Tracking Sentiment by Time Series Analysis

Anastasia Giachanou
Faculty of Informatics
Università della Svizzera italiana
Switzerland
anastasia.giachanou@usi.ch

Fabio Crestani
Faculty of Informatics
Università della Svizzera italiana
Switzerland
fabio.crestani@usi.ch

ABSTRACT

In recent years social media have emerged as popular platforms for people to share their thoughts and opinions on all kind of topics. Tracking opinion over time is a powerful tool that can be used for sentiment prediction or to detect the possible reasons of a sentiment change. Understanding topic and sentiment evolution allows enterprises or government to capture negative sentiment and act promptly. In this study, we explore conventional time series analysis methods and their applicability on topic and sentiment trend analysis. We use data collected from Twitter that span over nine months. Finally, we study the usability of outliers detection and different measures such as sentiment velocity and acceleration on the task of sentiment tracking.

CCS Concepts

•**Mathematics of computing** → **Time series analysis;**
•**Information systems** → **Sentiment analysis;**

Keywords

sentiment dynamics, sentiment change, time series analysis

1. INTRODUCTION

Recent years have seen the rapid growth of social media platforms that enable people to express their thoughts and perceptions on the web and share them with other users. Many people write their opinion about products, movies, people or events on microblogs, blogs, forums or review sites. The so-called User Generated Content (UGC) is a good source of user opinions and mining it can be very useful for a wide variety of applications that require understanding public opinion about a concept. For example, enterprises can capture negative or positive opinions of customers about products or about competitors and improve the quality of their services or products accordingly. It is also very important for government to understand the public opinion regarding different social issues and act promptly.

SIGIR '16, July 17-21, 2016, Pisa, Italy

DOI: http://dx.doi.org/10.1145/2911451.2914702

Sentiment analysis focuses on developing methods that can classify a text as expressing positive or negative sentiment [7, 3]. However, public opinion towards a specific topic changes over time. Time series models seem to be an appropriate tool for sentiment tracking. Leveraging time series analysis on public opinions can be useful to understand data and identify patterns, trends and seasonality. In addition, time series is helpful in identifying outliers which are likely to be related with events that caused a sentiment change. Finally, they can be useful for sentiment prediction and intervention analysis that capture future sentiment and detect the effect of a single event on the public opinion respectively.

The development of models that focus on sentiment dynamics has recently attracted much research interest. Bollen et al. [2] performed a sentiment analysis of all public tweets posted from the 1st of August to the 20th of December, 2008 and mapped every day to a six dimensional mood vector. In another study, O'Connor et al. [6] investigated the relation between opinion expressed in tweets and public opinion obtained by polls. To this end, authors retrieved relevant tweets to some specific topics and then estimated the sentiment score of every day. An et al. [1] explored whether mining social media data can be used to yield insights on climate change sentiment whereas Nguyen et al. [4] tried to predict the sentiment change using a feature based method.

The aim of this preliminary study is to explore the usability of conventional time series methods on sentiment tracking and how they can be leveraged to address other challenging tasks such as detection of reasons of sentiment change. We first explore and decompose the data we collected from Twitter about different topics with the aim to identify patterns, trends and seasonality. In addition, we try to detect the outliers and investigate their usability in detecting important events or reasons of change. This is very important as it is possible to use the data that are around those peaks to automatically detect the sentiment change reasons.

2. DYNAMIC SENTIMENT TRACKING

In this initial study we explore conventional time series analysis methods on data collected from Twitter. The study mainly focuses on presenting the preliminary steps that are critical to track sentiment from social media.

Frequency Analysis. The initial step in time series analysis is to explore the data and observe how the frequencies change. Let $X = (x_1, x_2, .., x_n)$ be a set of data observed at consecutive and equal time partitions denoted as $\{t_1, t_2, ..., t_n\}$. Based on this, we can observe how the number of tweets about a topic z changes per day, denoted as $N_t(z)$. This time series is an indicator of the popularity of the topic independent from the sentiment that is expressed.

To explore sentiment trends we need to measure the frequencies of tweets that express a specific sentiment. For reasons of simplification, we only consider positive and negative sentiment. However, the approach can be also applied on sentiments that represent emotions such as love, anger, sadness etc. Let $N_t(z,s)$ be the number of tweets that express a sentiment s towards a specific topic z posted during a particular time period t and $N_t(z)$ the number of total tweets posted towards z at t. Then, we can define the ratio of tweets that share a common sentiment s as:

$$r_t(z,s) = \frac{N_t(z,s)}{N_t(z)}$$

Based on this, we can measure the *sentiment velocity* that represents the rate of sentiment change and is defined as: $Vel_t(z,s) = N_{t+1}(z,s) - N_t(z,s)$. Based on this we can measure *sentiment acceleration* that represents the rate of change of sentiment velocity at a particular time t. The sentiment acceleration of topic z and sentiment s is defined as: $Acc_t(z,s) = Vel_{t+1}(z,s) - Vel_t(z,s)$. Plotting sentiment velocity and acceleration is useful not only to observe how a specific sentiment changes but also to detect if there is any emerging sentiment. For example, a negative emerging sentiment about a specific topic means that the company should be alerted and act promptly.

Time Series Decomposition. To get better understanding of the data we further apply time series decomposition. According to decomposition that is of crucial importance for the subsequent analysis and modeling, time series data can be decomposed into three components: the trend (T_t), the seasonal (S_t) and the random (R_t). Based on this, we can define the time series X_t as a function of these components: $X_t = f(T_t, S_t, R_t)$.

The decomposition is such that the three components add up to the original time series. One well known decomposition method is the additive decomposition defined as: $y_t = T + S + R$. The trend component reflects the long-term increase or decrease in the data. Another way to find the trend is to smooth the data and remove any wide variation such as the seasonality. *Moving average* is one of the most well known smoothing techniques and can be very helpful to identify patterns and trends in time series because it evens out short term fluctuations and makes the trend more apparent. According to this approach, the value of data at time t is the unweighted mean of the data observed at the k previous time periods. This is defined as:

$$MA_t = \frac{x_{t-(k-1)} + ... + x_{t-1} + x_t}{k}$$

The seasonal component represents patterns that are repeated at fixed periods like days, weeks, months etc. *Seasonal adjustment* is a method for removing the seasonal component of a time series. This is useful to observe the data without the seasonal effects that may have an influence on them. One typical example is that users may tweet more at specific days or specific time. Seasonally adjusted data can be constructed as: $SeasAdj_t = X_t - S_t$. Finally, the random component represents noise in data and can be constructed by removing the trend and seasonal components as: $R_t = X_t - T_t - S_t$.

Outlier Detection. The last tool we explore is the detection of outliers and its usability on identifying the reasons of sentiment change. The outliers that are visually depicted as sudden peaks may be caused because of false measurements or because of some important events. We are mostly interested on those caused by some events. These important events not only influence the popularity of a topic but also the public sentiment towards the topic. Therefore, we believe that the data related with outliers may be useful in detecting possible reasons of a sentiment change.

In order to identify the outliers we use the following equation: $e_i = x_i - x'_i$ where x_i is the observation i and x'_i is the prediction of the observation i. In other words, this equation calculates the ordinary residuals for each observation. We use *LOESS* that is also known as locally weighted polynomial regression model and interquartile range to detect the residuals. Let Q_1 and Q_3 be the lower and upper quartiles respectively, then the outliers are the observations that are outside the following range:

$$[Q_1 - k * (Q_3 - Q_1), Q_3 + k * (Q_3 - Q_1)]$$

where k represents the span of the range and it is usually set between 1.5 and 3.0.

3. EXPERIMENTAL SETUP

A tweets' dataset that spans over several months was required for our study. However, due to the restriction in tweets' redistribution, a large part of tweets in existing collections is missing. Also, it is not possible to extend the existing collections and include data from additional months. Therefore, we started collecting data from Twitter since the 10th of April 2015. Due to the required preprocessing, in this study we use data[1] collected until the 31st of December 2015. To collect tweets we used a list of 70 topics from different domains including politics, TV series, cities, products etc. In our study, we focus on the following topics: *android lollipop* (346.714 tweets), *Michelle Obama* (1.076.732 tweets) and *Merkel* (1.369.756 tweets).

Preprocessing was performed on the data before the subsequent experiments which involved stop-word removal and stemming. For identifying the opinionated terms we use the AFINN Lexicon [5]. AFINN contains more than 2000 words each of which is assigned a valence from -5 to -1 for terms with a negative sentiment or from 1 to 5 for terms with a positive sentiment. The sentiment score of each tweet is a sum of the scores of its opinion words. We chose to use a simple approach to assign the sentiment score since our focus is not on sentiment analysis but on sentiment tracking. However, in future we plan to use a more complex method.

4. RESULTS

Figure 1 shows the number of total, positive and negative tweets published every day for the topic *android lollipop*. From this figure, we observe that the popularity of the topic is decreasing. It is also clear that there are some peaks that may be related to some important events. However, it may be also the case that these peaks are related to seasonal effects. To have better understanding of the data we need to isolate the different components.

Figure 2 shows the decomposition of the topic *android lollipop*. The decomposition confirms that the trend of the topic is decreasing. Also, there is a seasonal effect on the

[1]To get access to the collection, please contact the authors

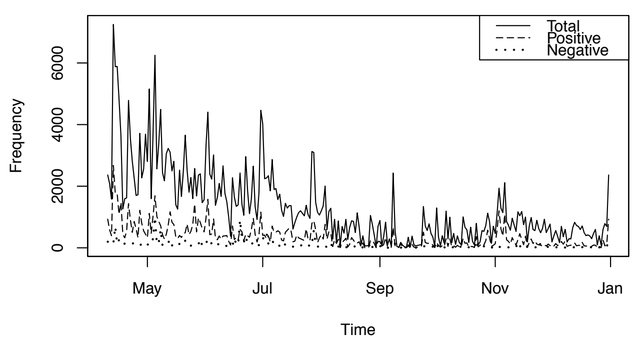

Figure 1: Number of total, positive and negative tweets of "android lollipop" topic per day

data. One possible explanation is that users tend to post more tweets on specific days of the week.

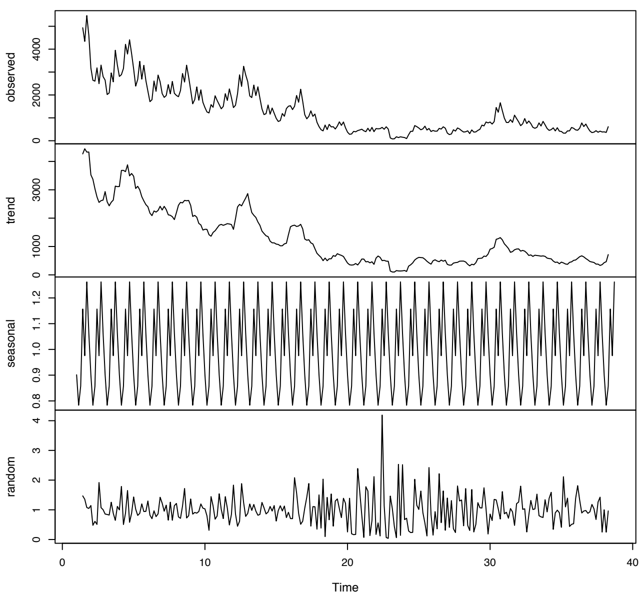

Figure 2: Decomposition of "android lollipop" topic

Another way to find the trend is to apply the moving average approach. Figure 3 shows the total, positive and negative number of tweets about the topic *Merkel* after applying the method of moving average using the past 10 days. In this case, we do not see any clear trend meaning that the popularity of this topic remains stable with a slight decreasing tendency. In its largest part, the positive and negative sentiment follow the trend of the topic's popularity. However, there are also cases that one sentiment seems to be stronger than the other. For example, in the end of November negative sentiment is stronger than positive. Looking through the collection and the news, we believe that this negative sentiment is related to the event mentioned in news as "Merkel under pressure from her own ahead of EU migration summit" that took place around the 27th of November.

The comparison between positive and negative tweets is better depicted on Figure 4 that shows the ratios of negative and positive over the total number of tweets collected for the topic *Merkel*. As expected, the sentiments tend to follow a contradictory behavior meaning that when positive sentiment trend is increasing the negative is decreasing and vice versa. For example, in the early of August the positive sentiment is clearly the dominant one whereas the negative has a very low ratio. This large difference implies that the positive sentiment was emerging at this specific time. Looking through the collection and the web, we discovered that on the 5th of August, "Merkel went to GlobalCitizen festival to show support for food security and nutrition for 500M people". We believe that this was the event that increased the positive sentiment. Knowing that this action was regarded positively by people is an important information for the press and information office of politicians.

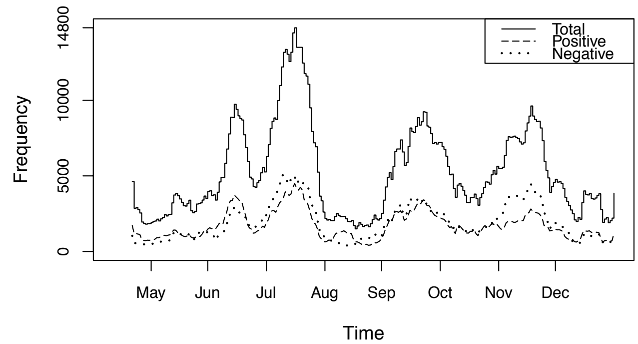

Figure 3: Smoothed data of "Merkel" topic

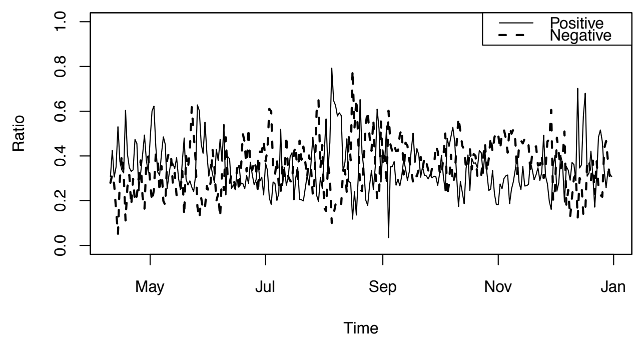

Figure 4: Ratio of positive and negative tweets about "Merkel" topic per day

Some important information is also reflected with sentiment velocity and acceleration that help us to understand how quickly a topic is gaining or losing preference. Figure 5 shows the positive and negative velocity and acceleration of the topic *Merkel*. Here, we observe that the negative sentiment grows faster in the middle of July whereas in August the positive sentiment has greater acceleration that lasts only few days.

Finally, we explore the usability of detecting outliers in finding reasons of sentiment change. Figure 6 shows the number of total, positive and negative tweets published every day for the topic *Michelle Obama* together with the detected peaks when retweets are considered and not. From

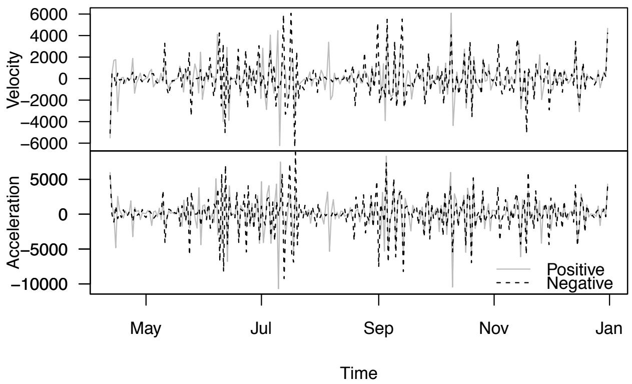

Figure 5: Velocity and acceleration of positive and negative tweets about "Merkel" topic per day

this figure we can observe the time when there was a sudden change in topic popularity or in sentiment towards the specific topic. Knowing the time of a potential peak is useful in identifying the events that caused those peaks. Apart from that, we observe that the peaks in topic's popularity, positive and negative peaks occur at different time periods. This implies that a sudden peak in a topic's popularity does not mean that there will be emerging sentiment. Another observation is that there are peaks that at different time points when we do and when we do not consider the retweets. This is an effect of the users's tendency to retweet some messages, an attitude that increases the popularity of the topic. Although retweets do not contain an explicit opinion, they can be seen as an endorsement of positive or negative sentiments. However, if few tweets are retweeted a lot, we may miss other events. Therefore, we believe that data should be explored in both scenarios. considering retweets and not.

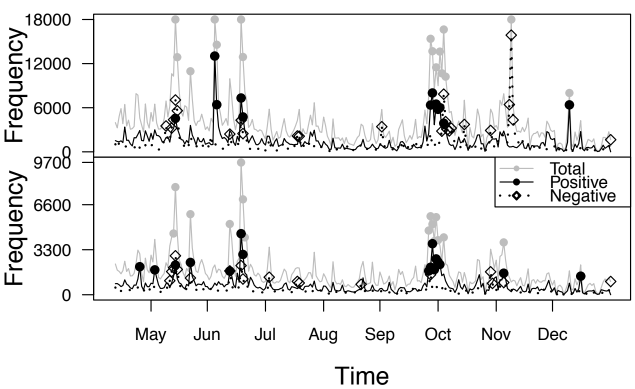

Figure 6: Number of total, positive and negative tweets about "Michelle Obama" when retweets are considered and when they are not per day

To understand the reasons of change on topic's sentiment, we need to look through the tweets that were published not only on the specific day but also a bit before and a bit after this day. Table 1 shows some manually detected tweets that we believe caused emerging sentiment. In some cases the tweets are not adequate to reveal the events but we need to look through the web. For example, on the 5th of May there was a great amount of tweets saying that "Michelle Obama ruined our lunch". Looking for more information on web, we found that this is related to the fact that Michele Obama changed schools' lunch program but students did not like the change. This example shows some of the challenges in addressing the task of detecting reasons of sentiment change.

Table 1: Dates and reason of emerging sentiment

Day	Emerging sentiment	Potential reason
5 May	negative	Michelle Obama ruined our lunch
12 May	negative	Michelle Obama gave a talk and mentions racism
2 June	positive	Michelle Obama responds to Kanye!
26 Sep	positive	Barack Obama & Michelle Obama shut the Internet down with class
16 June	negative and positive	First Lady Michelle Obama Meets With Prince Harry Over Tea
28 Oct	negative	Prince Harry and Michelle Obama to Meet With Wounded Vets

5. CONCLUSIONS AND FUTURE WORK

In this study we explored the usability of time series methods on sentiment tracking. We plotted frequencies and decomposed data from Twitter to identify patterns, trends and seasonality. Also, we tried to identify peaks in topic's popularity and in sentiment and we investigated their usability in detecting important events or the reasons of sentiment change. We believe that this is a good start for the development of methods that could automatically detect a sentiment change and the reasons that caused it. In future we plan to explore these problems more thoroughly.

Acknowledgments

This research was partially funded by the Swiss National Science Foundation (SNSF) under the project OpiTrack.

6. REFERENCES

[1] X. An, R. A. Ganguly, Y. Fang, B. S. Scyphers, M. A. Hunter, and G. J. Dy. Tracking Climate Change Opinions from Twitter Data. In *KDD '14*, 2014.

[2] J. Bollen and A. Pepe. Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. In *ICWSM '11*, pages 450–453, 2011.

[3] A. Giachanou and F. Crestani. Like it or not: A survey of twitter sentiment analysis methods. *ACM Computing Surveys*, in press.

[4] L. T. Nguyen, P. Wu, W. Chan, W. Peng, and Y. Zhang. Predicting Collective Sentiment Dynamics from Time-series Social Media. In *WSDOM'12*, 2012.

[5] F. Å. Nielsen. A new ANEW: Evaluation of a word list for sentiment analysis of microblogs. In *ESWC'11 Workshop on 'Making Sense of Microposts': Big things come in small packages*, pages 93–98, 2011.

[6] B. O'Connor, R. Balasubramanyan, B. Routledge, and N. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *ICWSM '10*, 2010.

[7] B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135, 2008.