

Controlling and Monitoring Crisis

Nuria García-Santa¹, Esteban García-Cuesta²(✉), and Boris Villazón-Terrazas¹

¹ Expert System Iberia, Campo de las Naciones, Madrid, Spain

`ngarcia@expertsystem.com`, `esteban.garcia@uem.es`

² Universidad Europea de Madrid, Madrid, Spain

`bvillazon@expertsystem.com`

Abstract. Nowadays there is an increase interest on using social media contents for detecting and help during natural disasters. That interest is mainly based on the successful use of this type of data in the different phases of a disaster management process which ranges from early detection to efficient communication during the management of disasters. This paper focuses on the first phases of disasters' management and presents a system that allows to analyze, enrich, and detect on real time needs from a set of web sources. By using the concept of crowd as a sensors, this application can help, jointly with other traditional systems, to improve the current applied procedures during the initial phases of the disaster and detect demanded needs.

Keywords: Disaster management · Social media · Emergency assistance · Crowdsourcing

1 Introduction

The combination of recent advances of information technologies and communications, with the advent of social media applications has fueled a new landscape of emergency and disaster response systems, by allowing affected citizens to generate geo-referenced real time information on critical events. Moreover, information from social media is increasingly becoming an important source of knowledge not only for situation awareness but also for obtaining users experiences, feelings, and critical local knowledge through their post and comments about the event [4]. The identification and analysis of such events is not straightforward and the application of automatic tools is needed for both language processing and semantic interpretation.

In this scenario the concept of “crowd as a sensor” is very relevant due to it provides a larger coverage of the situation and its context. Furthermore the concept of ‘crowd as journalist’ which can also provide an initial interpretation of the event is an extra which is obtained by analyzing the comments that people provide. That terminology has its roots in the “crowdsourcing” term which can be defined as a process of obtaining needed services, ideas, or content by soliciting contributions from a large group of people, and especially from an online community, rather than from traditional employees or suppliers [4].

At present, there are several crowdsourcing open platforms available on the web, which are being used by governments, crisis response teams, NGOs, business organizations and other individuals to collect data and use it to develop new policies, innovative ideas for new products, help victims of natural calamities to find refuges, medicines, and other emergency needs [2]. Following that approach we present in this paper an ongoing system for helping during the initial phases of a crisis by relying on the contents and extraction of knowledge from social media channels such as Twitter¹ and Facebook², and using an open crowdsourcing platform so called Ushahidi³. Our motivation is to convert social media messages into actionable units of information (AUIs) which can be consumed directly by final users or by other machines in order to generate an automatic response to a particular crisis. The system is based on Linked Open Data [1] to alleviate the integration problems of crowdsourced data and to improve the exploitation of such new generated data by other crisis control systems.

2 Crowd Crisis Control

Crowd Crisis Control⁴ is an under construction platform that aims at controlling and monitoring large areas humanitarian crisis within low-cost technologies. The platform makes use of XLike multi-linguistic pipelines, developed within the context of the XLike EU FP7 project⁵, which, among other functionalities, provide basic linguistic state-of-the-art tokenization, lemmatization, part of speech tagging, and name entity recognition. Moreover, one of the main advantages of such infrastructure is that it provides a common framework for the four major languages worldwide: English, Spanish, Chinese, and German.

The system has been implemented and it is based on the platform Ushahidi⁶, a framework that provides visualization, reporting, and data collection capabilities, and allows the interaction with external volunteers. During the training of the system we have collected a database of crisis alerts by monitoring in English and Spanish some agency channels from Twitter and Facebook, such as *112Madrid*^{7,8}, Emergency Center of Madrid, Spain and *nswpolice*^{9,10}, Police of Sydney, Australia.

Next, we present a high level overview of the complete workflow of the platform. The implemented system periodically monitors information from social networks channels such as Twitter and Facebook, extracts relevant data (considering relevant those posts or comments which contains any word considered

¹ <https://twitter.com/>.

² <http://facebook.com/>.

³ <http://www.ushahidi.com/>.

⁴ <http://crowdcrisiscontrol.isoco.net/crowdmap/>.

⁵ <http://www.xlike.org/>.

⁶ <http://www.ushahidi.com/>.

⁷ <https://twitter.com/112cmadrid>.

⁸ <https://es-la.facebook.com/112ComunidadeMadrid>.

⁹ <https://twitter.com/nswpolice>.

¹⁰ <https://www.facebook.com/nswpoliceforce>.

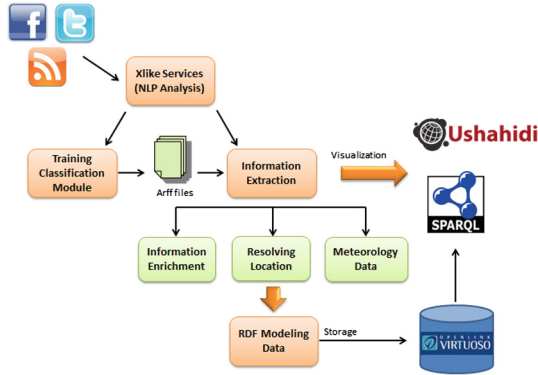


Fig. 1. Internal Workflow of CrowdCrisis Control

in the crisis vocabulary which has been previously defined automatically), analyzes the texts and automatically filter and enrich those that add value in the management of humanitarian crisis.

The Fig. 1 shows the system workflow which includes the following steps:

- **Identification and crawling of data sources:** it monitors social network channels through an API REST facilitated by these platforms. We have used the APIs provided by Twitter¹¹ and Facebook¹².
- **Analysis of text messages:** it processes texts obtained from the different sources by applying the natural language processing techniques which are provided by the multilingual XLike services [3].
- **Training at classification module:** it parses texts and uses machine learning techniques (e.g. naive bayes) to train and filter results for each one of the domains (fire, earthquake, flood, etc.).
- **Extraction of relevant information:** it includes the following tasks
 - modeling of the unstructured data to obtain information such as source, brief description, date, etc. and also include the predicted domain (type of emergency, fire, flood, etc.) of the texts which are obtained by the classifier module.
 - extraction of the location from mobile devices (whenever it is possible due to not all the users have this option activated); if that is not possible then the localization is extracted directly from the social networks platform whenever it is available.
 - extraction of cross-information by linking the analyzed data and the obtained entities with external sources of information. For example using named entities related to locations and link those to resources coming from Geonames¹³ and OpenWeatherMap¹⁴.

¹¹ <https://api.twitter.com/1.1/>.

¹² <https://graph.facebook.com/>.

¹³ <http://geonames.org>.

¹⁴ <http://openweathermap.org/>.

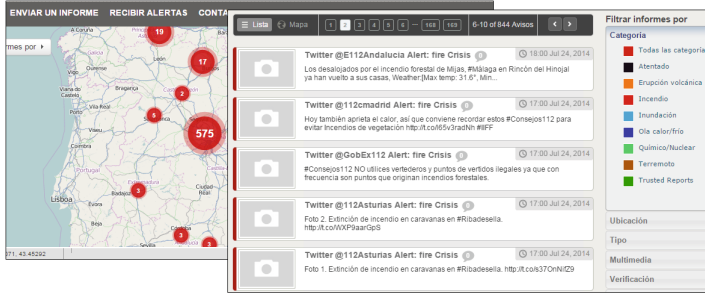


Fig. 2. Crowd Crisis Control Platform - map and report visualization of the incidents

- **Modeling and transformation of data:** it structures the analyzed data into RDF¹⁵ format conforming to standard vocabularies and ontologies, such as MOAC¹⁶ and WGS84¹⁷.
- **Publication of analyzed and structured data:** it stores the RDF data in a publicly accessible Virtuoso¹⁸ triplestore and also it is provided using Ushahidi platform¹⁹ to manage easily the crisis information reports.
- **Visualization of the data:** it presents crisis reports on Ushahidi platform which provides an end-user appealing visualization at the Crowd Crisis Control website²⁰. Ushahidi platform allows showing emergency reports, with textual and media resources, as temporal and spatial data. Besides it includes collaborative features, like the option of making comments or send reports by end users, it facilitates sending alerts about emergencies monitored to every registered user, and it supports PC and mobiles devices.

3 Evaluation

The presented system has been deployed within the context of C3 Spanish Project²¹ and it is monitoring crisis located in the countries Spain and Australia. Figure 2 shows the geolocalization of the incidents and the obtained reports resulting from the classification of detected incidents. Despite the interface has been implemented in Spanish, due to the C3 Spanish project main coverage and clients are in Spain, we want to point out that the systems can also analyze other languages such as English or any other which is compliant with the APIs provided by the XLike multilingual pipeline [3].

¹⁵ <http://www.w3.org/RDF>.

¹⁶ <http://www.observedchange.com/moac/ns/>.

¹⁷ http://www.w3.org/2003/01/geo/wgs84_pos.

¹⁸ <http://virtuoso.openlinksw.com/>.

¹⁹ <http://www.ushahidi.com/products/ushahidi-platform>.

²⁰ <http://crowdcrisiscontrol.isoco.net/crowdmap>.

²¹ <http://crowdcrisiscontrol.isoco.net/>.

As can be seen, the system automatically extracts those comments or posts which are associated to any of the categories that has been trained for and in this case the ones shown belongs to a fire crisis and contain different information such as prevention information to avoid forest fires (second one from the top) or some consequences due to a fire which has already happened such as evacuation of people from their houses. It worth to highlight that this type of information belongs also to different phases of a crisis management process which currently we are not able to automatically categorize.

4 Conclusions

In this paper we have presented an ongoing system for managing crisis which is based on the assumption that the information provided by the crowd through the social media channels can complement the traditional sources of information. For this purpose we have analyzed to major social media platforms such as Twitter and Facebook and implemented a new crowdsourcing platform, which is based on the open source platform Ushahidi.

The presented system include all the needed functionalites in order to perform a complete knowledge extraction process in a crisis scenario covering the functionalities of extraction and access to the source data, filtering of useful post/comments, natural language processes, structuring of the data, classification and finally representation and interpretation. We are able to process this workflow for Spanish and English though it can be easily extended to any other language which is covered by services of the XLike EU project. Despite the current system has many of the needed funcionalities already implemented, the automatically categorization of the texts regarding its membership to one of the phases during the management of a crisis is proposed as the next most important challenge to be achieved.

Acknowledgments. This work was supported by X-Like project (ICT-288342-STREP) and Spanish CrowdCrisisControl project (IPT-2012-0968-390000).

References

1. Bizer, C., Heath, T., Berners-Lee, T.: Linked data - the story so far. *Int. J. Semant. Web Inf. Syst. (IJSWIS)* **5**, 1–22 (2009)
2. Halder, B.: Evolution of crowdsourcing: potential data protection, privacy and security concerns under the new media age. In: *Democracia Digital e Governo Eletrônico*, Florianópolis, vol. 10, pp. 1–17 (2014)
3. Padró, L., Agic, Z., Carreras, X., Fortuna, B., Garcia-Cuesta, E., Li, Z.: Language processing infrastructure in the xlike project. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 14)* (2014)
4. Poblet, M., García-Cuesta, E., Casanovas, P.: Crowdsourcing tools for disaster management: a review of platforms and methods. In: Casanovas, P., Pagallo, U., Palmirani, M., Sartor, G. (eds.) *AICOL 2013. LNCS*, vol. 8929, pp. 261–274. Springer, Heidelberg (2014)