Opinion Dynamics of Elections in Twitter

Felipe Bravo-Marquez[1] Daniel Gayo-Avello[2] Marcelo Mendoza[3,4]
Barbara Poblete[1,4]

Department of Computer Science, University of Chile[1]

University of Oviedo, Spain[2]

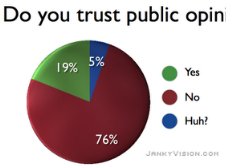
Universidad Técnica Federico Santa María, Chile[3]

Yahoo! Research Latin America, Chile[4]

October 25, 2012

Public Opinion

- A complex collection of beliefs held by the population about specific **subjects** or **topics**.
- Public opinion has been traditionally evaluated using **polls**.



Drawbacks of Polls

- Polls need to be conducted periodically in order to track a target topic over time.
- Polls are unable to detect instantaneous changes in public opinion [Akcora et al., 2010].
- The way that polls are worded can bias the overall evaluation (cf. [Schuman and Presser, 1996]).

Social Media

- Microblogging services are increasingly being adopted by people in order to access and publish information.
- **Twitter**: Massively used Microblogging platform where users post messages limited to 140 characters.
- Twitter users tend to publish **personal opinions** regarding certain topics and news events.

Opinion Mining or Sentiment Analysis

- Application of **NLP** and **text mining** techniques to identify and extract subjective information from textual datasets [Pang and Lee, 2008].

Opinion Mining Problems

1. Subjectivity Evaluation: subjective or objective?
2. Polarity Evaluation: positive or negative?

Approaches

1. Supervised: Based on corpora of **evaluated opinions** and machine learning.
2. Non-supervised: Based on a lexicon of **opinion words** and evaluation functions.

Public Opinion, Opinion Mining, and Social Media

- Opinions are provided **freely and voluntarily** by the users in Twitter.
- Posted opinions could be **aggregated** and used to measure the public opinion **implicitly**.

Limitations of Social Media for Public Opinion

- Population which uses social media platforms is not necessarily a **representative sample** of the entire population.

Benefits

- Cheaply process greater amounts of data [Yu and Kak, 2012].
- Opinions become available in continuous time streams $\Longrightarrow$ suitable for studying the **temporal** properties of public opinion or **opinion dynamics**.

Related Work

- There has **emerged** an interest in understanding **temporal aspects** of opinions and furthermore, in **predicting future events** from social media.
- [Mishne and de Rijke, 2006] showed that opinions exhibit a certain degree of **seasonality** in Twitter.
- [O'Connor et al., 2010] correlated polls and sentiment analysis series from Twitter. The sentiment series are able to capture broad trends in the survey data but show **great variation** among different datasets.
- Other examples: **movie-box-office** sales performance [Asur and Huberman, 2010, Mishne and Glance, 2006], the evolution of the **stock market** [Bollen et al., 2011].

Related Work (II)

Predicting Elections with Twitter

- According to [Tumasjan et al., 2010] the predictive power sentiment analysis applied to **elections** is **"close to traditional election polls"**
- [Gayo-Avello, 2011] states that this power is **greatly exaggerated**.

Discussions

- To the best of our knowledge, no other research work has performed a **deep statistical analysis** of opinion time-series created from social media.
- We strongly believe that the study of the **volatility** and other aspects of opinion time series will allow us to determine the limitations of **assessing opinion dynamics** from social media.

Our Case Study

- We conduct an **experimental exploration** of opinion time series extracted from Twitter related to the 2008 U.S. Presidential **elections**.

- We used ARMA/ARIMA and GARCH models to determine if the series are appropriate for making **reliable forecasts**.
- We focus on different aspects of opinion time series, especially on the notion of opinion *volatility*.

Building Twitter Opinion Time Series

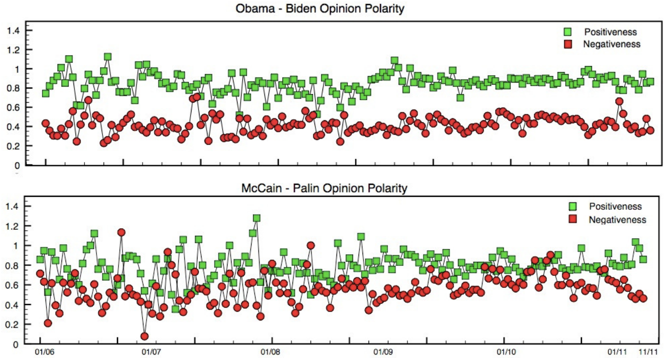
- An opinion value X_t is a measure which reflects a **dimension** of the public opinion regarding a specific topic.
- An opinion time series is **sequence** of opinion values spaced at uniform time intervals $X_1, \ldots, X_t, \ldots, X_n$.
- We need to convert a collection of **timestamped tweets** into opinion time series
- The **sentiment evaluation** of the tweets was made by **counting** their positive and negative opinion words from a given list

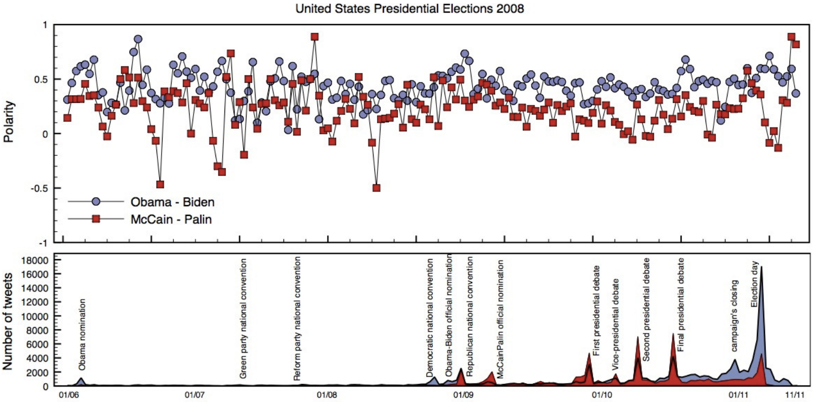
Opinion Values

- **Activity level**: the number of tweets associated to the event in the time period
- **Positiveness**: the average number of positive words per tweet in the period.
- **Negativeness**: the average number of negative words in the period.
- **Polarity**: Positiveness - Negativeness

Dataset

- 250,000 tweets from June 1, 2008 to November 11, 2008.
- All of the tweets are related to either the Democrat ticket (Obama) or to the Republican one (McCain)
- The Twitter Search API was used using one query per candidacy.




United States Presidential Elections 2008
Polarity
1
0.5
0
-0.5
-1
Obama - Biden
McCain - Palin
Number of tweets
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Obama nomination
Green party national convention
Reform party national convention
Democratic national convention
Obama-Biden official nomination
Republican national convention
McCainPalin official nomination
First presidential debate
Vice-presidential debate
Second presidential debate
Final presidential debate
campaign's closing
Election day
01/06
01/07
01/08
01/09
01/10
01/11
11/11

Time Series Analysis Tools

- To establish if an opinion time-series should be discarded as a basis for a **predictive model**, minimum amount of tests should be performed.
- We are interested in modeling the **conditional mean** and the **conditional variance** or **volatility** of the series
- The past conditional variance or volatility of a time-series given past observations, measures the **uncertainty** in the deviation of the time-series from its conditional mean [Cryer and Chan, 2009].
- We model the expected mean using the Box-Jenkins methodology [Box and Jenkins, 1994] based on ARMA/ARIMA models
- Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models [Engle, 1982, Bollerslev, 1986] for the volatility analysis.

Box-Jenkins Models

ARMA(p,q) model

$$X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \epsilon_{t-j} + \epsilon_t \qquad (1)$$

- Process Identification
- ACF and PACF
- ADF test for stationarity
- Differentiation
- Parameter Estimation
- Residuals Analysis
- Trend analysis
- Seasonal Models

Volatility

- ARIMA models cannot generate reliable predictive model in presence of **volatility**.
- The patterns of changing from quiet to volatile periods is named as *volatility clustering*.
- The conditional variance or volatility can be **estimated** and forecasted using **GARCH** models.

GARCH(q,p) model

- Let $\sigma^2_{t|t-1}$ be the expected conditional variance or volatility of a zero-mean time-series r_t at period t

$$\sigma^2_{t|t-1} = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon^2_{t-1} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j|t-j-1} \tag{2}$$

Volatility in Opinion Time Series

- Intuitively, during hectic periods, people tend to be more sensitive to information and hence opinion trends register larger **fluctuations**
- GARCH models are applied to zero-mean time series
- A common approach used in financial time series (e.g stock prices) is to model the returns of a positive time series (e.g stock prices) : $r_t = log(\frac{X_t}{X_{t-1}})$
- We could apply this transformation to the **Positiveness** and **Negativeness**

Exploratory Analysis

- The correlations of **Positiveness** and **Negativeness** were not significantly different from zero
- These results validate the idea of modeling positiveness and negativeness as separate time series
- To check the stationarity of the time series we conducted the **Augmented Dickey-Fuller** test

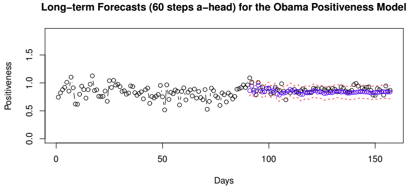
Time series	**ADF test**	***p*-value**
O.+	-7.117	< 0.01
O.-	-9.567	< 0.01
M.+	-10.715	< 0.01
M.-	-6.016	< 0.01

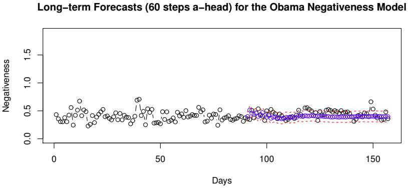
- We reject the hypothesis of non-stationarity!

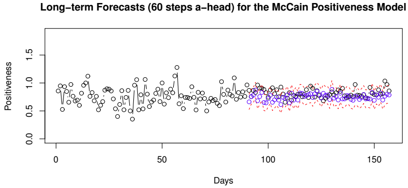
Fitting ARMA models to the Data

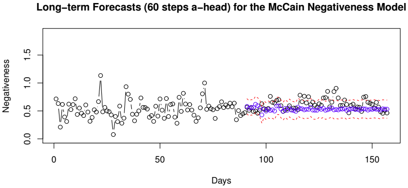
- Time series were separated into two parts, one for model fit/test and a second part for time series forecasting
- Model fitting/testing was conducted over the **first three months** of the US elections
- Model selection was performed by fitting **high order models** to each time series
- The idea is to identify coefficients with significant error standard measures of over-parameterized models

ARMA forecasting results









- Forecasting results are far from being accurate
- Fitted ARMA models can at least model the mean of future outcomes but not the variance

Volatility Analysis

- We used the financial risk management convention of converting the original time series to log return values

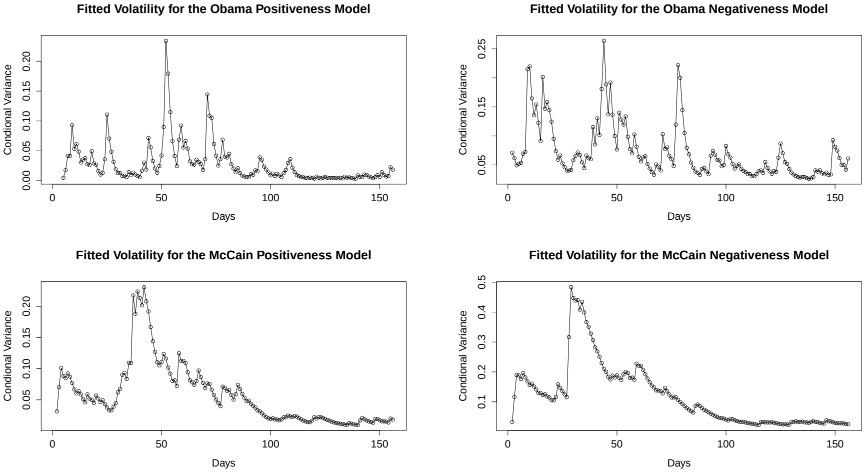
$$r_t = log(\frac{X_t}{X_{t-1}})$$

- We checked statistical conditions of GARCH models

	$R_O.+$	$R_O.-$	$R_M.+$	$R_M.-$
Kurtosis	2.09	0.978	0.346	1.99
Zero-mean t-test p-value	0.936	0.955	0.999	0.925
McLeod-Li avg. p-value	0.000	0.023	0.002	0.000
α_1 t-test p-value	0.000	0.015	0.004	0.001
β_1 t-test p-value	0.000	0.000	0.000	0.000
Mean Volatility	0.028	0.073	0.058	0.119

- We have statistical evidence that GARCH(1,1) models are **appropriate** for modeling our transformed time series

Volatility Values



Discussions

- The volatility tends in all cases to **decrease** while approaching the election day
- At the beginning of the election period people could have been **more open** to new information and hence there was more **uncertainty** about the voting preferences.
- While getting closer to the election day the preferences of the voters **became clearer** and hence the change in the opinion pattern was reduced

Conclusions

- The predictions of future outcomes in the long-term using ARMA models are **limited** because of the **volatility**
- The presence of volatility in our return transformed time series suggests that the past **conditional variance** can be modeled using **GARCH** models
- Calm and volatile periods could be **identified** and **predicted**
- Forecasting the volatility could be used as a **measure of risk** in public opinion as it is used in **finance**
- As better NLP methods for **opinion mining** are developed, opinion time series created with those methods will reflect in a **better manner** the **opinion dynamics** of the population

Questions?

Thanks for your Attention!

Acknowledgments

- CONICYT Grant for Master studies in Chile
- Fondef "Observatorios Escalables de la Web en Tiempo Real" D09I1185
- Yahoo! Research Latin America

References I

- Akcora, C. G., Bayir, M. A., Demirbas, M., and Ferhatosmanoglu, H. (2010).
  Identifying breakpoints in public opinion.
  In *Proceedings of the First Workshop on Social Media Analytics*, SOMA '10, pages 62–66, New York, NY, USA. ACM.
- Asur, S. and Huberman, B. A. (2010).
  Predicting the Future with Social Media.
  IEEE/WIC/ACM International Conference on Web Intelligence.
- Bollen, J., Mao, H., and Zeng, X.-J. (2011).
  Twitter mood predicts the stock market.
  J. Comput. Science, 2(1):1–8.
- Bollerslev, T. (1986).
  Generalized autoregressive conditional heteroskedasticity.
  Journal of Econometrics, 31(3):307–327.
- Box, G. E. P. and Jenkins, G. M. (1994).
  Time Series Analysis: Forecasting and Control.
  Prentice Hall PTR, Upper Saddle River, NJ, USA, 3rd edition.
- Cryer, J. D. and Chan, K.-S. (2009).
  Time Series Analysis: With Applications in R (Springer Texts in Statistics).
  Springer, 2nd edition.

References II

Engle, R. F. (1982).
Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation.
Econometrica, 50(4):pp. 987–1007.

Gayo-Avello, D. (2011).
Don’t turn social media into another ’literary digest’ poll.
Commun. ACM, 54(10):121–128.

Mishne, G. and de Rijke, M. (2006).
Capturing global mood levels using blog posts.
In *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*. AAAI Press.

Mishne, G. and Glance, N. (2006).
Predicting movie sales from blogger sentiment.
In *AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW)*, pages 155–158.

O’Connor, B., Balasubramanyan, R., Routledge, B. R., and Smith, N. A. (2010).
From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series.
In *Proceedings of the International AAAI Conference on Weblogs and Social Media*.

Pang, B. and Lee, L. (2008).
Opinion mining and sentiment analysis.
Found. Trends Inf. Retr., 2:1–135.

Schuman, H. and Presser, S. (1996).
Questions and answers in attitude surveys: Experiments on question form, wording, and context.
Sage Publications, Inc.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., and Welpe, I. M. (2010).
Predicting elections with twitter: What 140 characters reveal about political sentiment.
In *ICWSM*. The AAAI Press.

Yu, S. and Kak, S. (2012).
A Survey of Prediction Using Social Media.
ArXiv e-prints.