Galean: Visualization of Geolocated News Events from Social Media

Vanessa Peña-Araya, Mauricio Quezada, and Barbara Poblete
PRISMA Research Group
Department of Computer Science
University of Chile, Santiago, Chile
{vpena, mquezada, bpoblete}@dcc.uchile.cl

ABSTRACT
Online Social Networks (OSN) have changed the way information is produced and consumed. Organizing and retrieving unstructured data extracted from these platforms is not an easy task. Galean is a visual and interactive tool that aims to help journalists and historians, among others, analyze news events discussed on Twitter. In this tool, news events are visually represented by the very countries from where the news originated, the date they happened and their impact in the OSN. Galean considers countries as entities, as opposed to mere geographical locations as most of the tools in the state of the art. As a consequence, it allows users to explore and retrieve news not only by their geographical and temporal context, but also by the relationship among countries. With this tool users can search for behavioral patterns of news events and observe how countries are associated in specific events. We expect our work to become a public tool that helps conduct historical analyses of social media news coverage over time.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—Information filtering; H.5.2 [Information Interfaces and Representation]: User Interfaces—Graphical user interfaces

Keywords
Social media analysis; spatio-temporal analysis; geovisualization; text analysis; historical analysis

1. MOTIVATION AND CONTRIBUTION
An event is usually understood as something that happened in a particular place and time. When a news event happens, users from OSN react to them by sharing the information or opinions they have about them. However, it is not an easy task to extract valuable knowledge from this huge volume of information. Even if news events are detected and stored in local databases it is hard to retrieve them later. This is mainly because the user does not always know what he or she is looking for and, if so, might not know the exact words to describe it.

In this paper we present Galean, an interactive visualization tool that represents events extracted from Twitter by their geographical and temporal context. It aims to help users explore and retrieve news events from OSN by the political geographical entities that participate in it. In addition, it aims to reveal links between countries based on their participation in particular news events, and behavioral patterns of the news. With this representation of news events, we work to deliver an interactive tool that can answer questions such as what are the events related to this country on this date? Or, in which news events do these two countries participate together and how do they evolve?

Most state of the art tools that work with geographical features of events extracted from OSN, aim to position them over where they happened or aim to detect events happening in a particular place. For example, Sakaki et al.\[3\] use the geographical metadata in tweets\[1\] to detect where the epicentre of an earthquake is and Watanabe et al.\[5\] to identify small localized events. SensePlace2\[1\] gives a broader view by also considering the locations that tweets are about. However, this tool emphasizes dimensions of the events more than the events themselves. As a result, these tools ignore the relationships among countries that occur as a consequence of real world events, and therefore do not allow users to explore and query events by these relationships as our tool does.

We hope Galean will help journalists and ordinary OSN users explore and analyze news events from a historical point of view.

2. SYSTEM ARCHITECTURE
The system architecture of Galean is designed in a three-step process:

I.- Event identification: We detect news events that had high media coverage by gathering messages from verified news accounts from Twitter. The data set currently in use contains news events starting from August 2013. The detection and storage processes of news events are conducted every hour. Although Galean could use any news source, it contains news accounts from Twitter. The data set currently in use

II.- Ge-temporal contextualization: besides the date in which

1 Tweets: messages posted in the Twitter OSN
a news event was detected, we also extract the countries from where the news originated, and the countries from where Twitter users talk about the news. Both processes use CLAVIN [4], an open source geoparser. A more detailed description of the location-aware model of news events used by Galean can be found in [2].

III.- Visualization: The complete view of Galean is in Figure 1. To start using it, the user must select a particular date of interest in the filter area (section A, Figure 1). After this, all the news events detected on that date will be displayed as a bubble on the main map, as depicted in section B of Figure 1. Each of them are positioned over the country that was involved in the event with their size represent their impact in OSN, measured by volume of messages. When passing the cursor over international events (orange bubbles), the links between all the countries involved in the news event are revealed. By clicking a bubble, additional information about the news it represents appears in the view: miscellaneous data, propagation of the news over the world and tweets discussing the event, in sections C, D and E respectively in Figure 1.

3. DEMONSTRATION

Galean is an exploratory tool. By selecting a date to start with, the user has an idea of the news events that happened that day by country and impact. With this basic information the user navigates and finds news events according to what he or she is looking for. Figure 1 shows the overview of news events for April 5th, 2014. The selected event in that picture refers to the search of the Malaysian flight MH370 which disappeared March 4th of the same year. As can be observed, that event relates Malaysia, China and Australia. This is because a Chinese news agency declared at that time that a signal, that may have come from the black box of the missing flight, was received from a location near Australia.

It is possible to see that the event has a lot of impact on Twitter, as shown in the map below. This map also highlighted in red that Malaysia, China and Australia were the protagonists of that event. The filters available in the left part of the view allow users to focus on news associated with one or more countries. With these features, Galean allows users to follow the evolution of a news event.

4. ACKNOWLEDGEMENTS

The authors were partially supported by a FONDECYT Grant 11121511 and the Millennium Nucleus Center for Semantic Web Research Grant NC120004. We also thank Alexandre Bergel and Omar Miranda for their help in early stages of this work.

References


