Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

Tracking Political Elections on Social Media: Applications and Experience

Danish Contractor
IBM Research
New Delhi
dcontrac@in.ibm.com

Bhupesh Chawda
IBM Research
New Delhi
bhchawda@in.ibm.com

Sameep Mehta
IBM Research
New Delhi
sameepmehta@in.ibm.com

L. Venkata Subramaniam
IBM Research
New Delhi
lvsubram@in.ibm.com

Tanveer A. Faruquie
IBM T.J. Watson Research Center
Yorktown Heights, New York
tafaruqu@us.ibm.com

Abstract

In recent times, social media has become a popular medium for many election campaigns. It not only allows candidates to reach out to a large section of the electorate, it is also a potent medium for people to express their opinion on the proposed policies and promises of candidates. Analyzing social media data is challenging as the text can be noisy, sparse and even multilingual. In addition, the information may not be completely trustworthy, particularly in the presence of propaganda, promotions and rumors. In this paper we describe our work for analyzing election campaigns using social media data. Using data from the 2012 US presidential elections and the 2013 Philippines General elections, we provide detailed experiments on our methods that use granger causality to identify topics that were most "causal" for public opinion and which in turn, give an interpretable insight into "elections topics" that were most important. Our system was deployed by the largest media organization in the Philippines during the 2013 General elections and using our work, the media house able to identify and report news stories much faster than competitors and reported higher TRP ratings during the election.

1 Introduction

Social Media often complements or supercedes existing sources of information because of the variety of interactions, speed of availability and the collective expression of people. This is particularly true for applications which require analyzing information from a large number of people. Social media is thus becoming an increasingly popular medium for many election campaigns. On one hand it allows a candidate to mass communicate and reach out to a large section of the electorate, on the other hand it provides a medium for people to express their opinion on the proposed policies and promises of candidates. The resulting interaction that builds up to the day of the election provides a goldmine of information which can be used to analyze the political campaign, gauge public sentiment and establish causality between policies and peoples perception. Indeed, the successful application of online campaigns such as the 2008 Presidential Elections in the US, 2012 London Mayoral elections and the recent 2014 Indian Prime Ministerial Elections have shown the importance of online political opinions. However, analyzing social media comes with its own set of challenges. Much of the data is encoded in text, which is noisy, sparse and unstructured and sometimes even multilingual. Further, the information itself may not be completely trustworthy, particularly in the presence of propaganda, promotions and rumours. Moreover, the views of the people often swing because of external events.

Given the national prominence of elections and the influence of social media, the area has attracted the attention of several researchers. Their work ranges from using simple statistics to sophisticated linguistic analysis to predict the outcomes of elections[Gayo-Avello, 2012; Jungherr *et al.*, 2012; O'Connor *et al.*, 2010; Chung and Mustafaraj, 2011; William and Gulati, 2008; Sang and Bos, 2012]. Social media has also been used to better design election campaigns[1], reach out to supporters, raise money and analyze social expression as a feedback mechanism to improve the election campaigns. Work such as [Osborne and Dredze, 2014; Petrovic *et al.*, 2013] study whether information available on Twitter and other social media is indeed different from information available from other news sources.

In this paper we describe our work for analyzing election campaigns using social media data. Using data collected from the 2012 US Presidential Election and 2013 Filipino Election we describe our methods and experiments for detecting topics and sentiments that are important in an election. Our system identifies short term trending topics (emerging topics) as well as long-term topics, topics that are significant in a larger

[1] For eg: http://www.salesforce.com/uk/socialsuccess/social-media/london-mayoral-election-2012.jsp

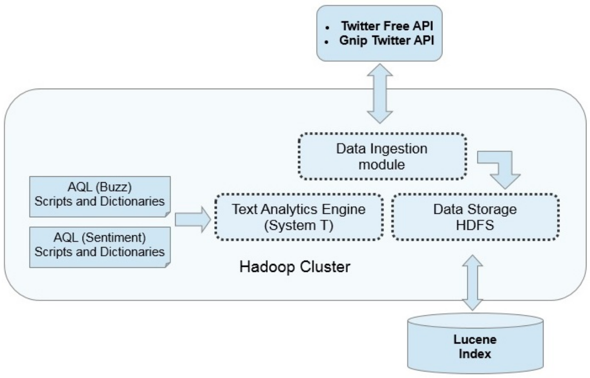

Figure 1: System Architecture

course of time. Our work was deployed live by ABS-CBN Corporation[2], which is the largest media house in Phillipines, during its coverage of the 2013 Filipino general elections. Reporters using our system were able to identify breaking news stories and run them much ahead of their competitors (Details in section 3).

In summary, our paper makes the following contributions :

- To the best of our knowledge, we are the first to actively deploy a system live in an election and demonstrate how social media data can be used to detect potential news stories that can be used by other main stream media.
- We describe our language independent methods for identifying topics that are most influential in elections and present experiments using two election data sets.
- Lastly, since there exists a large body of literature that uses social media data for prediction, we present our experiments on predicting election candidate approval ratings using causal topics identified by our system. While we do not claim our system can be used as a replacement of offline opinion polls and electoral surveys, our system can possibly be used as one more source to gauge the mood of the crowd during an election campaign.

The rest of our paper is organized as follows - the next section describes details of our system. In Section 3 we describe some experiments and results, in section 4 we review some related work and in section 5 we conclude the paper.

2 System

This section describes the implementation details and the methods employed for topic trend identification and candidate approval rating prediction.

2.1 Architecture

Our system was designed to analyze large amounts data from the social media[3]. We used a multi-node Hadoop cluster for building our system. Hadoop supports the processing of large amounts of data in a distributed fashion on a cluster of commodity machines. The ingestion components, composed of Java modules, were run on the master-server which connected to the Twitter APIs and fetched data directly to the HDFS storage on Hadoop. Data was stored in files of size approximately equal to the HDFS block size.

The data was analyzed using rules created in the SystemT [Krishnamurthy *et al.*, 2009] implementation. SystemT provides a rule based language called AQL (Annotation Query Language) for text analytics which is capable of executing in a Hadoop based environment. The rule language is convenient for identifying patterns from text, running the text through dictionaries, combining rules, filtering them etc.

In order to detect sentiment, we made use of a rule based sentiment analyzer built using SystemT. The sentiment rules defined in AQL make use of dictionaries containing negative and positive polarity words and the neighboring context captured in terms of part of speech tags. The sentiment module was coupled with another module called *Buzz*, which tagged any similar occurrences of top bigrams that occurred in the text. This allowed us to arrive at a sentiment polarity for a particular term (or bi-gram) as opposed to the sentiment for a sentence in general. In effect, the same tweet could be positive for some bigram while being negative for another bigram. The top-bigrams were updated periodically based on the volume of incoming tweets. The sentiment labeling module was found to have an F1 accuracy of 0.70 on tagged sentiment data set of 1000 english pre-labeled english language tweets. Figure 1 is a schematic representation of the architecture described.

Tagalog language: In order to customize the sentiment analyzer for the Tagalog language, language experts were requested to translate the rulesets used by the sentiment analyzer.

2.2 Topic trends in tweets

This section describes our modules for topic detection. We discover two types of topics - "*short term or emerging*" topics and "*long-term*" topics. As the names suggests, short-term topics are topics that are likely to appear and remain for short durations while long-term topics are broader themes or topics that remain active in social media discussions for larger periods of time.

Short-term (emerging) Topic detection

Emerging topics are the topics or keywords which appear in the "recent buzz" in social media. These topics are not very well known and are often short lived. They occur in bursts and disappear very quickly depending upon their significance or the entity they are linked with.

Since we were working with a news broadcaster, one of the key requirements was to detect "breaking" news off Social Media which can be delivered (after further development) on main stream news media. We used the Apriori algorithm [Agrawal *et al.*, 1993] for finding frequent item sets on tweets to find potential new topics. The requirement here was to not only identify the topics, but also link them to possible election entities (candidates and political parties). Due to the very nature of the data set we were dealing with, the names of candidates, political parties and common election keywords were bound to be frequent. Hence, an item set like ["elec-

[2]http://www.abs-cbnnews.com/

[3]Twitter, for example reports nearly 10,000 tweets per second.

tion","philippines"] is not interesting enough. However, in a set with more than one element, if we have a known entity relating one or more unrelated entities, then it becomes interesting enough and a subject for further investigation or reporting. Hence we filtered the sets which had all known entities, so as to link a known entity to an unknown entity. A very fitting example of such a set is ["nancybinay","viceganda"] which identified the event involving comedian Vice Ganda making a comment on political experience of Nancy Binay. Another such example is the set ["pcos, "malfunction"] which related to the alleged malfunction of PCOS machines. In fact "PCOS" became the Twitter trending topic[4] for most of the morning of the day this event occurred.

The discovered topics along with set of tweets were presented to reporters for manual verification, validation and further story development. While this module is fairly simple, it was found to be very effective in determining emerging topics in near real time. We now describe our work detecting long-term topics - topics that were important for a larger duration of time.

Long-term Topic detection

We define long-term topics to be those topics that are active in social discussions for a longer duration (a week or longer). In order to identify these topics, we use mentions of the most frequent bigrams, along with candidate names to build a regression model using the approval ratings of candidates (available from external sources - details in section 3). Using a regression model we then identify the most *causal* topics for the prediction made by the regression [Contractor and Faruquie, 2013].

Algorithm 1 Build prediction model using n-grams occurring in the data

Let A_d be the approval rating for a candidate on date d and m be the set of most frequent bigrams in the data
for each date d in the collection **do**
  for each of the political candidates **do**
    Let the set of users who mention candidate c, be U_c
    for each user $u_c \in U_c$ **do**
      for each ngram $ng \in N$ **do**
        Value of predictor feature f_c for candidate = number of occurrences of bigram ng in posts made by user u_c on date d
      end for
    end for
  end for
end for
Build linear regression model using the set of f_c's as predictors and approval ratings A_d as predictions.

Linear Regression

Let the dependent variable in the linear regression be y and let there be N independent variables x. Let the total number of data points be P, then the relationship can be expressed as

$$y_p = \beta_0 + \sum_{i=1}^{N} \beta_i(x_{ip}) + \epsilon_p \tag{1}$$

where ($p = 1, 2 \ldots P$) and ϵ_p is the error term for the p-th observation. The residual term e_p defined as

$$e_p = y_p - \hat{y_p} \tag{2}$$

gives the error between the prediction and the actual value of the dependent variable. Using the ordinary least squares method that minimizes the sum of the squared residuals, the parameters set β can be estimated. For our work the dependent variable y are the candidate approval ratings (A), while the independent variables are frequent bigrams.

Causality between topics (bigrams) and ratings

A commonly used technique to infer causal effects using time series observations is Granger causality [Granger, 1969]. A time series x is said to be be granger causal for another time series y if building a regression model combining the two series gives better predictions than a model built using only the time series y.

$$y_t \approx A.\vec{y_{t-1}} + B.\vec{x_{t-1}} \tag{3}$$

$$y_t \approx A.\vec{y_{t-1}} \tag{4}$$

If on applying a t-test or an F-test on the predicted outcomes from the two regression models above, shows a significant improvement in prediction when both time series, then the causality x is said to granger cause y.

Since we use multiple features in our models, performing a pair-wise granger test using each feature would be computationally expensive and therefore we use the Lasso Granger method for causality determination.

Lasso Granger method

The Lasso algorithm for linear regression performs variable selection using the L_1 penalty term to obtain a sparse estimate of the coefficient vectors β. The variable selection can be obtained by solving the following optimization problem:

$$min_\beta \sum_{p}^{P} || \ y - \sum_{i=1}^{N} \beta_i x_{ip} \ ||^2 + \lambda \ || \ \beta \ || \tag{5}$$

where λ is the penalty parameter that determines the sparseness of β. The series x is said to cause y iff β is a non-zero vector. The next section describes the data used and our experiments and results. We made use of Lasso Granger regression implementation provided by Matlab[5] for our experiments.

3 Experiments and Results

This section describes our experiments and results of our trial deployment during the 2012 US Presidential Elections and a live deployment used by the ABS-CBN News Corporation during the 2013 Filipino General Elections.

[4]http://www.abs-cbnnews.com/insights/05/13/13/citizens-use-social-media-complain-vs-defective-pcos

[5]http://www.mathworks.in/help/stats/lasso.html

3.1 Data set - US Elections

We used the US elections data set for development purposes and estimating how well our system would do when deployed live in a real election.

US poll data

We used the presidential candidate opinion poll released by Gallup[6]. Gallup conducts opinion polls concerning political, social, economic issues and regularly publishes poll outcomes. For the 2012 US Presidential elections, Gallup collected data[7] from approximately 3,050 registered voters by asking whom they would vote for if the elections were held at that time and reported seven day rolling averages for the opinion poll ratings. The models we described in this paper predict these ratings using tweets collected from Twitter.

Twitter data

Using the freely available twitter4j[8] Java API library we collected over 37 million tweets between Sept 7, 2012 and Nov 7, 2012. These APIs are restricted by the rate limits[9] imposed by Twitter and collect a small fraction of the total data available at any given time. The following meta-data was collected along with the content of the tweets:

- Screen name and Twitter ID of author of tweet
- Screen name and Twitter ID of original author of tweet, in case the tweet is a retweet
- Screen name and Twitter ID of author of parent tweet, if a tweet is a response.
- Twitter ID of tweet
- Date and time of posting of the tweet.

3.2 Data preprocessing

From each post in the collection, bigrams were extracted after stop word removal and stored along with the text. These were indexed along with the raw data from Twitter and stored in Apache Lucene[10], a popular text search engine library for Java.

3.3 Training - Long term topic detection

Using the poll ratings collected from Gallup as the predicted variable, and using each of the features defined in the model as predictors we trained regression models in a seven day window. Thus, data from the last seven days was used to train a regression model and predict the poll rating for both presidential candidates on the eighth day.

3.4 Causal analysis of bigrams - US Elections

By running the lasso granger regression model using the predicted variables and the bigram features, we identified bigrams (long-term topics) that were causal for the predictions. We found 227 bigrams that were causal for Barack Obama during the elections and 183 bigrams for Mitt Romney. Some examples of bigrams found to be causal for the two US presidential candidates are shown in Table 1. As can be seen from Table 1, bigrams related to "taxes", "job growth", "osama bin laden", " automobile bailout", "benghazi attack', etc were found to be contributing to the approval ratings. Causal bigrams for Mitt Romney include some with negative sentiments such as,"obama_lied", "failed_policies" etc (Note: These were causal bigrams and no sentiment analysis was done to identify these).

Barack Obama	Mitt Romney
american_soil_	acceptance_letter_
auto_bailout_	al_qaeda_
bin_laden_	arab_spring_
business_owners_	birth_certificate_
climate_change_	bush_americans_
fiscal_cliff_	cut_deficit_
fix_economy_	dnc_obama_
foreign_policy_	democratic_national_
job_growth_	dnc_speech_
hillary_clinton_	equal_pay_
trillion_debt_	failed_policies_
obama_hillary_	fiscal_cliff_
taxes_obama_	health_insurance_

Table 1: Examples of causal bigrams (long-term topics) found for the two candidates in 2012 US Presidential Election

3.5 Election day approval rating prediction - US elections

Gallup did not make available the poll ratings in the final week running up to the election. Therefore, we trained the model using the last seven days of available data and used it to predict the outcome on election day using the tweets. Our model predicted 47.8 % of the vote to Barack Obama and Mitt Romney receiving 47.2 %. These prediction values are very close to the poll predictions made by different polling agencies[11].

We also used candidate mentions in tweets as an estimate of popularity (commonly used in previous work), and were able to detect a clear win for Barack Obama but the vote shares predicted for the candidates were off from actual vote shares by a very large margin. We found no correlation between the number of tweet mentions of the candidates and the daily approval ratings. Barack Obama won the 2012 US Presidential election with 50.8% of the overall vote.

3.6 Philippines Elections

This section describes details about the live deployment and use of our system during the 2013 Filipino General elections. The elections in Philippines were held on the 13th May 2013 to elect 12 of the 24 seats in the senate. Together with the senators elected in 2010, the candidates elected in this election formed the 16th Congress of the Philippines. A total of

[6]http://www.gallup.com
[7]http://www.gallup.com/poll/150743/Obama-Romney.aspx
[8]http://twitter4j.org/
[9]https://dev.twitter.com/docs/rate-limiting
[10]http:/lucene.apache.org/
[11]http://www.realclearpolitics.com/epolls/2012/president/us/general_election_romney_vs_obama-1171.html


Home › Focus ›

Citizens use social media to complain vs 'defective PCOS'

by [redacted]

Posted at 05/13/2013 4:56 PM | Updated as of 05/13/2013 5:27 PM

The term 'PCOS' continues to trend in Twitter as reports of alleged mafunction pour in the social media.

MANILA, Philippines – Complaints against the precinct count optical scan (PCOS) machines echoed through the four walls of social media during Monday's Philippine midterm elections, pushing 'PCOS' to a Twitter trending topic for most of the morning.

The same was reflected in the monitoring of the [redacted] where

Figure 2: The news story regarding faulty voting machines as it appeared on the media house's website after it was detected by our system as a possible news story.

33 candidates contested for the election with two main coalitions - Team PNoy and United Nationalist Alliance (UNA).

Twitter data for the Philippines election was collected using the Gnip Powertrack stream which provides all the tweets which contains the user specified keywords. We collected only election related tweets using specific keywords formulated using the names of candidates, political parties, election campaigns etc. The structure of the data was similar to the Twitter data for US Elections and the pre-processing as described earlier was performed. A total of 7.5 million tweets were collected from March 20, 2013 up to the elections on May 13, 2013.

In order to train the long-term topics module during the Filipino election, we used the electoral surveys generated by Pulse Asia[12] that contained opinion polls for all candidates in the 2013 General Elections.

3.7 System Deployment and Results - Filipino Election

Deployment : Around 4 million Tagalog and 3.5 million English Tweets were analyzed over a 50 day period using our system. Data was analysed on a daily basis to identify the buzz and sentiment around candidates and political parties. In addition, trending topics were identified each day. However, during debates (held twice during the pre election period) and 2 days leading up to the election, the reports were delivered at a 4 hour interval. Each batch of tweets was run through the buzz and sentiment detection module before being consolidated for candidates and political parties. After this, the data was passed through the topic detection module which identified the trending topics in that batch. The topics delivered in each batch along with the sentiments for candidates helped the reporters to arrive at a news story which could be published immediately or investigated further. This story would contain the insights that were derived from our analysis in addition to the actual event that occurred, thus making the story interesting and unique for the news media.

[12] http://www.pulseasia.ph/EP2013.html

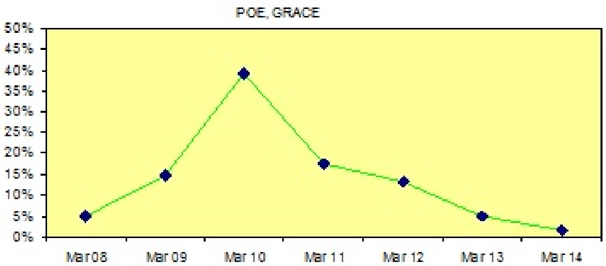

Figure 3: Grace Poe released her TV ad on March 10 which drew flak from viewers leading to a peak in negative sentiment.

Experiences and Results : During May 1 to May 13 (day of polling), all the insights were hosted on the media house's election page garnering about 1.5 million hits. More importantly, our system (specifically the local event detection algorithm) was able to discover *breaking news* off social media. For example, during election day, our tool discovered defective voting machines as emerging topics. The reporters at media house followed this story and were the first to report in through their news channel (much ahead of their competitors). Figure 2 shows a print version of the story as it appeared on the website of the media house. As an additional validation, we tracked the TRP of programs showing social insights from our work and found it to be 25% more than election programming by other channels. The hashtag (#Halalan2013) on which the social media insights generated by our system were disseminated was found to be top trending hashtag in Philippines[13].

Senatorial Candidate Nancy Binays tiff with comedian Vice Ganda generated a lot of debate on the social media. When reporters saw the keywords "Vice Ganda", "Comedy", "Binay" and "Risa" among the trending keywords, they immediately could relate to the actual story (Vice Ganda making a comment on Nancy Binay about not knowing enough about being a public servant). We also had instances when social media gauged public mood very strongly, for example, Figure 3 shows how negative sentiment around Grace Poe increased when a TV ad was aired.

As another example, Nancy Binay lead the share of tweets by 78% when compared to 22% share of another candidate, Grace Poe. However, Grace Poe actually lead by 16 million votes as against 13 million for Nancy Binay in the fi-

[13] http://www.medianewser.com/2013/05/abs-cbns-halalan-2013-is-most-watched.html

Grace Poe	Francis “Chiz” Escudero
maaga_ito_	scudero_para_
si_heart_	rt_pingmedina_
ate_grace_	gmanews_ang_
sa_pilipinas_	escudero_saychiz_
poe_sila_	isda_nsug_
sinimulan_ni_	alanpcayetano_sonnyangara_
evangelista_believes_	waterfall_restaurant_
tutuloy_nya_	rt_iamaljangeles_
escudero_para_	ipagmamalaking_karanasan_
sino_nga_	magalang_grace_
humor_painfulthings_	senator_saychiz_
escudero_saychiz_	chiz_mahal_

Table 2: Examples of causal bigrams (long-term topics) found for leading candidates in 2013 Filipino General Election. Bigrams are in the Tagalog language and some also refer to candidate names.

nal results. Further analysis showed that most of the tweets for Nancy Binay, talked about her birthday near the election day. Most others talked about her tiff with Vice Ganda. Furthermore the sentiment results showed that Grace Poe has more positive sentiments (31% vs 21%) and less negative (6% vs 14%) as compared to Nancy Binay, which seemed to agree with the actual electoral outcomes. In yet another example, emerging keywords unearthed Senator Miriam Santiago’s tweet supporting candidate Sonny Angara because he has “sex appeal”, also created a news story.

Table 2 shows some of the important long-term topics found in the elections. As can be seen, candidate Escudero Chiz’s controversial fight between his girlfriend Heart Evangelista and her parents was identified as a long-term topic (“si_heart_”, “chiz_mahal_”) and this event indeed featured prominently during the elections.

The supplementary notes[14] contain further details on the deployment as well other models we developed during the course of this work.

4 Related work

Researchers have tried to measure the impact of social media on many different applications including prediction of product sales [Bhatt *et al.*, 2010], movie award winners [Bothos *et al.*, 2010], stocks [Bollen *et al.*, 2011], project risks [Wolf *et al.*, 2009], disease transmission [Sadilek *et al.*, 2012], rumour detection [Finn *et al.*, 2014] and elections [Tumasjan *et al.*, 2010]. Most work involving social media data and elections has focused on tasks such as election forecasting.

[Tumasjan *et al.*, 2010] used Twitter for predicting German elections. They used mentions of political party names and concluded that the relative number of tweets mentioning the names is a good indicator to predict elections. Similar conclusion was reached by [William and Gulati, 2008] where Facebook supports were indicative of the 2008 US election results. But, this method was found to be sensitive to the period from which Tweets were chosen for counting and to omission of smaller parties from the counting process. The work in [Sang and Bos, 2012] found that just counting the number of tweets is not enough and it is important to filter high quality tweets for counting entity mentions along with using sentiment for prediction. They applied their system for predicting the 2011 Dutch elections. The work in [Chung and Mustafaraj, 2011] also contended that merely counting mention of party or candidate is not enough and recommended to discount spam tweets and incorporate the tweet sentiment in the prediction system. In [O’Connor *et al.*, 2010] the authors found that the sentiments in Twitter showed correlation to approval polls. The work in [Gayo-Avello *et al.*, 2011] experimented with previously proposed election prediction models including that of [O’Connor *et al.*, 2010] and [Tumasjan *et al.*, 2010] and in contrast, found that the accuracy of prediction was very low. All of the above work have concentrated on predicting the final outcome of the elections. Our experience in both the 2012 US General Elections and 2013 Filipino elections also demonstrates that hashtags or entity mentions and tweet volumes etc do not necessarily indicate election outcomes and we adopt a different approach. Instead we focus on using Twitter to track the issue based support of people and use granger causality, to identify long-term *causal* topics in an election.

[Xie *et al.*, 2014] present a system which identifies bursty topics and viral events by comparing the distribution of words in a time slice with an archival snapshot. One of the drawbacks of this method is that the number of unique words to keep track of can become very large. In contrast, our method for emerging topics detection uses frequent item set mining which is further constrained using seed words. Other related work by [Rill *et al.*, 2014] describes methods to detect emerging political topics on Twitter but relies heavily on the presence of topical hashtags which may not always be present.

5 Discussion and Conclusion

In this paper we described our work for monitoring social media during elections and demonstrate in a live election, how they can be useful in gauging public sentiment, detecting breaking news etc. We demonstrated our work on two different elections from two vastly different geographies, languages and electoral systems. We show that for some of the large scale events, social platforms may provide information ahead of traditional news sources. While many discussions on social media were more reactive, during our deployment, we found instances where new events were reported first on Twitter (traditional media lagged by few hours !). In the future we would like to develop richer online topic detection models as the methods developed for offline large scale text processing cannot directly be applied to live streaming data. We also plan to deploy and study our work in the context of other events such as sports competitions and movie releases.

[14] http://www.researchgate.net/publication/275539820_Tracking_political_elections_on_social_media__Applications_and_Experience_(Supplementary_Material)

References

[Agrawal *et al.*, 1993] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, SIGMOD '93, pages 207–216, New York, NY, USA, 1993. ACM.

[Bhatt *et al.*, 2010] Rushi Bhatt, Vineet Chaoji, and Rajesh Parekh. Predicting product adoption in large-scale social networks. In *Proceedings of ACM international conference on Information and knowledge management*, pages 1039–1048, 2010.

[Bollen *et al.*, 2011] Johan Bollen, Huina Mao, and Xiao-Jun Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011.

[Bothos *et al.*, 2010] Efthimios Bothos, Dimitris Apostolou, and Gregoris Mentzas. Using social media to predict future events with agent-based markets. *IEEE Intelligent Systems*, 25(6):50–58, 2010.

[Chung and Mustafaraj, 2011] Jessica Elan Chung and Eni Mustafaraj. Can collective sentiment expressed on twitter predict political elections? In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2011.

[Contractor and Faruquie, 2013] Danish Contractor and Tanveer Afzal Faruquie. Understanding election candidate approval ratings using social media data. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 189–190. International World Wide Web Conferences Steering Committee, 2013.

[Finn *et al.*, 2014] S Finn, PT Metaxas, E Mustafaraj, M OKeefe, L Tang, S Tang, and L Zeng. Trails: A system for monitoring the propagation of rumors on twitter. *Computation and Journalism Symposium, NYC, NY*, 2014.

[Gayo-Avello *et al.*, 2011] Daniel Gayo-Avello, Panagiotis T. Metaxas, and Eni Mustafaraj. Limits of Electoral Predictions Using Twitter. In *International AAAI Conference on Weblogs and Social Media (posters)*, July 2011.

[Gayo-Avello, 2012] Daniel Gayo-Avello. I wanted to predict elections with twitter and all i got was this lousy paper - a balanced survey on election prediction using twitter data. 2012.

[Granger, 1969] C. W. J. Granger. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3):424–438, August 1969.

[Jungherr *et al.*, 2012] A. Jungherr, P. Jürgens, and H. Schoen. Why the pirate party won the german election of 2009 or the trouble with predictions: A response to tumasjan, a., sprenger, to, sander, pg, & welpe, im predicting elections with twitter: What 140 characters reveal about political sentiment. *Social Science Computer Review*, 30(2):229–234, 2012.

[Krishnamurthy *et al.*, 2009] Rajasekar Krishnamurthy, Yunyao Li, Sriram Raghavan, Frederick Reiss, Shivakumar Vaithyanathan, and Huaiyu Zhu. Systemt: A system for declarative information extraction. *SIGMOD Rec.*, 37(4):7–13, March 2009.

[O'Connor *et al.*, 2010] Brendan O'Connor, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, 2010.

[Osborne and Dredze, 2014] Miles Osborne and Mark Dredze. Facebook, twitter and google plus for breaking news: Is there a winner? In *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[Petrovic *et al.*, 2013] Sasa Petrovic, Miles Osborne, Richard McCreadie, Craig Macdonald, and Iadh Ounis. Can twitter replace newswire for breaking news? 2013.

[Rill *et al.*, 2014] Sven Rill, Dirk Reinel, Jörg Scheidt, and Roberto V. Zicari. Politwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *Knowl.-Based Syst.*, 69:24–33, 2014.

[Sadilek *et al.*, 2012] A. Sadilek, H. Kautz, and V. Silenzio. Predicting disease transmission from geo-tagged micro-blog data. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[Sang and Bos, 2012] Erik Tjong Kim Sang and Johan Bos. Predicting the 2011 dutch senate election results with twitter. In *Proceedings of the Workshop on Semantic Analysis in Social Media*, 2012.

[Tumasjan *et al.*, 2010] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe. Predicting elections with twitter: What 140 characters reveal about political sentiment. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 178–185, 2010.

[William and Gulati, 2008] C William and G Gulati. What is a social network with facebook and vote share in the 2008 presidential primaries. In *Annual Meeting of American Political Science Association*, 2008.

[Wolf *et al.*, 2009] Timo Wolf, Adrian Schroter, Daniela Damian, and Thanh Nguyen. Predicting build failures using social network analysis on developer communication. In *Proceedings of International Conference on Software Engineering*, 2009.

[Xie *et al.*, 2014] Runquan Xie, Feida Zhu, Hui Ma, Wei Xie, and Chen Lin. Clear: A real-time online observatory for bursty and viral events. *Proc. VLDB Endow.*, 7(13):1637–1640, August 2014.