

What Should Count?: A Quantitative Approach to Scoping Rumors in Social Media

Logan Walls, Jim Maddock, Kate Starbird, and Emma Spiro

ABSTRACT

Background:

Social media are important for rapidly obtaining and disseminating information, particularly when information is time-sensitive and critical for decision making. However, as individuals try to make sense of information in non-routine situations, rumors and misinformation often proliferate, eroding users' confidence in these platforms. (Turner, 1994; Hill, 2012; Hiltz et al., 2014; Hughes & Palen, 2012) As such, empirical studies of rumor on social media—and informal communication dynamics more generally—have become a highly relevant area of study. One challenge facing such work is determining a method of identifying relevant content (e.g. tweets, posts) within the larger social media stream. Our work directly addresses this problem, using unsupervised learning methods to augment qualitative approaches to define datasets of interest within a large corpus of text-based data.

Objective:

Social media studies typically draw from a corpus of content, however, research questions often pertain to smaller subsets of the larger dataset. Our research, for example, divides a corpus of tweets containing a particular set of keywords into subsets of content related to particular rumors. The process of identifying rumor-related content involves several components, including: (1) identifying salient information threads (e.g. stories) in a tweet corpus that may constitute a rumor, (2) refining search terms or “scoping” the potential rumor, (3) validating the subset of tweets generated, and finally (4) classifying tweets and analyzing the resulting patterns. Previous research frequently glosses over the initial three steps of this process, focusing instead on classification tasks (Oh et al., 2013; Mendoza et al., 2010; Maddock et al., 2015). Yet the choice of search terms can have a significant effect on the resulting analysis; moreover, methods which rely on outside sources (e.g. Snopes or Wikipedia) could introduce a bias towards high-volume, prominent stories. This paper focuses on the process of refining search terms and scoping rumor stories.

Methods:

While human coders can intuitively generate an initial set of search terms to define scope using an iterative and grounded approach (Maddock et al., 2015), the size of massive social media datasets creates inconsistency during this process. Coders are likely to miss small-scale variations of a story or important variations of particular terms. Therefore, we introduce an approach which augments such techniques with machine learning methods of discovering patterns within text corpora. We develop and evaluate multiple procedures for altering the initial search terms (which define rumor scope) using Mikolov et al.'s (2013) Word2Vec algorithms. These word embedding models produce candidate terms (based on use within the

corpus) which can be added to or deleted from the initial search query. Proposed changes are evaluated based on their ability to refine or expand the search results, while not introducing excessive noise (i.e. unrelated content). We then manually code the resulting set of tweets to determine if they are related (or unrelated) to the rumor story of interest, quantifying improvement based on the new query's ability to increase the percentage of relevant results.

Results:

In this project we apply the above approach to a dataset of tweets from the December 2014 Sydney Siege, a 16 hour hostage and standoff situation that occurred in in the Lindt in central Sydney last year. We have implemented three different methods of adjusting search queries to refine rumor scope. We are currently in the process of human coding new tweets identified through this process. Preliminary results indicate the potential for increased ability to find rumor-related content and reduce noise.

Future Work:

Our work presents a framework for combining human and machine learning methods for the analysis of social media data, and illustrates the usefulness of this approach in studying rumoring behavior during crisis events. Future work will extend the set of automated and semi-automated approaches to refining the rumor scope and apply these techniques across a variety of rumor and event categories. In particular, we aim to explore how the effectiveness of this method varies depending on the scale or volume of the rumor.

References:

- Maddock, J., Starbird, K., Al-hassani, H., Sandoval, D. E., Orand, M., & Mason, R. M. (2015). Characterizing Online Rumoring Behavior Using Multi-Dimensional Signatures, *CSCW 2014*, ACM, 1–14.
- Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter Under Crisis: Can we trust what we RT? *Proc. of 1ST Workshop on Social Media Analytics 2010*, ACM, 71-79
- Oh, O., Agrawal, M., & Rao, H. R. (2013). Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises. *MIS Quarterly*, 37(2), 407–426.
- Hill, K. (October 30, 2012). Hurricane Sandy, @ComfortablySmug, and the Flood of Social Media Misinformation. *Forbes.com*. Available at: <http://www.forbes.com/sites/kashmirhill/2012/10/30/hurricane-sandy-and-the-flood-of-social-media-misinformation/>
- Hiltz, S. R., Kushma, J., & Plotnick, L. (2014). Use of social media by US public sector emergency managers: Barriers and wish lists. *Proc. of ISCRAM*.
- Hughes, A. L., & Palen, L. (2012). The evolving role of the public information officer: An examination of social media in emergency management. *J. of Homeland Security & Emergency Management* 9(1), 1-20.

Turner, R. H. (1994) "Rumor as Intensified Information Seeking: Earthquake Rumors in China and the United States." *In Disasters, Collective Behavior, and Social Organization*, R. R. Dynes and K. J. Tierney, Eds. University of Delaware Press, Newark, NJ, p244–256.