Relating Political Party Mentions on Twitter with Polls and Election Results

Eric Sanders
CLS/CLST, Radboud University Nijmegen
Erasmusplein 1
6525 HT Nijmegen
+31 24 3616087
e.sanders@let.ru.nl

Antal van den Bosch
CLS, Radboud University Nijmegen
Erasmusplein 1
6525 HT Nijmegen
+31 24 3611647
a.vandenbosch@let.ru.nl

ABSTRACT
In each of the last ten days preceding the parliamentary elections of 2012 in the Netherlands at least one election poll was published. Throughout the same period close to 170 thousand Dutch microtext messages with references to political parties were posted on Twitter, the microblogging platform. In this study we investigate whether these tweets can serve as an addition to, or even an alternative for the traditional polls as predictors of the election outcomes. We show that counts of mentions of political party names are strongly correlated with the polls and the election results. While polls remain more accurate as a predictor of the outcome (a mean absolute error of 1.1% and a correlation of about 0.98 with the actual percentage of votes cast for all parties), the Twitter statistics show a mean absolute error of 1.9% when aggregated over a number of days, and display a high correlation with elections and polls (in both cases, $r \approx 0.95$). We conclude that tweet mention counts form a good complementary basis for predicting election results.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms
Human Factors, Languages.

Keywords
Twitter, Elections, Polls.

1. INTRODUCTION
With a current average of about a half billion messages posted daily, Twitter hosts a massive amount of accessible messages, which in turn harbor vast amounts of information. Tweets are often related to personal affairs, but may also refer to popular events. One of the interesting uses of the information in tweets is to try to determine people's opinions about certain matters. Politics is an attractive subject to try to get opinions about from tweets. In terms of events, political elections typically evoke the posting of tweets containing political views.

A conventional way of assessing average opinions about politics during election periods is polling. The standard polling method is to ask a small but representative part of the population what party or person one is planning to vote for. On Twitter people give this information without being prompted. It would be an interesting addition to (or even alternative to) polls if we could extract this information from tweets. The most challenging part of it is to gather a balanced representation from the tweets of the people participating in the elections. In essence this is impossible; while the legal voting age in the Netherlands is 18, many users on Twitter have not reached that age, but demographic information regarding individual users is not available in any trustworthy way on Twitter. The sheer magnitude of data available on Twitter may compensate for this partly unrepresentative information.

In this paper a comparison between the predictive potential of tweets and polls with respect to the outcome of the Dutch parliament elections of 12 September 2012 is presented. The number of times a political party is mentioned in a Dutch tweet is compared to the polls and the election results without normalization. This was done for all eleven parties that won at least one seat in the parliament. The next sections discuss related work, describe the data, explain the experiment, discuss the results, and draw conclusions.

2. RELATED WORK
Work that has focused on predicting election outcomes through social media mining offers a mixed bag of results. Tumasjan *et al* [1] show that for the six biggest parties in the election of the German parliament held on 27 September 2009, the percentage of tweets in which a party is mentioned between 13 August and 19 September 2009 highly correlates with the election result of that party. Their particular selection of parties and the period over which counts were gathered is questioned in a responding paper by Jungherr *et al* [2]. They claim that the choices made by Tumasjan et al give overly optimistic results on badly grounded heuristics.

O'Connor *et al* [3] compare the sentiment ratio of tweets containing 'obama' with presidential job approval polls in 2009 and presidential election polls in 2008. The ratio correlates well with the first poll but does not with the latter. Marchetti-Bowick and Chambers [4] build on the work of O'Connor et al. and use distant supervision for both topic identification and sentiment analysis. The comparison of the results with Obama's job approval poll gives better correlation than earlier work.

Tjong Kim Sang and Bos [5] compare tweet mentions and election results for the Dutch senate elections of 2011. Beyond raw counts of tweets they test and compare the predictive power of four alternative counting methods, but they do not find large improvements with these methods.

The novelty of the work described in this paper is that it is based on a relatively large number of consecutive polls on each of the ten days before the elections.

Gayo-Avello [6] pinpoints a couple of problems with predicting elections based on tweets and gives some suggestions. Apart from those addressed in this paper, he indicates that only good results are published and analyzing afterwards is not predicting.

3. DATA

The Twitter data used in the experiments is taken from a substantial archive of Dutch tweets collected within the TwiNL project (ifarm.nl/erikt/twinl). The FAQ of the related search website twiqs.nl states that an estimated 40% of all Dutch tweets are collected since December 16, 2010. The present study makes use of all tweets gathered between September 2 to September 12, 2012, for which between 2.0 and 2.4 million tweets per day have been archived.

The poll data is taken from the website Alle Politieke Peilingen (www.allepeilingen.com) that has saved the poll results from 2000 onwards of the six most cited polling institutes in the Netherlands. These are: peil.nl, TNS NIPO, de politieke barometer, buzzpeil.nl, de Stemming and NOS Peilingwijzer. All these polls try to predict the result of the elections (if the elections were held on the day of the poll).

4. EXPERIMENT

For the eleven parties that won one or more seats in parliament we counted how often the party name was mentioned in a tweet in the ten days before the elections and on election day, 12 September 2012. This was done with a basic pattern match. First it was investigated by which names parties are mentioned in the tweets. Most parties are almost exclusively mentioned by their abbreviation and rarely by their full name. Most full names are therefore ignored. For instance, the acronym of the VVD occurs over thousand times more often than its full name, ‘Volkspartij voor Vrijheid en Democratie’. However, two parties are often mentioned by their full name: GroenLinks and ChristenUnie. Their respective abbreviations can also have other meanings: GL being a typical English shorthand for ‘good luck’ and CU for ‘see you’, but a manual inspection revealed that these abbreviations are rarely used in these meanings.

We needed to generate several specific pattern-matching expressions. Three parties have ‘van de’ (‘of the’) or ‘voor de’ (‘for the’) in their full name which can be expressed in many ways, e.g. ‘vd’, ‘v.d.’, ‘v/d’, ‘van de’, ‘v d’, which are all represented in the search pattern that was used. Matching is case-insensitive, so ‘SGP’, ‘sgp’, ‘Sgp’ etc. are all recognised. No effort was made to find misspelled party names. The party names can be preceded by ‘@’ (Twitter account names) or ‘#’ (Twitter hashtags) and preceded or followed by punctuation.

Table 1 lists the resulting regular expressions for the parties.

During the period of ten days before the election, for each day and each party, the percentage of tweets in which a party is mentioned is compared to the result of the average of all polls that came out that day. This was done to investigate how much the percentage of party mentions in tweets resembles the polls. Subsequently, the results of the averaged polls on the day before the election and the election results are compared to each other and to the tweet mentions of (1) election day, (2) the day before election day, (3) an aggregate of all tweets during the 10-day period before the elections, and (4) an aggregate over a 5-day period before the elections.

Table 1. The regular expressions that were used to detect the party names in the tweets

<table>
<tr><th>Party</th><th>Regular Expression Pattern</th></tr>
<tr><td>VVD</td><td>"vvd"</td></tr>
<tr><td>PVDA</td><td>"pvda","partij\s+v(oor\s+|an\s+|.)?d(e|.)?\s+arbeid"</td></tr>
<tr><td>SP</td><td>“sp”</td></tr>
<tr><td>PVV</td><td>"pvv","partij\s+v(oor\s+|an\s+|.)?d(e|.)?\s+vrijheid"</td></tr>
<tr><td>CDA</td><td>“cda”</td></tr>
<tr><td>D66</td><td>"d\'?66"</td></tr>
<tr><td>GL</td><td>"gl","groen.?links"</td></tr>
<tr><td>CU</td><td>"cu","christen.?unie"</td></tr>
<tr><td>SGP</td><td>“sgp”</td></tr>
<tr><td>PVDD</td><td>"pvdd","partij\s+v(oor\s+|an\s+|.)?d(e|.)?\s+dieren"</td></tr>
<tr><td>50PLUS</td><td>"50[^\d]?(\+|plus)"</td></tr>
</table>

An average of 0.7% of all daily tweets posted throughout the last ten days before the election mentions at least one political party. Table 2 shows that these nearly 170 thousand tweets are not uniformly divided over the eleven days; about one third of all tweets is posted on election day, and more tweets are posted closer to election day.

5. RESULTS

First, comparisons are shown in three figures (Figures 1, 2, and 3) between daily percentages of Twitter mentions and daily poll results of selections of two or three parties during the ten days before the elections.

The daily percentage of Twitter mentions for a particular party is computed as follows:

$$Perc = 100 * \#mentions_p / \sum_p \#mentions_p$$

where $\#mentions_p$ is the number of mentions of a particular party, and $\sum_p \#mentions_p$ is the total number of mentions of all eleven parties. The counts thus represent mentions, not tweets: if in a tweet two parties are mentioned, the tweet is counted twice.

The percentages of poll results are computed from the predicted number of parliament seats, which is the statistic by which they are reported and stored. As there are 150 seats in the Dutch parliament, each seat stands for 0.67%. The percentage used here is the mean percentage of the predicted number of seats of all polling institutes that released a prediction that day. For some days there is a poll of only one institute. The predictions of the polling institutes differ slightly. The largest difference between two predictions from poll estimates for the same party on the same day is 4.7%. On the day before the elections, 11 September, all polling institutes published results.

Table 2. Number of tweets with at least one political party mentioned in the 10 days before elections and on election day (0 days)

Days before election	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
#tweets	56,580	24,004	17,224	8,498	12,011	10,178	9,373	9,062	10,317	8,048	4,700

5.1 Twitter vs Polls Correlation

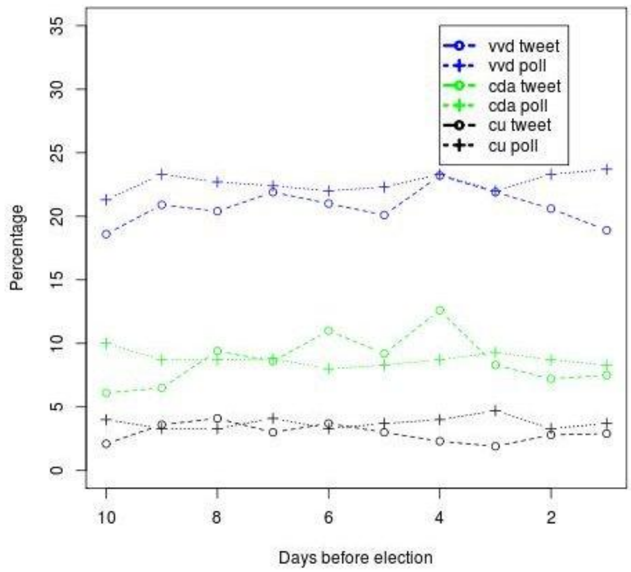

Figure 1. Twitter mentions and poll results for VVD, CDA and CU

Figure 1 displays the results for VVD, the party that won the elections, CDA, a middle party, and ChristenUnie (CU), a small party. The figure exemplifies the fairly strong correlation of the percentages of Twitter mentions and poll results during the whole period.

5.2 Twitter vs Polls Outliers

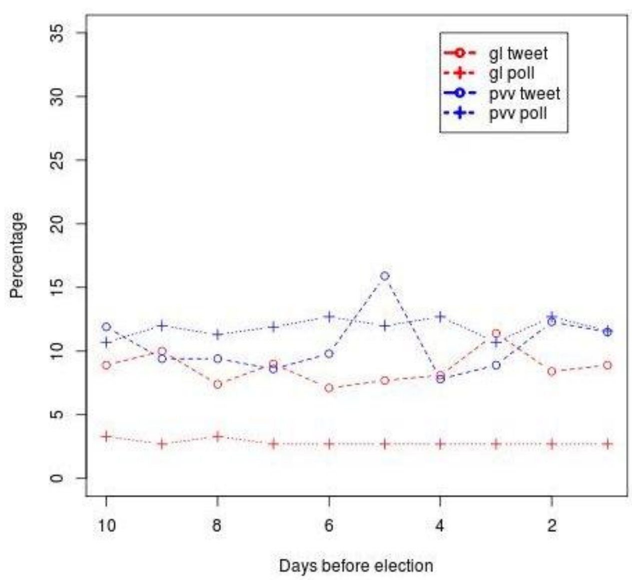

Figure 2. Twitter mentions and poll results for PVV and GL

This trend is typical for all but one party, GroenLinks (GL), as shown in Figure 2. For comparison, the GroenLinks estimates are compared against the predictions for the PVV. The figure displays an unexpected difference between the Twitter mentions and poll results for GroenLinks. This party is well known for its above-average use of and presence on social media in their campaign [7].

As an aside, the figure also shows a relatively high peak in the Twitter mentions of the PVV five days before the elections. This may be explained by the news that day that the PVV had falsely declared money from the European Union, while their campaign was outspokenly anti-Europe.

5.3 Twitter vs Polls Trend

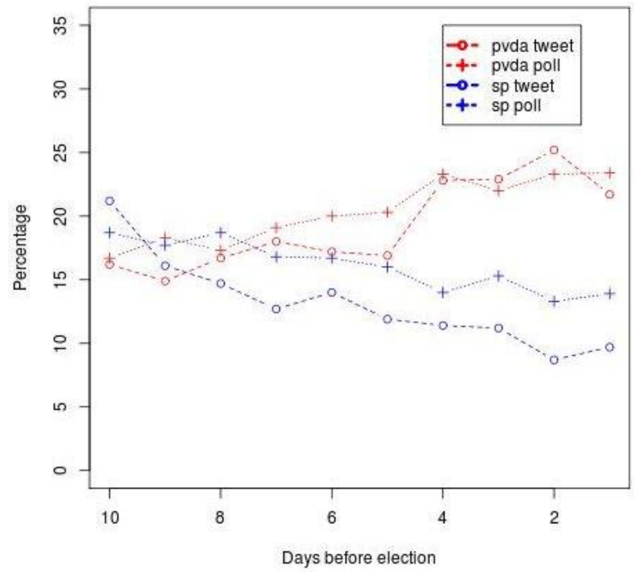

Figure 3. Twitter mentions and poll results for PVDA and SP

Figure 3 shows how both in the Twitter mentions as in the poll results the PvdA, one of the socialist parties and runner-up in the election results, was gaining in the last ten days before the elections while the SP, another socialist party, was losing voters. The SP started out popular, but in the debates the PvdA leader was doing well, while the SP leader's debating was considered disappointing.

5.4 Twitter vs Polls vs Election

Table 3 shows for all parties the difference between the election results on 12 September, the mean result of all polls on the day before the elections, and the relative percentage of tweets the party was mentioned on (1) election day, (2) the day before, (3) during all ten days and (4) during five days before the elections. The fourth and second rows from below list the mean absolute error (MAE) of the column with the election results (2nd column) and with the polls of the pre-election day (3rd column). The third last and final row show the correlation and the 95% confidence interval with the election and poll results.

The MAE of the polls with the election results is smaller than the MAE of the tweet mentions with the election results in all cases, meaning that polls are a better predictor of the election results than raw counts of party names in tweets. The table also shows that tweet mentions of a time span of several days (five or ten) before the elections are closer to the election results than the tweet mentions on one specific day (election day or the day before). Tweet mentions gathered during five days before the elections are closer to the election results than all tweet mentions from ten days before the election results. Finally, the correlation coefficient and the confidence interval show the same trend as the MAE, and are very high in all cases; 0.93 or higher.

Table 3. Comparison between election results, polls and tweets from different time slots in %

Party	Election 12 Sep	Polls 11 Sep	Tweet 12 Sep	Tweet 11 Sep	Tweet 2-11 Sep	Tweet 7-11 Sep
VVD	*26.8*	23.7	24.6	18.9	20.7	20.6
PVDA	*25.1*	23.4	18.5	21.7	20.2	22.2
PVV	*10.2*	11.6	13.6	11.5	10.7	11.4
SP	*9.8*	13.9	8.7	9.7	12.0	10.3
CDA	*8.6*	8.3	6.0	7.5	8.6	8.6
D66	*8.1*	7.9	9.8	9.7	9.0	8.5
CU	*3.2*	3.7	2.6	2.9	3.0	2.7
GL	*2.4*	2.7	7.0	8.9	8.6	8.8
SGP	*2.1*	1.7	3.2	4.4	2.9	2.8
PVDD	*2.0*	1.8	3.6	3.5	3.2	3.2
50PLUS	*1.9*	1.7	2.4	1.3	1.1	1.1
MAE elections		*1.1*	*2.4*	*2.4*	*2.2*	*1.9*
Corr elections		*0.98 (0.93 -1.0)*	*0.95 (0.82-0.99)*	*0.94 (0.78-0.98)*	*0.95 (0.83-0.99)*	*0.96 (0.84-0.99)*
MAE poll	*1.1*		*2.4*	*2.3*	*2.0*	*1.7*
Corr poll	*0.98 (0.93-1.0)*		*0.93 (0.76-0.98)*	*0.94 (0.78-0.98)*	*0.96 (0.87-0.99)*	*0.96 (0.83-0.99)*

6. DISCUSSION

The results of our comparative study on the 2012 Dutch parliament elections provide case-based evidence that tweets are a good basis for predicting election results. Purely on the basis of raw counts of party name mentions (with flexible pattern matching rules), without further domain knowledge, a strong correlation with the poll results can be observed (around 0.95). In a number of cases the difference between the Twitter mentions and the polls is larger than 5%, but the difference between the various polls is also almost 5% in a few cases. Although the polls more accurately predict the election outcome, the correlation between tweet-based estimates and the outcome is observed to be as high as 0.96, with a mean absolute error of only 1.9% (the polls attain 1.1%), provided that the tweet counts are aggregated over a number of days.

As Gayo-Avello rightly points out in his paper [6] our kind of approach lacks information that could improve the prediction of election outcomes or poll results based on Twitter. First, who is tweeting? If the Twitter account is from a party member or official the tweet could be filtered out as it may be used to steer social media opinions or even statistics. However, it is hard to ascertain whether a Twitter account is from a party member. Automatic profiling based on machine learning and text classification may help in this respect. Second, is the tweet polar or neutral? A Twitter user who will vote for a party is likely to compose positive tweets about that party. Automatic sentiment analysis (perhaps trained on political opinions to capture domain-specific sentiment markers) might be used to reweight counts.

Negation and hedging may be a third factor that could partially be determined automatically and improve estimates. A tweet such 'I will not vote for partyX' could then be left out of the count for partyX. This is a very challenging task, though. Morante and Daelemans [8] provide pointers on how this may be addressed. Fourth, can we account for factors that cause an increase in the number of tweets of a certain party? The detection of other events involving entities that also play a role in the focus event (such as the PVV scandal mentioned in the discussion of Figure 2) may be used to discount tweets about this event.

Finally, we observed that estimates based on counts aggregated over several days better approximated the election results than the counts on a specific day; five days seem to represent a reasonable aggregation window. A further study could be carried out to see whether an optimal time window can be found for events similar to the single case studied here.

We do not share Gayo-Avello's conclusion that elections cannot be predicted with Twitter, but acknowledge that further research has to be carried out before we say Yes we can! (predict elections with Twitter).

7. ACKNOWLEDGMENTS

We thank Ruut Brandsma from www.allepeilingen.com for providing the data from the polling institutes.

8. REFERENCES

[1] Tumasjan, A., Sprenger, T., Sander, P., Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the fourth International AAAI Conference on Weblogs and Social Media, Washington D.C., USA, 2010*

[2] Jungherr, A., Jürgens, P., Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T.O., Sander, P.G., Welpe, I.M. "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment", *Social Science Computer Review*

[3] O'Connor, B., Balasubramanyan, E., Routledge, B., Smith, N. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceedings of the fourth International AAAI Conference on Weblogs and Social Media, Washington D.C., USA, 2010*

[4] Marchetti-Bowick, M., Chambers, N. (2012). Learning for Microblogs with Distant Supervision: Political Forecasting with Twitter. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon, France, 2012*

[5] Tjong Kim Sang, E. and Bos. J. (2012). Predicting the 2011 Dutch Senate Election Results with Twitter, *Proceedings of SASN 2012, the EACL 2012 Workshop on Semantic Analysis in Social Networks, Avignon, France, 2012*

[6] Gayo-Avello, D. (2012). No, You Cannot Predict Elections with Twitter, *IEEE Internet Computing, vol. 16, no 6. pp. 91-94, Nov.-Dec. 2012*

[7] http://www.sax.nu/Portals/615/docs/Rapport_Saxion_Social_Media_Politiek_2012.pdf

[8] Morante, R., and Daelemans, W. (2009). A metalearning approach to processing the scope of negation. *Proceedings of the Thirteenth Conference on Computational Natural Language Learning, pp. 21-2), 2009*