

A REAL-TIME SOCIAL MEDIA MONITORING SYSTEM AS AN OPEN SOURCE INTELLIGENCE (OSINT) PLATFORM FOR EARLY WARNING IN CRISIS SITUATIONS

Elena ȘUȘNEA

“Carol I” National Defence University, Bucharest, Romania
esusnea@yahoo.com

Abstract: *In the past years, the number of active users on social networking websites has increased exponentially. The popularity of these websites is due to the fact that users aren't the only content consumers (e.g., news, information, entertainment), they can also generate content and collaborate with other users. In case of the occurrence of unexpected events, such as natural disasters or human atrocities, active users can react in real time by uploading or sharing comments, photos, audio and video recordings, meaning that the amount of data generated and related to the event will be huge. These immediate and dynamic reactions of the users provide insight on the severity of the event and can help identify present or potential factors which could cause the crisis situation to destabilize. The first goal of intelligence is to obtain information on the situation's awareness. In this context, a real-time social media monitoring system can provide a snapshot of the happenings around the unexpected event, therefore it can be an open source intelligence (OSINT) platform for early warning in crisis situations. Furthermore, using intelligent technology, such as data mining, the system can support decision makers to anticipate the evolution of the crisis and take preventive actions to avoid escalating the crisis.*

Keywords: OSINT, intelligent technologies, social media, early warning, crisis situation

1. Introduction

Multiple changes across the globe in recent decades have led to substantial changes in the intelligence paradigm. Globalization and the information revolution have increased the complexity of the current security environment. Also, political disputes, armed conflicts, technological incidents or natural disasters have contributed to the emergence of political, military or humanitarian crises. In this context, the use of the latest technologies for the collection and real-time analysis of open source information for the creation of open source intelligence (OSINT) products

and the dissemination of these products to decision-makers involved in crisis management can prevent the escalation of the situation and even avoid unwanted effects. Decision support templates successfully applied in predictable conventional situations, based on doctrines and regulations, do not work as well in case of unconventional, unpredictable situations. The complexity of current crises accentuates the need to develop early-warning computer tools in order to support decision makers in making proactive decisions.

In this article, we are presenting a social media monitoring tool that will allow the collection and processing of information from open sources in order to obtain OSINT. The combination of state-of-the-art intelligent technologies, such as Web crawling, data mining, natural language processing, with the intelligence analysts' experience, is the key element for supporting decision makers to anticipate the evolution of the crisis and take preventive actions to avoid escalation of the crisis.

2. Crisis and technological revolution in the context of current security environment

Despite the frequent use of the concept of crisis, there is currently no unanimously accepted definition. The difficulty of formulating a unitary definition of the crisis stems both from its complexity and from the multitude of approaches of different entities (political decision-makers, military or non-military organizations, civil society, etc.) involved in its settlement. Also, the concept has different meanings that have varied according to certain factors, such as the context sensitivity, the perspective from which the situation is studied, etc. In essence, a crisis may be of a political, military or humanitarian nature and may be caused by political disputes or armed conflicts, technological incidents or natural disasters.

As there is no internationally accepted definition, based on the previously formulated definitions [1], [2], [3], we would like to propose the following definition: *the crisis is a complex situation perceived as a threat to the values, interests or goals of the entities involved, which often requires near-real-time interventions to reduce the disruptive interactions between them.*

The definition contains some keywords whose significance is detailed below. By *situation*, we understand any dispute, problem or condition in which an entity

(e.g. a country, nation or other non-state entity) finds itself at some point, economically, politically or socially and in which decision makers may intervene through amending, transforming or improving. The situation is *complex* if it is difficult to define or if it causes unpredictable changes or transformations that lead to the deterioration of the existing order, it significantly influences the reaction to some solutions, has many interdependent causal forces, is new, and often requires *near-real-time interventions*.

In the recent years, technological revolution has contributed, on the one hand, to the increase in complexity of the space in which current crises occur and, on the other hand, it continues to exert a significant influence on the development of specific crisis management activities.

First of all, a factor that increases the complexity of the space in which current crises occur is due to the widespread use of Web 2.0 specific technologies by persons and organizations, especially social media platforms, for misinformation and radicalization purposes.

Secondly, the IT revolution is manifested in the crisis management process. The transition from peace to mistrust and then to confrontation or armed conflict can be detected by carefully defining and monitoring certain indicators using appropriate IT tools. These indicators should shape the key characteristics of the crisis (entities and relationships between entities) in order to estimate the risk of emergence of triggering factors that would degenerate the situation.

Analyst Sherman Kent estimated that in peacetime 80% of the information needed by decision-makers is public [4]. So, from this perspective, OSINT is considered the first resource and, in this sense, it is a precursor and catalyst for INTs.

3. Towards Open Sources

The rapid development of Web 2.0. in the

past ten years has turned virtual space into the largest data open source worldwide. The open sources in the Web space are both opportunities and challenges for the OSINT production.

Opportunities derive from the ability to access public data generated by various entities on the Internet, especially data from social media sites (such as YouTube, Facebook, Twitter) that can provide feedback on the occurrence of a trigger factor.

The challenges of using open sources consist, in particular, of the big data characteristics that these data have: volume, variety, velocity, and veracity [5]. The volume characterizes the magnitude of the data. The open source data dimensions are measured in terabytes and petabytes. Variety refers to the heterogeneous structure of the data set. Advanced technologies allow for the collection of various types of data structures, so we have structured data, semi-structured data and unstructured data. Currently, structured data is only 5% of the total data [6]. Velocity refers to the speed at which open source data is generated and analyzed. The proliferation of digital devices has led to an unprecedented rate of data creation and has spurred the development of real-time analysis techniques. Veracity emphasizes the inherent instability of open sources. For example, the feelings of social networking users are uncertain. So working with imprecise and uncertain data is another aspect of the big data.

Therefore, data collected from open sources require innovative forms of processing that provide an improved insight into the crisis and support timely decision making.

4. A Real-Time Social Media Monitoring System as an OSINT Platform for Early Warning in Crisis Situations

We must emphasize that the field of early warning must be approached with dainty expectations, because early warning does

not mean prophecy. Early warning means the provision of timely information, that allows decision makers to analyse the data in detail and set up intervention measures if necessary to avoid or reduce its unwanted consequences and prepare for effective response.

This timely information is provided by early warning systems. These systems can collect, store, analyze, and disseminate information with the purpose of “understanding and mapping the hazard; monitoring and forecasting impending events; processing and disseminating understandable warnings to political authorities and the population, and undertaking appropriate and timely actions in response to the warnings” [7].

The intrinsic value of early warning systems lies in their ability to help prevent or mitigate the undesirable effects of the crisis.

The platform for collection and processing of open source data has as a general objective the provision to the decision-makers of OSINT products in real time. In this respect, the platform must support all five stages of the OSINT process.

We have come up with an architecture for the Real-Time Social Media Monitoring System, designed for early warning, consisting of four modules. The architecture allows OSINT to be obtained in crisis situations and is based on the following four modules:

- the *data collection module* allows for the automatic or semi-automatic collection of data from open sources (mass media, grey literature, data and information commercial sources, online communities and user-generated digital content) in order to achieve digital format data collections.
- the *data storage module* has the role of clearing the data collected from multiple open sources and ensuring their efficient access by storing it in a data warehouse.

The statistics and analytical reports are generated in the warehouse.

- the *data analysis module and intelligence product achievement* performs the transformation of data into information and of the information into intelligence. It allows for the establishment of appropriate models and parameters, and the application of data mining methods and techniques (e.g. artificial neural networks, regression, decision trees, association rules, etc.) in order to obtain OSINT.
- the *OSINT reporting and dissemination module* supports the interaction between decision makers by providing a collaborative virtual space.

The data mining process for obtaining the OSINT from the web is the same as the process traditionally used on structured data. The difference is in the development of new methods that allow for the analysis of data from open sources available in the Web space through the features of the big data. This difference is underlined in the Web Data Mining concept.

Decision makers need a high level of interaction and communication, and ways to help generate ideas. The benefits of using the latest data collection, storage, analysis and visualization technologies in developing IT systems to assist decision-making are obvious. They provide more time for the preparation, analysis and planning of a coherent response and, in case

of intervention, increase the likelihood of success.

Decision-support IT systems contribute not only to the formulation and setting of the objectives to be achieved, the development of courses of action and their comparison, which eventually leads to the implementation of the chosen options but also to the prediction of the potential reactions of the parties involved in the crisis management. Thus, in the crisis management and conflict prevention procedures, the importance of information systems increases especially in the early stages of a crisis, based on the acquisition, evolution and analysis of relevant information for decision-makers.

6. Conclusions

Current crises - whether military, political or humanitarian - can no longer be addressed only in terms of military capabilities; nor are information requirements still only limited to the military aspects of the mission and the battle space.

Nowadays, information from open sources is critical for the decision-makers. The data generation process happens much faster than our ability to access them, so we need to constantly adapt our methods and techniques in order to collect, process, and produce information in good time so that the dissemination of intelligence to the decision-makers takes place in a timely manner to mitigate the effects of the crisis.

References

- [1] Glenn Herald Snyder, Paul Diesing, *Conflict Among Nations: Bargaining, Decision Making and System Structure in International Crisis*, Princeton University Press, Princeton, 1977.
- [2] Michael Brecher, Wilkenfeld Jonathan, *Crisis, Conflict and Instability*, Pergamon Press, New York, 1989.
- [3] Jean-Louis Dufour, *Les crises internationales de Pekin (1900) au Kosovo (1999)*, Corint, București, 2002.
- [4] Allen Dulles, *Memorandum Respecting Section 202*, Central Intelligence Agency, April 25, 1947.
- [5] <https://www-01.ibm.com/software/sg/data/bigdata>

- [6] Cukier Kenneth, *Data, data everywhere*, The Economist, 2010.
- [7] <https://www.unisdr.org>