

Using Temporal Analytics for Early Discovery of Upcoming Threats

Joel Brynielsson*, Pontus Hörling*, Fredrik Johansson*, Michael Malm*, Christian Mårtenson*, Magnus Rosell†, and Staffan Truvé†

*Swedish Defence Research Agency (FOI), Sweden

†Recorded Future AB, Sweden

Abstract—An ocean of data is available on the web. From this ocean of data, information can in theory be extracted and used by analysts for detecting emergent trends. However, to do this manually is a daunting and nearly impossible task. We describe an implemented semi-automatic system in which data is automatically collected from selected sources, and to which linguistic analysis is applied to extract, e.g., entities and events. The extracted information is automatically transformed into OSINT reports which are clustered and visualized. The user can interact with the system in order to obtain a better awareness of historic as well as emergent trends.

I. INTRODUCTION AND BACKGROUND

During the last years there have been several crises around the world that have called for different kinds of international humanitarian operations. Obviously it is important to discover these kinds of crises as early as possible, allowing for better planning and preparedness before crisis management teams are deployed. We argue that early signs of new conflict areas and trends such as suicide bombings, piracy, drought, and religious and ethnic persecution in many cases are reported by non-governmental organizations (NGOs) working in this kind of areas long before the trends become common knowledge.

In recent years there has been a growing interest in automatic and semi-automatic harvesting and mining of data from web pages, social media, RSS feeds, etc. [1]. A system that can be used for web harvesting is Recorded Future, which offers advanced linguistic analysis capabilities such as entity and event extraction. In this work, we combine the capabilities of Recorded Future with the open source platform Ushahidi. Ushahidi has been used in several recent crises, such as in Japan. Ushahidi makes it possible for local observers to submit eyewitness reports using SMS or web forms and increases the situational awareness of what the crisis situation looks like by presenting the reported events in, e.g., Google Maps.

II. THE IMPLEMENTED TOOL

We have implemented a prototype system which combines the strengths of Recorded Future and Ushahidi. Recorded Future is used to harvest text from data sources such as RSS feeds and web sites using event extractors. Once the events have been extracted from the harvested data, the event type is together with location, date and time, etc., used to construct reports which are exported into Ushahidi. In Ushahidi, they are clustered based on their geographical position and shown on a map, as illustrated in Figure 1. The user of the tool can choose

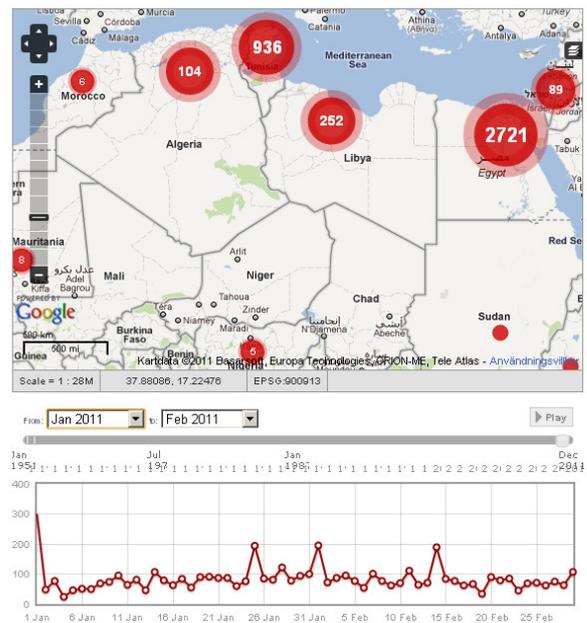


Figure 1. Clustered reports of event type Protest.

a start date and an end date and visualize how the number of reports with the selected event or category type evolves over time geographically. In this way the user can become aware of the development of trends in retrospect, and hopefully also as new trends emerge. The implemented prototype system has been used to collect data during the first quarter of 2011. Initial studies have shown that the tool can be useful for studying the spread of the Arabian Spring, but more systematic tests of the tool are still left for future work.

ACKNOWLEDGEMENTS

This work was supported by the R&D programme of the Swedish Armed Forces and by Recorded Future AB, and is also a supporting work for the Alert4All research project, which is funded under the European Union Seventh Framework Programme under contract no 261732.

REFERENCES

- [1] N. Memon, D. Hicks, and H. Larsen, “Harvesting terrorists information from web,” in *Proceedings of the 11th International Conference on Information Visualization*, 2007, pp. 664–671.