2016 International Conference on Computational Social Science
June 23-26, 2016 IC^2S^2

Monitoring Electoral Violence through Social Media: A Machine Learning Approach

Xiao Yang, Iadh Ounis and Craig Macdonald
School of Computing Science, University of Glasgow, UK
{firstname.lastname}@glasgow.ac.uk

Keywords: electoral violence, social media, twitter sampling, agreement study, machine learning

In recent years, there has been a growing research interest in social media platforms such as Facebook and Twitter, which have been shown to be influential in reporting news and events around the world [1]. Twitter is a communication platform allowing users to report and comment on a wide range of topics, including political events. Indeed, both citizens and politicians are increasingly embracing social media to disseminate information and comment on various topics particularly during significant political events and campaigns [2]. The conduct of regular, peaceful, and democratic elections is a central aim to a just and fair society. Unfortunately, electoral malpractice and violence continue to persist across a number of countries, preventing the consolidation of the democratic process. In this paper, we develop upon existing information retrieval and text classification approaches to develop effective classifiers that can detect electoral violence incidents from the users' posts and discussion in Twitter. In particular, we create a large test collection by collecting and monitoring tweets pertaining to the Venezuela parliamentary election in 2015. Using state-of-the-art machine learning approaches, we develop a new text classifier tailored to the detection of electoral malpractice and violence, and demonstrate its effectiveness on the aforementioned 2015 Venezuela election test collection.

Dataset and overview. We monitored Twitter for posts about the 2015 Venezuela parliamentary election, which was held on 6th December 2015. A number of election-related keywords and hashtags were used to collect a live tweet stream during the parliamentary election event, using the Twitter Streaming API. The crawled data contains over 7 million of tweets collected from 2nd November 2015 to 15th January 2016, which covers the pre-election period, the election day and a period following the election. The number of collected tweets per day is illustrated in Fig 1. As can be seen from the figure, a significant number of tweets were collected around two important dates, which are the election day (6th December 2015) and the day when the new National Assembly controlled by the opposition parties has been sworn in (5th January 2016). The Venezuela parliamentary election constitutes a good use case for demonstrating the effectiveness of our proposed text classification approach for detecting violence-related incidents from Twitter. Indeed, according to news media, a number of violence-related events have occurred in Venezuela during the election process such as the following: (1) Opposition politician Lilian Tintori was attacked in Cojedes [3]; (2) Opposition politician Miguel Pizarro was attacked in Petare [4]; and (3) Opposition politician Luis Manuel Diaz was killed [5].

Our approach to develop, train and evaluate our new tailored text classifier involved two stages. In the first stage (the micro-level analysis), we use the Terrier information retrieval (IR) platform [6] and the well-known pooling method [7] to develop a test collection from the collected data. Using the pooling method, a large sample of tweets has been sampled from the collection for labelling (aka coding) by social science experts in electoral violence and malpractice. The test collection is further analysed through an agreement study, whereby the level of agreement among the expert judges has been examined. Using a 5-fold cross validation approach, the labelled tweets are then used to train and evaluate our newly developed classifier tailored to separating violence-related tweets from the whole dataset. In the second stage (the macro-level analysis), we automatically group the violence-related tweets into clusters, where each cluster describes a specific violence-related event with its corresponding metadata (e.g. date, involved actors, etc.). Therefore, at the end of the process, we obtain several groups of tweets, each describing an electoral incident with its corresponding metadata description. In summary, by investigating the whole dataset at both the micro and macro levels, we are able to automatically identify all violence-related incidents in the 2015 Venezuela parliamentary election, and to assess the overall accuracy and effectiveness of our approach. The two stages of our approach are further described in detail below.

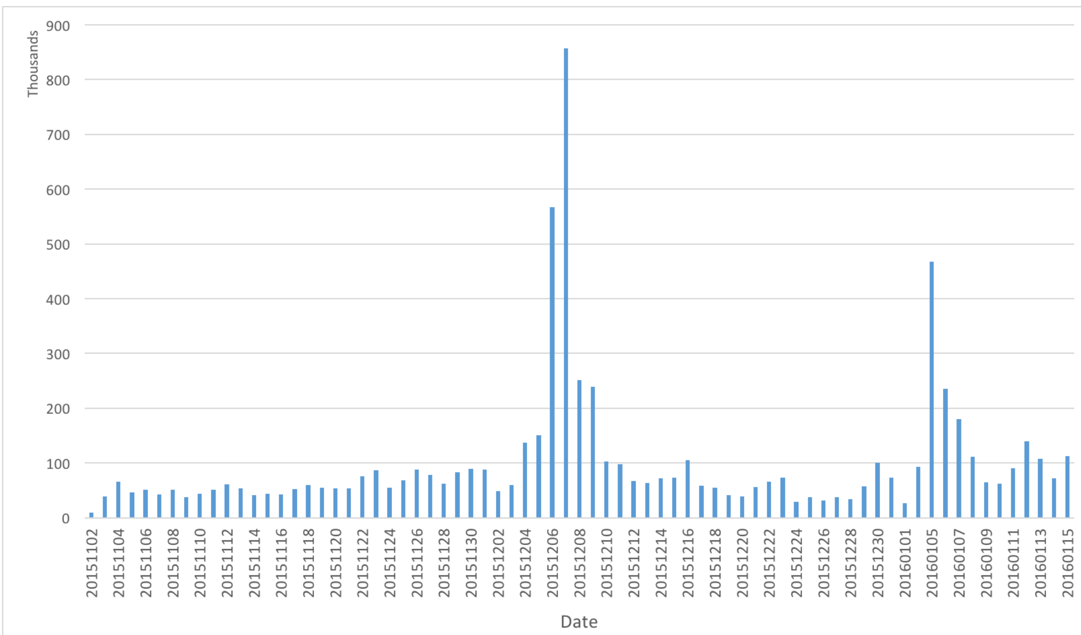

Figure 1. Number of tweets collected per-day in the dataset

The micro-level analysis: We use the state-of-the-art Terrier information retrieval platform to retrieve tweets that are relevant to a carefully selected list of election-related queries. For each query term, only the top *k* tweets are retrieved to build a pool for each day. After the entire dataset is queried day by day, all the query pools are aggregated to remove duplicates such as tweets retrieved by multiple queries or retweets which have the same content. By applying the pooling sampling method and eliminating duplicates, the test collection mostly contains those tweets of interest to the election, making it possible for human judges to examine and label the tweets related to electoral violence, while at the same time helping us to understand how Twitter is used when electoral violence happens. The latter provides useful clues on which features to use to help the classifier automatically detect the tweets of interest. Several electoral violence events are observed from our test collection, which confirms that our dataset is indeed representative and a useful case study for studying electoral violence using Twitter. In particular, each tweet in our generated test collection is categorised by our human assessors into three groups such as “election-related”, “electoral violence-related” or “unsure”. In order to study whether the labelled test collection is credible, a large sample of tweets has been carefully selected and judged by 5 experts. We study the level of agreement among the experts by comparing their judgements using classical statistical correlation methods. Such an agreement study allows us to learn more about the difficulty of the task, as well as to identify further features to integrate into the machine-learned text classifier. Indeed, both the discussion of the labelling task with the expert judges as well as the agreement study have been essential in identifying the most important features that distinguish electoral violence-related tweets from the rest of the dataset. These features have been used to train and develop an SVM classifier that can effectively identify the key violence-related tweets from the whole dataset.

T**he macro-level analysis:** After evaluating and refining the first stage of our approach, we automatically aggregate the relevant tweets by clustering them into different violence-related incidents using a number of clustering algorithms such as K-means. Topic modelling approaches are used to identify representative terms for each cluster. Furthermore, each cluster is inspected by our experts for both accuracy and additional metadata. Importantly, metadata such as the date of the incident, its type, the actors and victims involved in the incident are recorded from the observed tweets in each of the generated clusters (i.e. incidents). By studying the dataset from such a higher level, it is possible to summarise the electoral violence incidents that happened during the election period, including the provision of links to news articles and photos (linked to from the tweets), which help social scientists assess the credibility of the events. The proposed approach could be deployed on other elections and countries to evaluate its generalisation. Indeed, if successfully generalisable, the generated electoral violence datasets with their rich information at both the micro and macro levels are valuable resources for political scientists to study and mitigate electoral violence.

Acknowledgement. The authors acknowledge financial support from Economic and Social Research Council (ESRC), grant No ES/L016435/1.

References:

[1] Petrovic, S., Osborne, M., McCreadie, R., Macdonald, C., Ounis, I. and Shrimpton, L.. 2013. Can Twitter Replace Newswire for Breaking News? In *Proc. of International AAAI Conference on Weblogs and Social Media*.
[2] Fang, A., Ounis, I., Habel, P. and Macdonald, C. 2015. Topic-centric Classification of Twitter User's Political Orientation. In *Proc. of ACM SIGIR Conference*.
[3] EL NACIONAL. 2015. *Lilian Tintori reveló como fueron las agresiones hechas por colectivos en Cojedes.* Available: http://www.el-nacional.com/politica/Lilian-Tintori-agresiones-colectivos-Cojedes_0_740926097.html. Last accessed 20th Jan 2016.
[4] EL UNIVERSAL. 2015. *Miguel Pizarro reports on attacks by gov't supporters during caravan.* Available: http://www.eluniversal.com/nacional-y-politica/151123/miguel-pizarro-reports-on-attacks-by-govt-supporters-during-caravan. Last accessed 20th Jan 2016.
[5] BBC. 2015. *Venezuela opposition politician Luis Manuel Diaz killed.* Available: http://www.bbc.co.uk/news/world-latin-america-34929332. Last accessed 20th Jan 2016.
[6] Macdonald, C., McCreadie, R., Santos, R.L. and Ounis, I.. 2012. From puppy to maturity: Experiences in developing Terrier. In *Proc. of OSIR workshop at SIGIR*.
[7] Sanderson, M.. 2010. *Test collection based evaluation of information retrieval systems*. Foundations and Trends in Information Retrieval, 4(4). Now Publishers Inc.