Mining Multimodal Information on Social Media for Increased Situational Awareness

Stephen Kelly
Trinity College Dublin, Ireland
kellys51@scss.tcd.ie

Xiubo Zhang
Trinity College Dublin, Ireland
zhangx6@scss.tcd.ie

Khurshid Ahmad
Trinity College Dublin, Ireland
kahmad@scss.tcd.ie

ABSTRACT

Social media platforms have become a source of high volume, real-time information describing significant events in a timely fashion. In this paper we describe a system for the real-time extraction of information from text and image content in Twitter messages and combine the spatio-temporal metadata of the messages to filter the data stream for emergency events and visualize the output on an interactive map. Twitter messages for a geographic region are monitored for flooding events by analysing the text content and images posted. Events detected are compared with a ground truth to see if information in social media correlates with actual events. We propose an Intrusion Index as part of this prototype to facilitate ethical harvesting of data. A map layer is created by the prototype system that visualises the analysis and filtered Twitter messages by geolocation.

Keywords

Spatio-temporal, Social media analysis, Multimodal analysis, Geolocation.

INTRODUCTION

In this work we present a system that combines the spatio-temporal metadata of Twitter messages with an analysis of the content of the messages, which includes text and images. An evaluation is presented that motivates the system to be used in an emergency or disaster scenario, and will allow emergency managers to increase their understanding of events as reported on social media. The system aims to mine and extract latent information in social media text, analysing the content of the messages, performing image recognition, and creating a spatial data layer that can be used to visualise the results on a web map. Combining the spatial and temporal data of social media with the content of the messages in this way can reveal potentially sensitive information. These privacy concerns must be addressed in any system that performs collection, processing, analysis, and visualisation of social media data.

This paper describes a system that integrates content analysis methods for social media messages and image analysis with the spatial and temporal data inherent in each message to extract meaningful information about ongoing events. It also includes a tool that allows a user to observe data privacy. We look at disaster and emergency management as the domain of application. An output and demonstration of the system is presented where messages and conversations related to flooding events and storm landings in Ireland are detected by looking at geolocated messages in these regions. For Ireland this is during the period of December 2015 to January 2016.

The system uses the spatio-temporal metadata of the Twitter messages to track the unfolding of an event on social media as discussed by members of the public. A method and implementation of monitoring the potential *intrusiveness* of this system is also described that uses natural language processing to allow a user of the system to engage in ethical data harvesting and analysis as part of work conducted by the authors previously (Kelly &

Ahmad, 2014, 2015). Visualisation of the analysis and system output is also illustrated by creating a map layer that can be incorporated into any geographical information system or web map. We show that by utilising the spatial and temporal aspects of social media data and analysis of the text and image content of the messages, effective and timely information can be extracted to increase situational awareness and can be summarised and visualised to aid decision making for an emergency manager.

RELATED WORK

Information published on social media platforms has been known to contain valuable information that is timely, often geographically tagged, and descriptive of ongoing events, all of which is information that can increase awareness of a situation (Vieweg, Hughes, Starbird, & Palen, 2010). Situational awareness is an important concept for crisis and emergency management where informed decision making is vital for efficient operation and management of a crisis event (Yin, Lampert, Cameron, Robinson, & Power, 2012). The benefits and use of social media for organisational purposes and as a new channel of information has been demonstrated early on since the emergence of microblogging (Sutton, Palen, & Shklovski, 2008). Methods and systems have been developed to do topic and trend detection, prediction, and forecasting (Althoff, Borth, Hees, & Dengel, 2013; Jin, Gallagher, Cao, Luo, & Han, 2010). Visualisation and methods of displaying this information have also been investigated in order to increase situational awareness for users (Mathioudakis & Koudas, 2010). Detecting events has been an integral part of these systems, where burst detection in the volume of messages has been a typical approach (Dou, Wang, Skau, Ribarsky, & Zhou, 2012; Li, Lei, Khadiwala, & Chang, 2012; Mathioudakis & Koudas, 2010; Robinson, Power, & Cameron, 2013). Machine learning techniques have also been employed to filter the deluge of information published on microblogging platforms with reliable levels of accuracy (Imran, Elbassuoni, Castillo, Diaz, & Meier, 2013).

Spatio-temporal models have been investigated in a number of studies to find the location of a targeted event such as earthquakes (Sakaki, Okazaki, & Matsuo, 2010; Yin et al., 2012), flooding events and fires (Vieweg et al., 2010), conflicts and political disasters (Dou et al., 2012; Ritter, Etzioni, & Clark, 2012; Robinson et al., 2013), with some systems attempting to automatically detect events (Li et al., 2012). The importance of publication time of these data has been highlighted for use in prediction (Li et al., 2012).

Some systems, such as *TwiCal* (Ritter et al., 2012) and *Tweet4act* (Chowdhury, Imran, Asghar, Amer-Yahia, & Castillo, 2013), have attempted to analyse the content of the messages also. It has been shown that by using natural language processing techniques on social media messages that contain a particular topic useful information can be extracted for a particular domain (Aiello et al., 2013). This example is encouraging and indicates that social media contains useful and relevant information. Key difficulties in this research area include verifying information, extending to other media types, going from situational awareness to decision support, and data privacy issues (Adam, Shafiq, & Staffin, 2012; Imran, Castillo, Diaz, & Vieweg, 2015).

SYSTEM OVERVIEW AND METHODS

The following section discusses the system with conceptual outline of shown in Figure 1. The Slandail Social Media Monitor provides a framework for performing domain-specific multimodal analyses on geographically bounded social media data streams. To begin analysis a geographical region on a map is specified through the GUI and domain filters are set which specify which semantic dictionaries to use. The system connects to Twitter via Twitter's Streaming API to pull in all messages from the geolocation specified by the user of the system. The demonstration described here uses the Twitter API but can connect to the APIs of other social media platforms (Facebook, Instagram) using the same architecture.

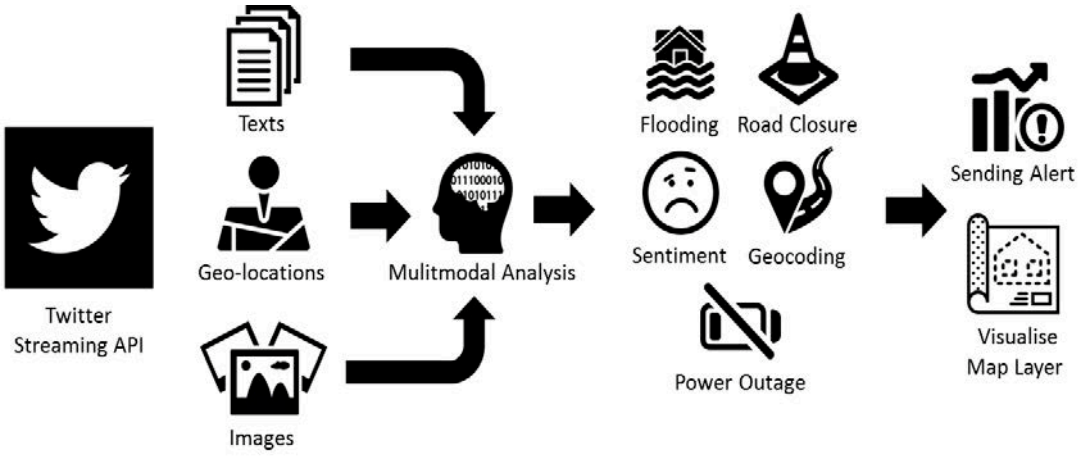

Figure 1. Domain-specific multimodal analyse of a social media stream to identify messages with images and text content related to emergency scenarios with geolocated information to produce an alert from the filtered social media stream and a map layer to visualise results

When the system is running, the message stream is pulled from the API and pre-processing is performed to extract feature, which are then passed to a classifier to see if they are relevant for emergency management. To analyse the text content of each message Stanford's CoreNLP package is used to perform tokenisation and part-of-speech tagging of the raw text. A Twitter specific part-of-speech tagging model is used in place of the general language model included in the CoreNLP package (Derczynski, Ritter, Clark, & Bontcheva, 2013). Technical details of the text analytics system, and the Slandail Social Media Monitor which has been developed in Java, have been presented in Zhang et al (Zhang, Kelly, & Ahmad, 2016). The text features were generated using the frequency occurrence of the disaster terms in the social media messages (Spyropoulou, 2016). Images contained in the messages are passed to an image recognition model that has been trained specifically to identify classes of images that are relevant for emergency scenarios and has been developed, tested, and published as part of the Slandail project (Jing, Scotney, Coleman, & McGinnity, 2016; Jing, Scotney, Coleman, McGinnity, et al., 2016). Processing messages in this way allows the text and image content to be analysed and to determine whether relevant information about an emergency or disaster event is present in the message.

The system can perform a real-time analysis of a social media stream by using an asynchronous processing pipeline. This allows collected messages to be processed in separate threads and improves error handling by ensuring no single message disrupts the analysis pipeline (Zhang et al., 2016).

Processed messages are then filtered for relevancy where only messages related to an emergency event are kept. The filtering of the text is performed by identifying domain terms defined in a set of domain dictionaries where only messages with these domain terms will be kept for subsequent analysis. A semantic dictionary, in the context of this paper, simply refers to a mapping from linguistic patterns to their semantic categories. The linguistic patterns in such dictionaries are expressed in TokensRegex (Chang & Manning, 2014) a regular expression implementation that matches token sequences rather than character sequences. Defining linguistic patterns at the token level enables the matching of complex patterns. The domain dictionaries referred to here include a categorised list of terms for emergency situations or concepts, such a flood or hurricane, which have been compiled into a term base and published as part of the output of the Slandail project (Spyropoulou, 2016) and available on the Slandail website[1].

The result of analysing each message can be presented in the front end of the system in two ways. First, by aggregating the frequency of the occurrence of disaster-related concepts into a time series and visualising them as a dynamic time series plot that can be adjusted by time. The second way is in the main interface, which includes the base map, and each geolocated messages is indicated as a point on this map. An example of the front end of the social media monitor is shown in Figure 2.

The location of each message collected and analysed is used to produce a map layer to visualise where messages were published or are referring to, this can be seen in Figure 2 where OpenStreetMap is used as the base map. If the geographic coordinates are included in the metadata of the tweet then this is used as the exact location of the message (red marker). In many cases this location data is not included, in this case the location attributed to the

[1] www.slandail.eu

user profile can be used. This location is more frequently included in the metadata of the tweets. This can be combined and cross-referenced with references to locations in the text (green marker). The coordinates of places names detected are found by using Google Maps Geocoding API[2]. As the query formulated for collecting tweets using the API is set to only include tweets that are geolocated in the area of interest, for the purpose of demonstration, we assume that all tweets returned are relevant. The front end shows how the frequency of relevant terms and topics being discussed in a location on social media has evolved over time as several time series graphs. Each message with a location is represented as a point on the base map and has a clickable dialogue that allows the message, its accompanying metadata and any images posted to be viewed. A breakdown of all relevant messages for the region is also shown (right most dialogue window Figure 2). This dialogue window allows a user to quickly scan through up-to-date messages that have been posted and contain either an image or message content related to one of several categories in the Slandail disaster terminology, or one of the classes of the image classifier. The particular example shown in Figure 2 includes messages containing information about floods, power outages, and road damage.

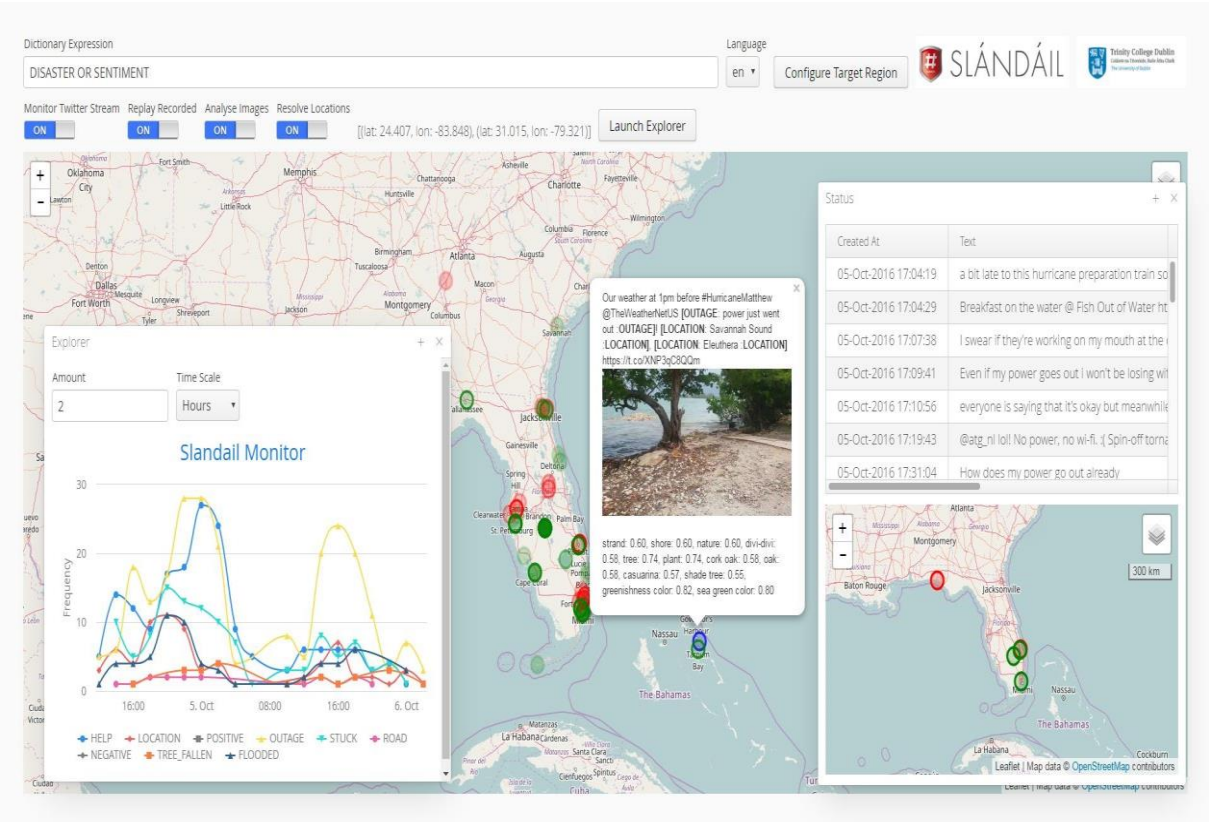

Figure 2. Front end of the social media monitor system showing filtered and relevant geolocated messages and images for the location under investigation, the frequency at which these messages have occurred, and a drill down menu to examine the messages and accompanying metadata.

A novel framework has been proposed and published as part of the Slandail project that allows the integration of image and text data from social media messages to improve image recognition in the event of a natural disaster being reported on social media. An implementation and trained model has been incorporated into the social media monitor presented here to facilitate the real-time recognition of images. For image feature extraction, all images are first resized to a standard size to have maximum height of 480. Colour images are converted to greyscale in order to calculate SIFT or SURF features based on the key interest points detected (Jing, Coleman, Bryan, & McGinnity, 2015; Jing, Scotney, Coleman, McGinnity, et al., 2016). For SIFT or SURF features, the key interest points are first detected and then the feature descriptors are calculated. The final image feature used for classification is the histogram of visual words based on the Bag-of-Words model (Niebles, Wang, & Fei-Fei, 2008).

Image features are extracted using the Bag-of-Words model and has been evaluated using two different flooding image corpora, one from the US Federal Emergency Management Agency's media library and another from

[2] https://developers.google.com/maps/documentation/geocoding

publicly available Facebook pages (Jing, Scotney, Coleman, & McGinnity, 2016; Jing, Scotney, Coleman, McGinnity, et al., 2016). The work presented by Jing et al (Jing, Scotney, Coleman, McGinnity, et al., 2016) utilises the proposed SIPF algorithm (Jing et al., 2015) and highlights the feasibility of real-time image analytics and incorporates the method into the monitor presented here for the recognition of images related to an emergency event.

Image and text features are fused during the learning process, where the model is trained on sample data. In one implementation of the image and text recognition system contained in the Slandail monitoring system, the image and text features are combined on the feature level. Images and their accompanying text captions or text descriptions are turned into features using the approach described previously in this section and are then concatenated into a single feature vector. It is assumed that these features are related semantically. The text accompanying an image often describes or elaborates on the image. A corpus was built containing images and accompanying text captions for classes related to emergency and disaster topics. Training a classifier with these feature vectors allows both text and image content related to the disaster or emergency classes to be detected. The sample corpus has been made available online and linked the Slandail Social Media Monitor[3].

Classification is performed by using a Support Vector Machine (SVM) and trained on the fused features with a test set consisting of fused features also (Jing, Scotney, Coleman, McGinnity, et al., 2016). The main approach presented shows an increase in the performance of image recognition by incorporating the text features extracted from the content of the accompanying messages. The mean average precision of the classifier is seen to improve in every instance where text and image features are used together to classify a social media message as opposed to just the image alone and demonstrates an advantage in recognition and reducing the burden of interpretation for an emergency manager (Jing, Scotney, Coleman, & McGinnity, 2016; Jing, Scotney, Coleman, McGinnity, et al., 2016).

Privacy preserving measures are incorporated into the system to help monitor the level of sensitive information being detected. To do this the presence of entities such as individual names, institutions, places and events that occur in the messages are detected using named entity recognition. The frequency occurrence of these messages in time is monitored and an *intrusion index* is formed to log whether the system is detecting a high degree of potentially sensitive information and was first presented in (Kelly & Ahmad, 2014, 2015). The system uses the Stanford Name Entity Recogniser (NER) for detecting entities (Toutanova, Klein, Manning, & Singer, 2003). The frequency of these entities, specifically person and place names, is recorded for each time period to compute a time series of entity occurrence. By locating and keeping a record of entities that have occurred in text, emergency managers and users of the system can measure the level of intrusion and personally identifiable information being collected. This may help emergency managers using the Slandail Social Media Monitor to be aware of the potential intrusiveness of the system and to help in deciding whether to anonymise stored data or delete data collected after the period of an emergency. An example of the output of the intrusion index includes a time series log of entities detected in messages (Figure 3).

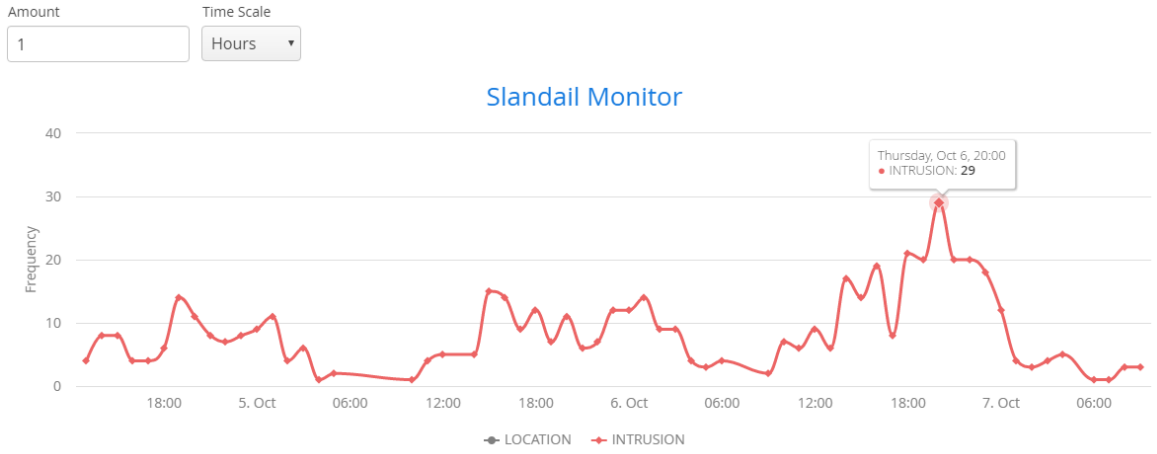

Figure 3. Slandail Social Media Monitor's visualisation of the *Intrusion Index* displaying the volume of entities.

Figure 3 shows a screen shot of the dialogue window from the Slandail Social Media Monitor indicating the frequency of entities occurring in the corpus of messages. This data is shown as a graph in the monitor front end

[3] https://sites.google.com/a/tcd.ie/slaindail-image/

but also stored as a log file. Typically an increase in message volume correlates with an increase in the number of entities detected. Normalising the series by accounting for the total number of words versus the occurrence of entities can give a better indication of periods where a higher than normal number of sensitives messages have been detected. The individual messages are flagged as containing named entities and have a higher potential of containing sensitive information. This can be logged and stored as meta-data with any messages that may be kept. End users and emergency managers within the Slandail project expressed interest in being able to recount the historical data that they may have made decisions on and to review the efficacy and justifications for these decisions at a later date. For the Slandail Social Media Monitor an online version of the system has been deployed on a standalone web server which allows the system to operate independently of a local machine and can be accessed from both mobile and standard web browsers for demonstration purposes[4].

DEMONSTRATION

To demonstrate the system's ability to filter a stream of social media data we use a corpus of Twitter messages retrieved using the public Twitter API for the period between the 20th of December 2015 to the 2nd of January 2016 (Table 1). We constructed the query for the API so as to only return messages that were geolocated in Ireland. This introduces a spatial dimension to the data as we are aware that the corpus is relevant to the location of Ireland. This time period was chosen as a number of flooding events and storms had occurred in Ireland during this time. The system was used to detect flooding events, or mentions of such on social media, during this period based on the content of the messages being posted in Ireland. Of the messages collected during this period we find that just over 6% of the messages randomly sample from the geolocation of Ireland were in reference to the storms landing (Table 1). There is little difference in negative or positive sentiment expressed during this time. An inspection of the disaster and emergency related messages shows people mainly discussing factual events or expressing information and opinions about the adverse weather conditions.

Table 1. Shows the number of Twitter messages with terms from the dictionary categories and the total number of terms in the text collection for each category. The sample includes Twitter messages geolocated in Ireland for the period from 20/12/2015 to 02/01/2016.

Category	Messages	Terms
Total	7944	161,885
Negative	2542	3569
Positive	2682	3183
Disaster &Emergency	484	500

The temporal component of each message together with their known location is used to identify conversations about flooding events on Twitter in Ireland during the sample period. By analysing the content of Twitter messages for an area, the topics of conversation can be compared to real world events happening on the ground. This type of information is very beneficial for informing emergency managers about ongoing events.

[4] Access to the online system can be granted by request to the authors

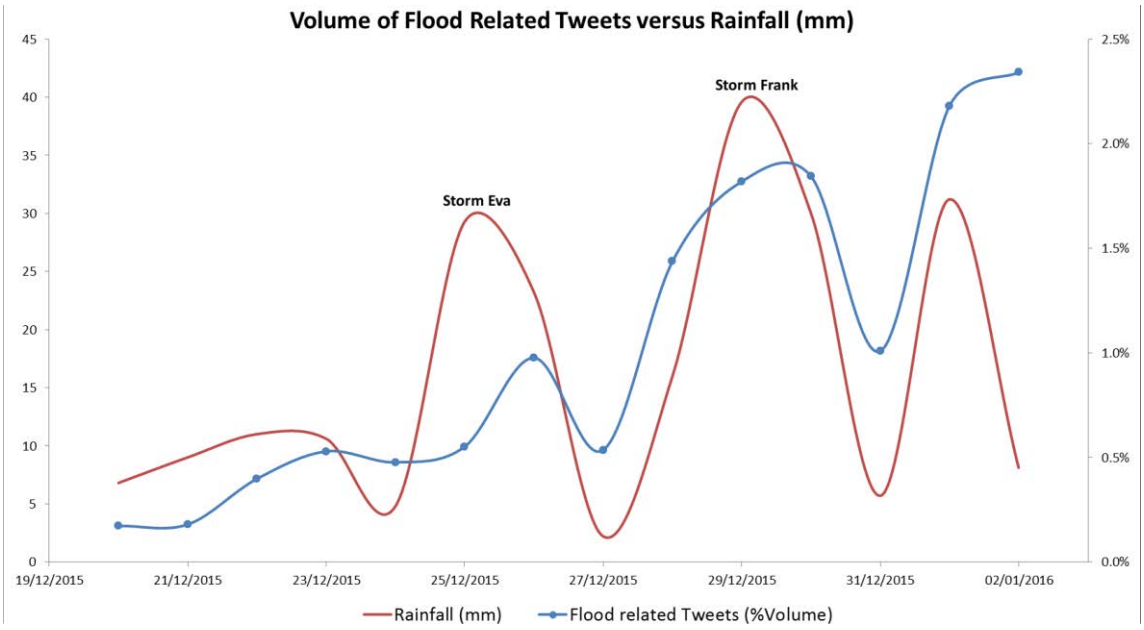

Figure 4. Shows the normalised volume of tweets tagged as begin geolocated in Ireland and identified as being relevant for flood related topics as compared to rainfall data during the landfall of several storms in Cork, Ireland between December 2015 and January 2016.

For the event detection, a ground truth was referenced based on average daily rainfall data for the region of Ireland. Meteorological data from the Irish national weather authority MetEireann[5] was used to create a time series of rainfall data for the sample period of December 2015 and January 2016. This period included the time of landfall of several major storms in Ireland. The rainfall increased at the time of landfall of these storms. The objective was to determine if the Slandail Social Media Monitor would detect an increase in the conversation about flooding and adverse weather conditions in Ireland on Twitter during this period.

From Figure 4 an increase in the volume of tweets in Ireland about flood related topics is seen and by observation looks to be correlated with the landfall of major storms and the increase in rainfall. The correlation of the two series during this period is seen to be 54%. This demonstrates the system's ability to detect conversations regarding real world events on Twitter by including the spatio-temporal characteristics of the data.

LIMITATIONS AND FUTURE WORK

A limitation of the system has been in user engagement and incorporating the output of the geolocated social media analysis into existing decision making processes for emergency managers. The objective of the Slandail Social Media Monitor has been to analyse and aggregated relevant social media messages containing content that would be of interest to an emergency manager and present it in a way that is familiar to them. This has included filtering the social media stream and presenting it as geolocated information on a map. Future and on-going work has included aggregating the output of the social media analysis in the Slandail monitor with data and variables that an emergency manager might typically use such as location of potential hazards, population density of an area, or proximity to bodies of water that may flood. A model has been proposed in forthcoming work by the authors that relies on fuzzy aggregation theory to incorporate the output of the social media analysis with additional data that emergency managers can use in their decision making process. This model weights the severity of information measured at a particular time and uses a type of weighted average to produce a single output that can be used to inform and support decision making (Mesiarová-Zemánková, Kelly, & Ahmad, 2017). Another concern for emergency managers and the Slandail monitor is how to handle rumours and false information. In this case, methods are being investigated to weight the reliability of a message based on heuristics about the user and their credibility (Gupta & Kumaraguru, 2012).

A powerful application of this system is to visualise analysis by location. Those who are first to encounter these areas are often locals of the area of people using the roads or amenities in the flooded areas. These people may also take to social media to announce problem areas and flood roads as a matter of interest or to inform authorities and the general public. To this end the system identifies messages and images that are relevant to the

[5] http://www.met.ie/climate/daily-data.asp

topic of flooding and combined with the geolocation of the messages can visualise this data to see potential areas of interest for an emergency responder. An example of this output for the Florida storms in the U.S has been shown in Figure 2. Here a larger scale study is currently being investigated where over a million tweets consisting of messages and images have been compiled during the landfall of Hurricane Mathew in October 2016. This study which is part of our future work looks at the time serial reporting of events and categorizes messages into the phases of emergency including preparedness and recovery in order to better understand the dynamics of how disasters are reported and discussed on social media.

CONCLUSION

This work has described the implementation of a system that combines the spatio-temporal aspects of social media data with the content of messages and images posted to create insights into unfolding emergency related events as reported on social media. The system has used content analysis and image analytics to filter messages and identify content related to emergency and disaster situations. It has been demonstrated that a correlation exists with a corpus of data which tracks the topic of conversation on Twitter in Ireland about flooding and the landfall of two major storms. Given the large volume of messages published by individuals on Twitter, it is clear that by combining the spatio-temporal attributes of messages with the content and images posted, sensitive information about people and their locations can be identified. A method has been integrated into the system that facilitates the ethical harvesting of data in the form of an intrusion index. This allows a user of the system to be more informed about the potential intrusiveness of the system. Lastly, the output of the system includes the creation of a map layer and time series data that allows the results and analysis of the system to be visualised. These visualisations are incorporated into a front end that displays the geolocation of the messages and system output on an interactive map with dynamic graphs in separate dialogue windows. The system described here can be used to better inform decision making for a user, and to better organise and filter relevant content from social media. The demonstration of the system shows an application in the area of emergency and disaster management where receiving timely, relevant information about ongoing events and data privacy can be addressed.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European community's Seventh Framework Programme under grant agreement No.607691 (SLANDAIL). The materials presented and views expressed here are the responsibility of the author(s) only. The EU Commission takes no responsibility for any use made of the information set out.

REFERENCES

Adam, Nabil R, Shafiq, Basit, & Staffin, Robin. (2012). Spatial computing and social media in the context of disaster management. IEEE Intelligent Systems, 27(6), 90-96.

Aiello, Luca Maria, Petkos, Georgios, Martin, Carlos, Corney, David, Papadopoulos, Symeon, Skraba, Ryan, . . . Jaimes, Alejandro. (2013). Sensing trending topics in Twitter. IEEE Transactions on Multimedia, 15(6), 1268-1282.

Althoff, Tim, Borth, Damian, Hees, Jörn, & Dengel, Andreas. (2013). Analysis and forecasting of trending topics in online media streams. Paper presented at the Proceedings of the 21st ACM international conference on Multimedia.

Chang, Angel X, & Manning, Christopher D. (2014). TokensRegex: Defining cascaded regular expressions over tokens. Tech. Rep. CSTR 2014-02.

Chowdhury, Soudip Roy, Imran, Muhammad, Asghar, Muhammad Rizwan, Amer-Yahia, Sihem, & Castillo, Carlos. (2013). Tweet4act: Using incident-specific profiles for classifying crisis-related messages. Paper presented at the 10th International ISCRAM Conference.

Derczynski, Leon, Ritter, Alan, Clark, Sam, & Bontcheva, Kalina. (2013). Twitter Part-of-Speech Tagging for All: Overcoming Sparse and Noisy Data. Paper presented at the RANLP.

Dou, Wenwen, Wang, Xiaoyu, Skau, Drew, Ribarsky, William, & Zhou, Michelle X. (2012). Leadline: Interactive visual analysis of text data through event identification and exploration. Paper presented at the Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on.

Gupta, Aditi, & Kumaraguru, Ponnurangam. (2012). Credibility ranking of tweets during high impact events. Paper presented at the Proceedings of the 1st Workshop on Privacy and Security in Online Social Media.

Imran, Muhammad, Castillo, Carlos, Diaz, Fernando, & Vieweg, Sarah. (2015). Processing social media messages in mass emergency: A survey. ACM Computing Surveys (CSUR), 47(4), 67.

Imran, Muhammad, Elbassuoni, Shady Mamoon, Castillo, Carlos, Diaz, Fernando, & Meier, Patrick. (2013). Extracting information nuggets from disaster-related messages in social media. Proc. of ISCRAM, Baden-Baden, Germany.

Jin, Xin, Gallagher, Andrew, Cao, Liangliang, Luo, Jiebo, & Han, Jiawei. (2010). The wisdom of social multimedia: using flickr for prediction and forecast. Paper presented at the Proceedings of the 18th ACM international conference on Multimedia.

Jing, Min, Coleman, Sonya A, Bryan, Scotney, & McGinnity, TM. (2015). Multiscale "Squiral"(Square-Spiral) Image Processing.

Jing, Min, Scotney, Bryan W, Coleman, Sonya A, & McGinnity, Martin T. (2016). The Application of Social Media Image Analysis to an Emergency Management System. Paper presented at the Availability, Reliability and Security (ARES), 2016 11th International Conference on.

Jing, Min, Scotney, Bryan W, Coleman, Sonya A, McGinnity, Martin T, Zhang, Xiubo, Kelly, Stephen, . . . Heyer, Gerhard. (2016). Integration of text and image analysis for flood event image recognition. Paper presented at the Signals and Systems Conference (ISSC), 2016 27th Irish.

Kelly, Stephen, & Ahmad, Khurshid. (2014). Determining levels of urgency and anxiety during a natural disaster: noise, affect, and news in social media. Paper presented at the DIMPLE: DIsaster Management and Principled Large-Scale information Extraction Workshop Programme.

Kelly, Stephen, & Ahmad, Khurshid. (2015). Propagating Disaster Warnings on Social and Digital Media. Paper presented at the International Conference on Intelligent Data Engineering and Automated Learning.

Li, Rui, Lei, Kin Hou, Khadiwala, Ravi, & Chang, Kevin Chen-Chuan. (2012). Tedas: A twitter-based event detection and analysis system. Paper presented at the Data engineering (icde), 2012 ieee 28th international conference on.

Mathioudakis, Michael, & Koudas, Nick. (2010). Twittermonitor: trend detection over the twitter stream. Paper presented at the Proceedings of the 2010 ACM SIGMOD International Conference on Management of data.

Mesiarová-Zemánková, Andrea, Kelly, Stephen, & Ahmad, Khurshid. (2017). Bonferroni mean with asymmetric weights.

Niebles, Juan Carlos, Wang, Hongcheng, & Fei-Fei, Li. (2008). Unsupervised learning of human action categories using spatial-temporal words. International journal of computer vision, 79(3), 299-318.

Ritter, Alan, Etzioni, Oren, & Clark, Sam. (2012). Open domain event extraction from twitter. Paper presented at the Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining.

Robinson, Bella, Power, Robert, & Cameron, Mark. (2013). A sensitive twitter earthquake detector. Paper presented at the Proceedings of the 22nd International Conference on World Wide Web.

Sakaki, Takeshi, Okazaki, Makoto, & Matsuo, Yutaka. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. Paper presented at the Proceedings of the 19th international conference on World wide web.

Spyropoulou, Maria. (2016). Emotive and Terminological Content in Disaster Related Messages. EMOT: Emotions, Metaphors, Ontology and Terminology during Disasters, 9.

Sutton, Jeannette N, Palen, Leysia, & Shklovski, Irina. (2008). Backchannels on the front lines: Emergency uses of social media in the 2007 Southern California Wildfires: University of Colorado.

Toutanova, Kristina, Klein, Dan, Manning, Christopher D, & Singer, Yoram. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. Paper presented at the Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1.

Vieweg, Sarah, Hughes, Amanda L, Starbird, Kate, & Palen, Leysia. (2010). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.

Yin, Jie, Lampert, Andrew, Cameron, Mark, Robinson, Bella, & Power, Robert. (2012). Using social media to enhance emergency situation awareness. IEEE Intelligent Systems, 27(6), 52-59.

Zhang, Xiubo, Kelly, Stephen, & Ahmad, Khurshid. (2016). The Slandail Monitor Real-time Processing and Visualisation of Social Media Data for Emergency Management. Paper presented at the MUININ -

International Conference on Availability, Reliability, and Security