

Emergency Identification and Analysis with EAIMS

Richard McCreadie, Craig Macdonald and Iadh Ounis
 {firstname.lastname}@glasgow.ac.uk
 University of Glasgow, UK

ABSTRACT

Social media platforms are now a key source of information for a large segment of the public. As such, these platforms have a great potential as a means to provide real-time information to emergency management agencies. Moreover, during an emergency, these agencies are very interested in social media as a means to find public-driven response efforts, as well as to track how their handling of that emergency is being perceived. However, there is currently a lack of advanced tools designed for monitoring social media during emergencies. The Emergency Analysis Identification and Management System (EAIMS) is a prototype service that aims to fill this technology gap by providing richer analytic and exploration tools than current solutions. In particular, EAIMS provides real-time detection of emergency events, related information finding, information access and credibility analysis tools for use over social media during emergencies.

1. INTRODUCTION

Social media platforms are some of the most popular information sharing mediums in use today. They provide a wealth of real-time information, which, if used effectively can provide insights about what is happening in the world. Importantly, during emergency situations such as natural disasters, the information shared on social media has been shown to be valuable to enhance the emergency response [8]. In particular, experts in a variety of domains have developed tools to use social media to identify and track events [1, 2, 3, 10, 11, 12]. For example, the Crisis Tracker system [12] makes use of a manually created set of search terms to crawl crisis-related content, which is then clustered automatically and manually annotated by volunteers. In general, social media tracking and analysis tools can collect, enrich, aggregate and visualize information from people involved in an event, enabling both the general public and emergency response agencies to better monitor what is happening in real-time.

However, many current crisis tracking and analysis tools rely on real-time volunteer (“the crowd”) effort to enable

or enhance core functions, such as information categorization [3], location identification [9] or information verification [12]. This reliance on humans for core functionalities can be disadvantageous, particularly during the outset of a disaster, where volunteers may not yet be in place. Furthermore, current fully automatic platforms such as Twitris [11] and Twitcident [1] are tied to one social media platform, and hence are not applicable to deployment in regions where those platforms are less prevalent. On the other hand, a wider array of automatic solutions for real-time event detection in social streams [7], sentiment analysis [13], timeline summarization [5] and credibility estimation of social content are becoming available, which present new opportunities to provide enhanced automatic tools for disaster monitoring on social media.

Hence, we present a new crisis tracking and analysis toolkit, named: EAIMS (Emergency Analysis Identification and Management System), which aims to provide automatic identification, tracking, summarization and exploration tools over social content generated during emergency events. EAIMS differs from most existing crisis analysis toolkits, in that it provides a richer suite of automatic tracking and analysis tools, namely: automatic event detection, timeline construction, targeted sentiment analysis, user-community identification and information credibility estimation, as well as information access support tools in the form of social media search and discussion threads ranking and visualization. EAIMS is social media platform-independent, making it applicable to deployment in regions that use alternative platforms.

2. EAIMS FUNCTIONALITY

To illustrate the functionalities of the EAIMS toolkit, we describe how EAIMS might be used during a real event. More precisely, we use the large earthquake that occurred during 24th of August 2016 in central Italy as our example when using Twitter as a social media data source, as illustrated in Figures 1 and 2.

EAIMS supports two operational levels, namely strategic and tactical, based on requirements defined by civil protection and law enforcement agencies [8]. The strategic level supports the early identification of an event for the end-user using social media as a sensor for what is happening in the world. In particular, users are able to define tracking profiles targeted for particular types of events (earthquakes in our example). Once a profile is registered, social media content is collected and first story detection (FSD) technologies [7] are used to identify posts representing candidate

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SWDM 2016 October 28th, Indianapolis, USA



Figure 1: Strategic Layer Views.

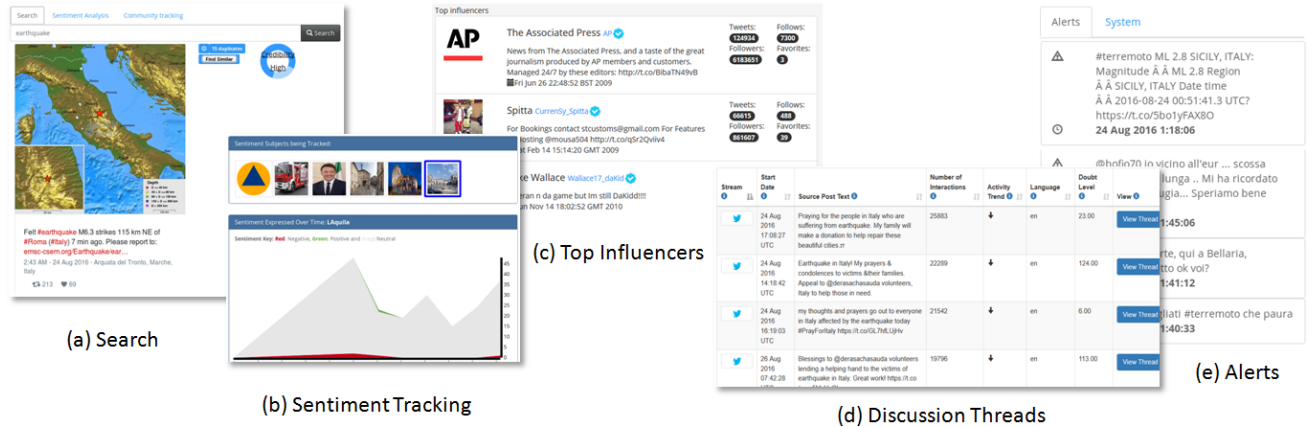


Figure 2: Tactical Layer Views.

events, which are then pushed in real-time to a map screen as pings (Figure 1 (a)). The user can view the content of these posts within the strategic UI (Figure 1 (b)). Multiple tracking policies can be used concurrently by different users, which enables the identification multiple events of different types. Once an event has occurred, the user can transition from the strategic to the tactical level. The tactical level provides a targeted data collection for a particular event, as well as a suite of data enrichment, data exploration and visualization tools for use of the collected content. In particular, once the user is confident that an emergency event is in progress, they can use the *Event Builder* function of EAIMS to configure the tracking of that event, e.g. by setting event-specific tracking terms to collect data for (Figure 1 (c)). Once the event tracking is configured, the user opens the tactical level UI for that event. The tactical level UI visualizes the output of a series of social sensors that select or enrich (e.g. by automatically adding sentiment or credibility labels) the social data collected, as well as provides social media content search and discussion threads visualization. In particular, within the tactical UI, EAIMS provides: real-time search using the Terrier Information Retrieval platform [4] with automatic credibility estimation over the ranked content (Figure 2 (a)); sentiment tracking for particular named people, organizations or places (Figure 2 (b)); top influencer identification within the social network(s) (Figure 2 (c)); visualization of the most active relevant discussion threads within the social network(s) (Figure 2 (d)); and automatic event timeline generation (Figure 2 (e)). The underlying

EAIMS services that feed the tactical visualization support content collection and analysis from multiple events concurrently, as long as those events have been registered with the event builder.

3. RECENT ADDITIONS

An earlier version of the EAIMS toolkit was published in [6], where the additional technical details of the platform can be found. The current version of the toolkit has been extended with real-time automatic extraction and storage of active discussion threads within the social media platform, as well as ranking and visualization capabilities for those discussion threads (Figure 2 (d)).

In particular, based on consultation with civil protection and police agencies, it was suggested that discussion threads on social media are a desirable means to both represent discussion topics and user communities during emergencies. Indeed, when considering reasoning about information veracity within social networks, discussion threads are a natural unit to analyze, as they are topically coherent.

However, within some social networks, discussion threads cannot be accessed directly. Instead, discussion threads are exposed as annotations on each individual post. For instance, within Twitter, discussion threads are represented by either 'in_reply_to_status_id' or 'retweet.id' annotations added to each tweet. For this reason, it is necessary to construct the discussion thread structure in real-time.

EAIMS uses an inverted index structure to construct the



Figure 3: Discussion Thread Visualization.

various discussion threads for an event over time. In this case, the inverted index stores the posting list (documents contained) for each discussion thread, while the lexicon stores the identifiers for each thread. When a new post t_n arrives on the stream, if that post is not a reply to a previous post, then it is considered the start of a new thread, and is hence added to the lexicon (we use the identifier of the first post as the thread identifier). If the post t_n is a reply, then a three-stage update operation is performed. First, the inverted index is used to get the posting list for the source post that t_n replies to, denoted t_s . This posting list is in effect the discussion thread that t_n belongs to. Second, a new ‘document’ is generated that contains the identifiers from all of the posts in the t_s posting list, as well as the identifier for t_s , i.e. this ‘document’ represents the new thread for t_n . This document is indexed, which both constructs the posting list for t_n , but also updates the posting lists for the (sub-)threads within t_n .

Conceptionally, under this approach, the inverted index maintains k copies of a discussion thread, where k is the length of that thread. Each copy is keyed by a different post within the thread. In this way, the full discussion thread can be retrieved in a single operation regardless of which post in the thread is replied to. For long-running events with very large thread sizes, storing multiple copies of each thread can be avoided by having all k posts within a thread ‘map’ down to a single posting list.

Within EAIMS, the inverted index structure can be used to find the largest discussion threads at the current moment for the user. These are visualized as a tabular overview, as shown in Figure 2 (b). A user can then select a particular thread to open a more detailed view of the post within that thread and their statistics, as illustrated in Figure 3.

4. OUTLOOK

EAIMS is a prototype tool designed to aid emergency management organizations analyze and explore social media content in real-time during emergency and/or large public events. It was developed as part of a larger EC FP7 project called SUPER (Social sensors for security assessments and proactive emergencies management). The outcomes of SUPER, including the social sensors illustrated by EAIMS are being deployed and evaluated by end-users from the Civil Protection Service of the Campania Region (Italy) during

flash flood emergencies, as well as end-users from the General Inspectorate of Romanian Police, for safety monitoring during large public events.

5. ACKNOWLEDGMENTS

This work has been carried out in the scope of the EC co-funded SUPER Project (FP7-606853).

6. REFERENCES

- [1] F. Abel, C. Hauff, G-J. Houben, R. Stronkman and K. Tao. Twitcident: Fighting Fire with Information from Social Web Streams. In *Proc. of WWW*, 2012.
- [2] P. Heim, D. Thom and T. Ertl. Semsor: Combining social and semantic web to support the analysis of emergency situations. In *Proc. of SEMAIS*, 2011.
- [3] M. Imran, C. Castillo, J. Lucas, P. Meier and S. Vieweg. AIDR: Artificial intelligence for disaster response. In *Proc. of WWW*, 2014.
- [4] C. Macdonald, R. McCreadie, R. L. Santos, and I. Ounis. From puppy to maturity: Experiences in developing terrier. *Proc. of OSIR*, 2012.
- [5] R. McCreadie, C. Macdonald, and I. Ounis. Incremental update summarization: Adaptive sentence selection based on prevalence and novelty. In *Proc. of CIKM*, 2014.
- [6] R. McCreadie, C. Macdonald, and I. Ounis. EAIMS: Emergency Analysis Identification and Management System. In *Proc. of SIGIR*, 2016.
- [7] R. McCreadie, C. Macdonald, I. Ounis, M. Osborne, and S. Petrovic. Scalable distributed event detection for Twitter. In *Proc. of IEEE Big Data*, 2013.
- [8] R. McCreadie, K. Kappler, M. Kardara, A. Kaltenbrunner, C. Macdonald, J. Soldatos and I. Ounis. SUPER: Towards the use of Social Sensors for Security Assessments and Proactive Management of Emergencies. In *Proc. of SWDM*, 2015.
- [9] O. Okolloh. Ushahidi, or “testimony”: Web 2.0 tools for crowdsourcing crisis information. In *Participatory Learning and Action Journal*, 2009.
- [10] M. Osborne *et al.* Real-Time Detection, Tracking, and Monitoring of Automatically Discovered In *Proc. of ACL*, 2014.
- [11] H. Purohit and A. P. Sheth. Twitris v3: From Citizen Sensing to Analysis, Coordination and Action. In *Proc. of ICWSM*, 2013.
- [12] J. Rogstadius, M. Vukovic, C. A. Teixeira, V. Kostakos, E. Karapanos, J. A. Laredo. CrisisTracker: Crowdsourced social media curation for disaster awareness. *IBM Journal of Research and Development*, 57(5), 2013.
- [13] S. Vargas, R. McCreadie, C. Macdonald, Craig and I. Ounis. Comparing Overall and Targeted Sentiments in Social Media during Crises. In *Proc. of ICWSM*, 2016.