the First Workshop on Social Media Analytics, SOMA ’10, pages 71–79, 2010.

Muhren, W. J & Van De Walle, B. Exploring decision-relevant information pooling by humanitarian disaster response teams. May (1-5) 2010.

Munro, R. Subword and spatiotemporal models for identifying actionable information in Haitian kreyol. In Proc. of CoNLL, pages 68–77, 2011.

McClendon, S; & Robinson, A. C. Leveraging geospatially-oriented social media communications in disaster response. In Proceedings of the International ISCRAM Conference, pages 2–11, 2012.

Qazvinian, V.; Rosengren, E.; Radev, D. R.; & Mei, Q. Rumor has it: Identifying misinformation in microblogs. In Proceedings of EMNLP, EMNLP ’11, pages 1589–1599, 2011.

Sakaki, T; Okazaki, M; & Matsuo, Y. Earthquake shakes twitter users: Real-time event detection by social sensors. In Proceedings of WWW, pages 851–860, 2010.

Starbird, K.; Muzny, G.; & Palen, L. (2012). Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground Twitterers during mass disruptions. Proc. ISCRAM.

Stefanidis, A.; Crooks, A.T.; Radzikowski, J.; Croitoru, A. & Rice, M. (2014), Social Media and the Emergence of Open-Source Geospatial Intelligence, in Murdock, D.G., Tomes, R. and Tucker, C. (eds.), Human Geography: Socio-Cultural Dynamics and Global Security, US Geospatial Intelligence Foundation, Herndon, VA, pp. 109-123.

Tapia, A. H.; Bajpai, A. H.; Jansen, B. J.; & Yen, J. Seeking the trustworthy tweet: Can microblogged data fit the information needs of disaster response and humanitarian relief organizations. In Proceedings of ISCRAM, pages 1–10, 2011.