

TORCIA

A Decision-support Collaborative Platform for Emergency Management

Chiara Francalanci and Paolo Giacomazzi

Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, P.zza Leonardo da Vinci 32, Milano, Italy

Keywords: Emergency Management, Crisis Management, Social Media, Twitter, Semantic Engine.

Abstract: The TORCIA platform has been developed as part of a project funded by the Lombardy Region. The main goal of the project is the development of a tool that leverages social media in emergency management processes. With a continuous and real-time collection of information from social media, TORCIA can detect situations of potential emergency and identify their geographical position. The TORCIA platform supports emergency operators from different organizations and institutions with a decision-support dashboard and favors the creation of a collaborative process combining the contributions of citizens and institutions by means of a mobile app that is also integrated within the platform.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

TORCIA is a project funded by the Lombardy Region (Italy) with the support of the European Fund for Regional Development. The main goal of the project is the development of a software platform that leverages social media in emergency management processes and, more generally, in the planning and control processes of crisis situations (e.g. accidents or natural disasters), with a focus on roads and transportation issues. The project has started in June 2012 and ended in June 2014. The partnership of the project includes Alcatel-Lucent, main contractor of TORCIA, the Department of Electronics, Information and Bioengineering (DEIB) of Politecnico di Milano, scientific coordinator of TORCIA, Fondazione Politecnico di Milano, ACT Solutions, Beta 80 and Vidiemme.

Italian cities, on the basis of their local specificities, have defined plans to respond to possible emergencies designed in cooperation with the National Civil Protection. These plans include a series of tasks, roles and policies to be enacted in case of a given natural or man-made emergency situation. The transport network is a key infrastructure to be managed in emergency situations. The emergency plans of the Civil Protection can include a number of changes to the transport infrastructure in case of emergency, such as the direction of travel in specific avenues or highways, or access restrictions to specific areas that

are needed by emergency vehicles. Ensuring safety is key to an effective management of emergency situations and controlling the transport infrastructure is among the main drivers of safety.

In this context, the TORCIA project aims at supporting the collection of information that allows institutions to detect critical situations that could cause an emergency and to make decisions that decrease the risk of emergencies as well as their consequences when they cannot be avoided. Citizens involved in an emergency situation that are also Web 2.0 users can provide an important contribution to emergency management by sharing information on social networks, Facebook and Twitter above all. In order to collect, organize and interpret this information and make it useful to the emergency management teams, the TORCIA project has implemented a software platform that is based on a geographically distributed cloud infrastructure connected by a broadband optical network. TORCIA can take advantage of the cooperation of citizens by analyzing the information that they share, using social networks according to an overall information workflow that is integrated and technically innovative. Figure 1 shows the technical architecture of the TORCIA platform.

2 RELATED WORKS

Leveraging social media to improve disaster

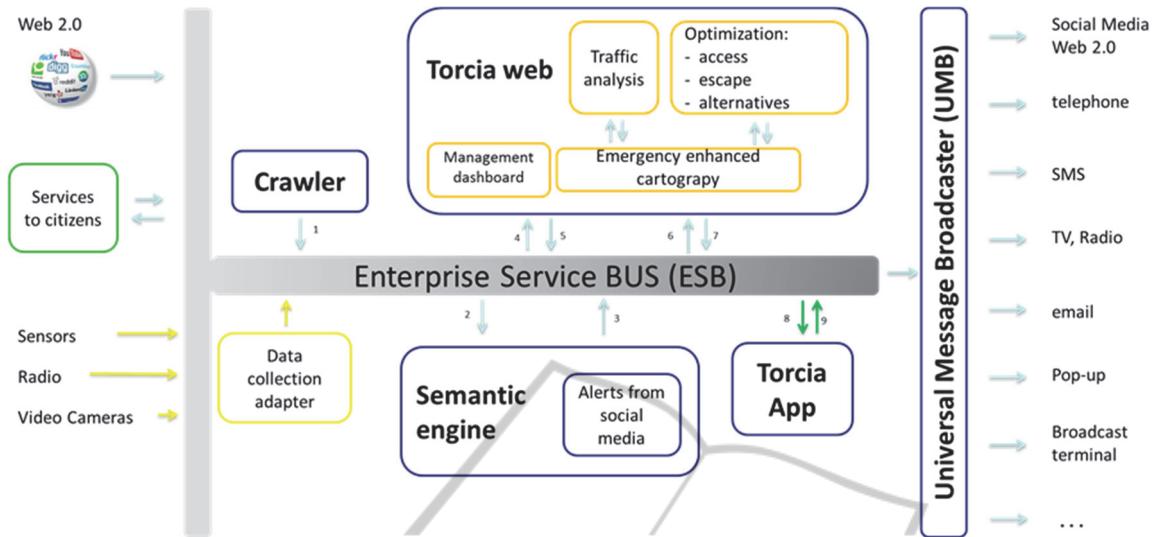


Figure 1: Architecture of the TORCIA platform.

management processes has attracted the attention of researchers in several different areas (Purohit, 2013). On one hand, there is a whole body of literature focusing on how to integrate social media into official emergency management plans and practices (Brugemanns, 2012). This literature points to the challenges raised by social media and their unique collaboration features.

On the other hand, significant research contributions focus on the technical issues to be addressed in order for social media information to be fully exploited. The technical challenges with gathering, analysing, and sharing social information span from the identification of relevant information from a potentially overwhelming number of social media messages to classifying the conversation topics and locating the geographical areas impacted by the emergency. For example, authors of (Saravanou, 2015) provides a methodology to identify the areas that have been hit the most by a disaster based on geographical clustering. In (Imran, 2015), tweets are analysed semantically to dynamically gather the topics of social conversations. The assessment of the damage caused by a disaster is performed automatically based on social media information in (Cresci, 2015).

The TORCIA project is positioned in this second stream of literature, focusing on technology rather than processes. The main contribution of the TORCIA project is to provide an end-to-end platform that integrates multiple technical components that are typically designed and tested in isolation in previous literature (see Figure 1). The technical components of the TORCIA platform are

coordinated according to the TORCIA workflow, which represents an overall approach to the management of the social information lifecycle during. By integrating multiple technical components and testing the platform on several real emergencies over a two-year time frame, TORCIA has proved to be able not only to effectively gather and classify information, but also to raise alerts with a significant time advantage compared to the standard alert channels.

3 SOCIAL MEDIA

The TORCIA workflow starts from the real-time continuous collection of information from the most important social media (for a high-level visual description of the TORCIA workflow see <http://sostorcia.it/it/news/141-il-video-del-progetto-torcacia.html>). The first research question that has been faced is whether social media can really provide information that is relevant in the context of emergency management and useful to reach our project objectives. The answer to this question is yes if social media users, that is citizens, talk about emergency situations on social media and if that buzz can be useful for emergency management. These conditions have been verified as part of a preliminary feasibility analysis that has been conducted at the beginning of the TORCIA project and is summarized in the following.

A first, straightforward consideration is that if an emergency occurs in a scarcely populated geographical site, it is very likely that there is not a

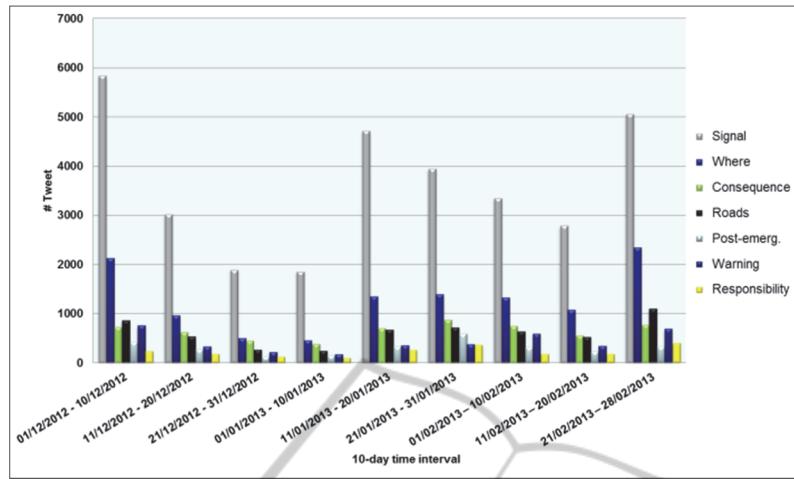


Figure 2: Topics of buzz on floods (buzz in Italian).

sufficient number of social media users among the citizens involved and, as a consequence, social media will provide very little or no information. For example, a landslide blocking a solitary mountain path is very likely not to receive attention on social media by the few people involved. Conversely, a landslide in a famous path (such as the landslide that has blocked the “love path” at the Cinque Terre in Italy in 2013) will be immediately signalled on social media. In general, the higher the number of people involved in a same emergency situation, the greater the value of social media information in the management of that emergency. We have chosen to focus on the most favourable scenario, that is city emergencies and, in particular, floods, which have recently become more frequent, especially in Italy, and, thus, suitable to provide numerous test cases within the time frame of our project. With reference to floods in the context of a city, we have conducted our preliminary analysis with the following objectives: (1) verify the volumes of buzz on floods, particularly from Twitter; (2) analyze the topics of the buzz on floods; (3) verify whether the buzz on floods provides information useful to geolocalize the emergency; (4) evaluate the presence of the institutions on social media, particularly on Twitter; (5) compare the Italian case with international best practices. Crawling, i.e. the collection of posts from social media, has started in September 2013 on roughly 60 keywords that define the domain of interest, i.e. floods, in Italian and English. A sample list of english keywords is reported in Table 1. Overall, we have collected over 40 million tweets during the project. The volumes of buzz reported in Table 1 represent the number of occurrences of each keyword in the Italian language in a typical day with

bad weather (without any real emergency situation). It can be noted how buzz concentrates on a subset of keywords. Our preliminary qualitative analyses of buzz have shown that the subset of most frequent keywords tends to change over time, as a consequence of the changing weather, but also as a result of a tendency to imitate each other in the choice of terms when describing common weather conditions. A first finding is that the volumes of buzz in the flood domain in Italian are significant, on average 40,000 tweet/month and worth of further analysis (see also (Rossitto, 2012)).

Table 1: Crawling keyword used for data collection and related number of daily occurrences.

Term	Number of occurrences
Flood/s, flooding	362
Hailstorm/s, hail	96
Storm/s	337
Lanslide/s	792
Hurricane	69
Cloudburst	62

Our next step has been to analyze the social media buzz on floods in order to gather the most frequent topics (Cesana 2012).

The goal was to verify whether social media buzz includes information that could be potentially interesting in emergency management (Cameron, 2012), (Shih, 2012). The most frequent topics are: (a) Signal, i.e. posts signaling bad weather conditions and evaluating their severity (from mere inconvenience to warning of a potential or ongoing emergency situation). (b) Where, i.e. posts explicitly referring to the place where the user is experiencing bad weather conditions. (c) Consequence, i.e. posts describing facts that have occurred as a consequence

of bad weather conditions (flooding, landslide, collapse of a building, etc.). (d) Roads, i.e. posts reporting on traffic and road conditions. (e) Post-emergency, i.e. posts commenting on recent events that have created emergency situations, often with reference to ongoing activities aimed at restoring normal conditions.

(f) Warning, i.e. posts commenting on bad weather forecasts or informing on the potential risks associated with future bad weather conditions. (g) Responsibility, i.e. posts expressing opinions on the causes and related responsibilities of the consequences of bad weather conditions.

Figure 2 reports the distribution of buzz into the topics discussed above. It can be noted how buzz concentrates on the signal category, which, in turn, indicates a prevalence of social interactions during the emergency, as opposed to the phases preceding and following the emergency. The fact that pre- and post-emergency receive a lower degree of attention is partly related to a very limited presence of Italian institutions and authorities on Twitter. It can also be noted that the volumes of buzz in the where and roads categories are significant. This indicates that social media buzz may provide information useful for the definition of the geographical zone impacted by the emergency and for the assessment of the conditions of the road infrastructure.

Overall, these observations indicate that: (1) The collection and analysis of information from social media, Twitter in particular, should be performed in real time, given the high percentage of posts shared during the emergency (a time frame of a few hours). (2) The information that is collected can indeed provide useful insights on the geographical position

and extent of the emergency and, to a more limited degree, on issues with the road infrastructure.

4 THE SEMANTIC ENGINE

Not all information is useful for emergency management. Roughly half of the information that is collected syntactically by means of crawling keywords is not related to the domain. Figure 3 shows how the floods domain is not an exception, indicating that 50% of the posts collected with the crawling keywords exemplified in Table 1 are not discussing flood-related issues. In the TORCIA architecture (see Figure 1), the semantic engine (Caracaci, 2012) is in charge of the identification and removal of irrelevant information. The semantic engine classifies relevant posts into the topics shown in Figure 3 (also called categories). This classification is performed by a software module based on a semantic network that associates metadata with each topic and weighs them according to their importance in the application domain. Table 2 shows a few instances of the metadata supporting the classification activity. The selection of metadata and the tuning of weights in the semantic network have been repeated several times during the project to continuously improve the quality of the classification process.

Finally, the semantic engine analyzes the information that has been previously cleaned and classified in order to identify situations of potential or ongoing emergency. In the first case (potential emergency), the engine raises an alert, while in the

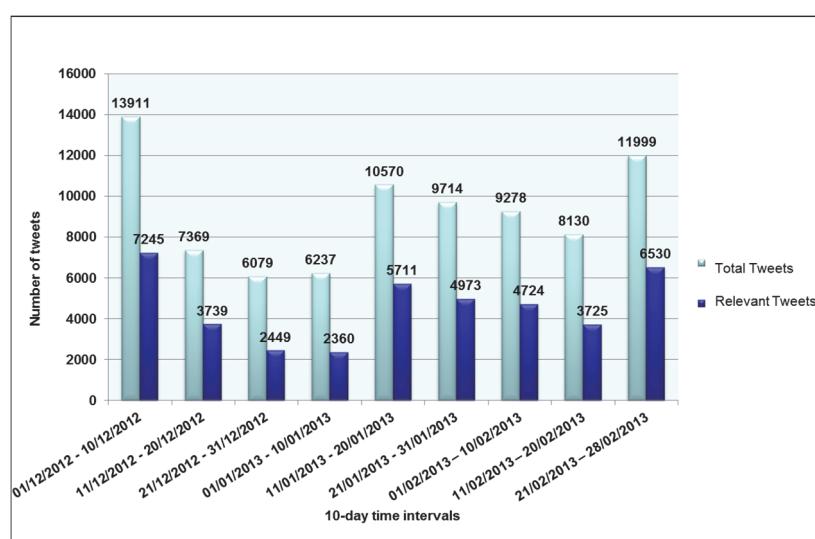


Figure 3: Volume of posts collected syntactically vs. Volume of posts relevant to the floods domain (example).

second case (ongoing emergency) it raises an alarm. The identification of a potential or ongoing emergency is based on the detection of peaks in the volumes of buzz compared to the average volumes. We have observed how not all peaks correspond to an actual emergency situation, either potential or ongoing, and, therefore, setting volume thresholds to create alerts or alarms results into an overwhelming number of false positives which is significantly higher than the correct signals. For example, bad weather conditions in a vast zone of the national territory causes a considerable increase in the volumes of posts, but in most cases those peaks do not correspond to emergencies. Those peaks can be even higher than the levels of buzz during a real emergency impacting on a specific geographical zone, such as the volumes reached during the flood experienced by the city of Catania (Sicily, Italy) in 2013 or by the Island of Sardegna in 2014. Both floods have been severe and have generated corresponding peaks that we have been able to identify by combining two thresholds: 1) a threshold on the overall volumes of buzz and 2) a second threshold on volumes of posts consistently referring to the same geographical zone (e.g. a region, such as Sardegna) or place (e.g. a city, such as Catania). This second threshold is applied to posts in the “where” and “roads” categories (see Figure 2). The identification of geographical zones or places is supported by a geographical mapping service designed by Vidiemme, a partner in the TORCIA project. Given that the majority of users disable their

Table 2: Topics and sample metadata.

Topics	Metadata (examples)
Where	City, Center, Road, Underground, North, Spain, Venice, Site, Town, Citym Region, Zone...
Signal	Flood, Rain, River, Water, Storm, Thunderstorm, Hailstorm, Hurricane, Cloudburst...
Consequence	Dead, Rescue, Landslide, Situation, Damage, Collapse, Drown...
Responsibility	Police, Judge, Court, Fire brigades, Manager, Ministry, Civil Protection, City Hall...
Roads	Open, Closed, Accessible, Clear, Blocked, Congested...
Warning	Risk, Danger, Forecast, Sewer, Clogged, River bank, Dam, Dyke...
Post emergency	Refund, Complaint, Money, Fund, Statement, Notification, Charge, Restoration, Reinstallation...

smartphone’s location service, geographical information is gathered from text. tweets are cross-checked with the geographical mapping service to verify whether they are names of places.

5 THE DASHBOARD

Alerts and alarms are conveyed to the management dashboard (implemented by Beta80, partner of the TORCIA project). During the project, we have been able to test the TORCIA architecture on several emergency situations, including: (1) Flood, Catania (Italian city), February 2013; (2) Flood, Vicenza (Italian city), May 2013; (3) (first) cloudburst, Rome (Italian city), July 7/8 2013; (4) (second) cloudburst, Rome (Italian city), July 21/22 2013; (5) (third) cloudburst, Rome (Italian city), July 27/28 2013; (6) Flood, Tuscany (Italian region), October 21/22 2013; (7) Flood, Sardegna (Italian region), November 2013; (8) Flood, Marche (Italian region), April 2014; (9) Flood, Senigallia (Italian city), May 2014; (10) Flood, Seveso (Italian river), September 2014.

By analyzing these emergencies, we have realized that buzz on social media is not straightforward to use, unless it is organized, i.e. disambiguated, categorized and geolocalized. This analysis allows users to obtain interesting information that complements that obtained from more traditional information sources that are commonly used by institutional operators. With reference to the flood in Catania (see list above), The alerts raised by the semantic engine are notified to the management dashboard with a significant time advance compared to the time when the official alert is broadcasted by the Civil Protection (a few hours). For example, in the Catania flood (February 2013), the official alert is at 6 PM, while the TORCIA alert would have been raised at 4:15 PM. A similar time advance has been consistently found for all the emergencies that have been considered for testing. However, the management dashboard is strongly user-centric. All information, including alerts and alarms, are proposed to the user, assuming that he/she has to manually validate it. This validation is considered necessary both because the semantic engine, notwithstanding our continuous refinement, is subject to errors, and because not all social information is dependable. The quality of information can benefit from the insight of an expert, for example a civil protection operator, who has access to all information sources, as opposed to social media only.

6 THE MOBILE APP

The information organized by the semantic engine and validated by operators is made available to citizens by a mobile application (implemented by Beta80, a partner of the TORCIA project). In the TORCIA architecture, the mobile application represents the medium with which citizens and institutional operators cooperate before, during, and after an emergency (Estelles, 2012).

It constitutes a coordination tool and, for the institutions, a crowdsourcing mechanism. Through the app, citizens can access validated information and, directly from the app, they can contribute with additional information which can be addressed to the operators, to Twitter, or both. Operators can collect information from citizens in real time and, in turn, send institutional messages critical for emergency management. With the registration mechanism, the users of the mobile app represent a set of selected citizens whose identity is known to the TORCIA system. In case of recurring emergencies, the app can build a profile of mobile app users and design cooperation processes that assign specific roles to users depending on their profile characteristics. For example, it is possible to assess the frequency and impact of user contributions, with well-known user scoring mechanisms, and implement automated evaluations of user influence and dependability based on the scores (Cha, 2010).

Through the app, mobile users can access the same geolocalized information that is available to institutional operators through the management dashboard. In fact, institutional operators can be users of the mobile app (Capelli, 2013) and, as a consequence, the app can also be used as an on-site coordination tool.

7 CONCLUSIONS

The TORCIA workflow starts with the real-time collection of information from the main social media. This information is analyzed by a semantic engine that can identify and geolocalize potential or ongoing emergency situations based on the online buzz. When it identifies a potential emergency situation, the semantic engine sends an alert to a management dashboard that has been designed for institutional operators, such as the Civil Protection. If the alert is validated by an operator, an emergency alarm is created and emergency management procedures are activated. In particular, within the TORCIA platform, it is possible to communicate

with citizens with a multi-channel approach, by including social media among the set of active communication channels. Citizens can use the mobile app to obtain information on the emergency situation. For example, they can access all the information that the crawler collects from social media, provided that it is previously disambiguated, classified and geolocalized by the semantic engine. They can request the app to calculate escape routes based on their current position as well as operating constraints imposed by institutional operators (this software module has been developed by ACT solutions, a partner of the TORCIA project).

REFERENCES

- Capelli, P., 2013, "A Methodology for the Classification of the competitive models of mobile applications," Graduation Thesis in Computer Engineering (in Italian), Politecnico di Milano, Dec. 2013.
- Carcaci, C., 2012, "Syntactic text classification with set theory", Graduation Thesis in Computer Engineering (in Italian), Politecnico di Milano, Dec. 2012.
- Cesana, G. , D'Ettorre, A. "Analisi, Metodologia e Sviluppo di un Sistema Integrato per la Gestione delle Emergenze in Ambito Social", Graduation Thesis in Computer Engineering (in Italian), Politecnico di Milano, Sept. 2012.
- Rossetto R., 2012, "Twitter's numbers in Italy", Wired, 26, Sept. 2012, <http://daily.wired.it/news/internet/2012/09/26/numeri-twitter-italia.html>.
- M. Cameron, R. Power, B. Robinson, and J.Yin., 2012, "Emergency situation awareness from twitter for crisis managment. In SWDM Workshop, 305-308, Lyon, France, 2012.
- E. Estelles, F. Gonzalez, 2012, "Towards an integrated crowdsourcing definition," Journal of Information Science, 38(2), 189-200, 2012.
- M.Cha, H.Haddadi, F.Benevenuto, and K.P.Gummadi, 2012, "Measuring user influence in twitter: The million follower fallacy," In Proceedings of international AAAI Conference on Weblogs and Social, 2010.
- G. Shih, 2012. "In hurricane, Twitter proves a lifeline despite pranksters," Reuters, 2012.
- B. Brughemans, 2012, "Using Information from Social Media for Emergency Management," EENA Thematic Newsletters, <http://www.eena.org/uploads/gallery/files/doc/socialmedia.pdf>.
- A. Saravanou, D. Gunopoulos, G. Valkanas, G. Andrienko, 2015. "Twitter floods when it rains: a case study of the UK floods in early 2014," Workshop Series on Social Web for Disaster Management, Florence (IT), May 18-22, 2015.
- M. Imran, C. Castillo, 2015. "Towards a data-driven approach to identify crisis-related topics in social media streams," Workshop Series on Social Web for

- Disaster Management, Florence (IT), May 18-22, 2015.
- S. Cresci, M. Tesconi, A. Cimino, F. Dell'Orletta, 2015. "A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages," Workshop Series on Social Web for Disaster Management, Florence (IT), May 18-22, 2015.
- H. Purohit, P. Meier, C. Castillo, A. Sheth, 2013. "Tutorial: Crisis Mapping, Citizen Sensing and Social Media Analytics- Leveraging Citizen Roles for Crisis Response Coordination," *ICWSM 2013*, <http://www.knoesis.org/hemant/present/icwsm2013>.

