Review of Journal Articles

1. Breaking News Detection and Tracking in Twitter

This paper discussed a means of using the microblog Twitter to detect and track breaking news, which they define as “news that has either just happened or is currently happening.” The authors follow three steps to detect a story: sampling, indexing, and grouping. The sampling stage identifies messages with words like “breaking news” or the hashtag #breakingnews. Next, the messages are indexed using Apache Lucene, an open-source text search engine. The indexed messages are then grouped with similar messages, based on comparison using TF-IDF (Term Frequency-Inverse Document Frequency). Finally, the groups were scored based on reliability and popularity. Reliability was determined from the number of followers of users in the group, and popularity was determined from the number of retweets in the group.

This paper is very relevant to our project, as detecting breaking news related to emergencies is a large part of the task which must be accomplished. The only missing element is classifying the breaking news as a type of emergency and notifying users. We will be able to follow a similar approach for detecting our emergency situations, although additional filtering will be needed so that the product only focuses on emergencies. Finally, the approach to scoring message groups based on popularity and reliability may be a good solution to our concerns regarding ranking input from users.

1. Collective Intelligence for the Design of Emergency Response

This paper discusses how collective intelligence or crowdsourcing can be used for emergency response. It provides several examples where this has been successful. One example discussed is FireMash, which is a system to report bushfires in Australia. Users can notify the system of a fire by using a special hash tag (#nswfires) in their tweets. They can also use an online interface which displays a map to identify a fire. Additionally, they can add their own location to the map to ensure they are notified of a fire in their vicinity. Another example the authors presented was asking the public to identify the location of fire hydrants so that firefighters would not need to search for them when a fire occurs. This example also allows users to act using microblog posts to identify fire hydrants in their area. Users can even take pictures of hydrants with their smartphones, as such pictures are geo-tagged.

Similar to the first paper reviewed, this paper also provides ideas that could be useful for our project. The FireMash example in particular has very similar goals to our project. We may even consider adding an interactive map to the user interface for our project to easily allow users to identify the location of an observed emergency. The main difference between this approach and our project is the users must be cognizant of the FireMash system and intentionally communicate to the system by providing the #nswfires hash tag in their messages. Our project hopes to get input from all users, even those who have no knowledge of our system. The second example provided is less applicable to our project as it focuses on gathering information in preparation for an emergency, which is currently now something we plan to do.

1. TwitterReporter: Breaking News Detection and Visualization through the Geo-Tagged Twitter Network

This paper discussed an approach to detecting real-time news in Twitter utilizing the geo-tagging feature of Twitter. The authors grouped tweets by time and space. Tweets were analyzed in one hour batches in .5 by .5 decimal degree geospatial areas. A stemming algorithm was applied to change words to their roots (for example, raining becomes rain), in order to simplify comparison of the posts. The simplified tweets were then scored using document frequency weighting for pre-determined topics (such as tornado, earthquake, etc.). Topics with a high frequency were shown as events on a Google Maps overlay.

The techniques in this paper are somewhat more simplistic than those shown in some of the other papers reviewed, resulting in some anomalous results due to retweets and other issues. However, these techniques may be combined with those from the other papers to further improve those algorithms. The use of the Google Maps API to visually display the events is something that might work well for this project, as Google Maps has been adapted for both browser applications and mobile devices. This may be a good way to notify emergency personnel, as they may have existing Google Map overlays showing the location of emergency services facilities or personnel.