# Information Reliability in Twitter

When using social media applications for situational awareness, one important consideration is the quality of information provided. In the case of Microblog 911, an application which is designed to detect and report emergency situations based on data posted to Twitter, if an emergency was falsely reported due to faulty information, it could result in wasted time and effort that could be used to respond to a real emergency. In “Toward Formal Reasoning with Epistemic Policies about Information Quality in the Twittersphere”, the authors discuss current approaches to information evaluation, as well as several novel methods of addressing the problem of information reliability from social media [1].

Throughout their paper, the authors work upon the framework of NATO Standard Agreement (STANAG) 2022, which addresses the evaluation of information used in intelligence reports. Per STANAG 2022, each individual item is assigned a confidence rating that indicates the source’s reliability as well as the credibility of the information. The confidence rating is made up of source reliability, assessed on a scale of A through F, and the credibility of the information, which is rated on a scale of 1 to 6. Unfortunately, the source reliability rating is traditionally based on how accurate previous statements by the source have been. In the realm of social media such as Twitter, tracking the reliability of a source to assess accuracy of all previous statements is impractical and in some cases nearly impossible. As such, a different method must be used for source reliability assessment.

A method for determining source reliability that is used by many existing applications is analysis of network properties, such as centrality, overlap, and distance. This means that the reliability of the source is determined by evaluating the source’s network. If the source has a dense network of reliable sources, then it also it likely to be reliable. This is the premise behind Google’s PageRank algorithm. The PageRank algorithm calculates the quality of a document by assessing the quality of inlinks to the document recursively. Similarly, a blog search engine called Technorati uses the number of links from other blogs as the quality metric. When this metric is combined with reader engagement, measured by blog comments, it outperforms the PageRank algorithm [2].

There are also several existing methods in place for determining information credibility. One method, utilized by the computational knowledge engine Wolfram Alpha, is to limit the information provided to that from highly regarded sources. Conversely, Wikipedia uses the approach of crowdsourcing, which means the accuracy of the information is determined by letting anyone revise the information until consensus is reached. Neither of these methods is applicable to Twitter, as there is no limit to who can provide information, and there is no means for others to edit incorrect information in posts by others. Although search engines such as Google makes efforts to assess source reliability as described above, they do not attempt to evaluate the consistency of information provided by different sources.

After analyzing the existing methods for assessing source reliability and information credibility, the authors proposed methods for performing these assessments in Twitter. For source reliability, the use of network analysis is particularly suited to Twitter due to its intrinsic nature as a network. Of the available network analysis algorithms, they selected the TunkRank metric, proposed by Daniel Tunkelang in [3]. The algorithm scores the source based on the expected number of people who will see a message posted by the source and the probability that a user will retweet a post. This algorithm was found to be far more useful than merely counting followers, a metric which can be inflated by creating fake users.

For assessing the credibility of a message, it is necessary to determine if the content of a message is independently confirmed by another message, as this is an important part of credibility for STANAG 2022. However, it is not desirable for the messages to be exactly the same, as this would imply that they are dependent. Thus, it is necessary to identify messages that have the same meaning but different wording. For this purpose, the authors selected the Rouge-S metric, which is used to compute similarity of document summaries. The authors then added additional logic to ensure that the similar message was not a negation of the first message. For example, Rouge-S would indicate the message “Hetherington is dead” is similar to the message “Hetherington is not dead”, when in reality they are contradictory.

The authors provided an assessment of several possible methods for determining source reliability and information credibility in Twitter. The TunkRank algorithm is a much more useful method for determining the influence of a source than merely counting followers due to solving the problem of follower inflation through fake accounts. The Rouge-S metric is one of several methods that can be used to determine the similarity of posts, a factor that feeds into credibility. The minor modifications to the method proposed by the author to deal with cases of message negation make it the superior means for this task. Overall, the results of the paper are very applicable to the Microblog 911 application and can be easily incorporated to result in a more reliable and less error-prone application.

# References

[1] B. Ulicny, M. Kokar. Toward Formal Reasoning with Epistemic Policies about Information Quality in the Twittersphere. Proceedings of the 14th International Conference on Information Fusion (FUSION), Chicago, IL, July 5-8, 2011.

[2] B. Ulicny, C. Matheus, M. Kokar. Metrics for Monitoring a Social-Political Blogosphere: A Malaysian Case Study. IEEE Internet Computing, Special Issue on Social Computing in the Blogosphere. March/April 2010.

[3] Daniel Tunkelang. A Twitter Analogy to PageRank. The Noisy Channel Blog. <http://thenoisychannel.com/2009/01/13/a-twitter-analog-to-pagerank/>.