Effects of the Recession on Public Mood in the UK

Thomas Lansdall-Welfare
Intelligent Systems Laboratory
University of Bristol
Bristol
United Kingdom
Thomas.Lansdall-Welfare@bristol.ac.uk

Vasileios Lampos
Intelligent Systems Laboratory
University of Bristol
Bristol
United Kingdom
Bill.Lampos@bristol.ac.uk

Nello Cristianini
Intelligent Systems Laboratory
University of Bristol
Bristol
United Kingdom
Nello.Cristianini@bristol.ac.uk

ABSTRACT

Large scale analysis of social media content allows for real time discovery of macro-scale patterns in public opinion and sentiment. In this paper we analyse a collection of 484 million tweets generated by more than 9.8 million users from the United Kingdom over the past 31 months, a period marked by economic downturn and some social tensions. Our findings, besides corroborating our choice of method for the detection of public mood, also present intriguing patterns that can be explained in terms of events and social changes. On the one hand, the time series we obtain show that periodic events such as Christmas and Halloween evoke similar mood patterns every year. On the other hand, we see that a significant increase in negative mood indicators coincide with the announcement of the cuts to public spending by the government, and that this effect is still lasting. We also detect events such as the riots of summer 2011, as well as a possible calming effect coinciding with the run up to the royal wedding.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*Text analysis*; G.3 [**Probability and Statistics**]: Statistical Computing; I.5.4 [**Pattern Recognition**]: Applications—*Text processing*; J.4 [**Social and Behavioral Sciences**]: Sociology

General Terms

Experimentation, Measurement

Keywords

Event Detection, Mood Analysis, Sentiment Analysis, Social Media, Twitter

1. INTRODUCTION

Social media allows for the easy gathering of large amounts of data generated by the public while communicating with each other. This can give social scientists – among others – a tool to access public opinion and public sentiment, a task so far accomplished only by opinion polls. Twitter in particular gives researchers access to real-time data that is suitable for the analysis of public sentiment on a large scale. This type of analysis is of interest because it avoids self-reporting and opinion-polling, and therefore opens the possibility to access much larger populations. On the flip side, this can only be accomplished with text mining technologies. We make use of standard tools for mood detection [9] and we apply them to a dataset of 484,033,446 tweets collected from July 2009 to January 2012 in the United Kingdom. This period has been marked by protracted economic downturn, and some social tensions.

Our goal was to see if the effects of social events can be seen in the contents of Twitter, and to speculate if some of them could even be predicted. The first part of our analysis provides a sanity check, in that it corroborates our assumption that word-counting methods can provide a reasonable approach to sentiment or mood analysis. While this approach is standard in many applications [3, 4], we felt that a sanity check in the domain of mood detection via Twitter was necessary. By making use of lists of words that are correlated with the sentiments of Joy, Fear, Anger and Sadness [9], we observe that periodic events such as Christmas, Valentine and Halloween evoke the same response in the population, year after year.

The second part of the analysis focuses on a visible change-point occurring in October 2010, when the government announced cuts to public spending, testing its statistical significance. We can show that the change point is real, and that its effects can still be observed. In other words, public mood still has not recovered from that announcement. The same testing technique shows another important period, that of Summer 2011, when riots broke out in various UK cities, leading to looting and even loss of life. Our method seems to suggest that an increase in public anger preceded – and not followed – those events. While we leave the interpretation of these findings to social scientists, we also observe how the period preceding the royal wedding seems to be marked by a lowered incidence of anger and fear, which start rising soon after that. Of course, other events also happened in early May 2011, so they may also be responsible for that increase.

A description of our data set can be found in Section 2. Section 3 details the occurrence of periodic events, with the detection of significant events covered in Section 4. Finally, we conclude the paper in Section 5.

2. DESCRIPTION OF THE DATA

Simple in concept, Twitter provides users with a platform on which they can publish brief, textual communications of up to 140 characters that are publicly available and thus

WWW 2012 Companion, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1230-1/12/04.

easy to capture. The communications, or 'tweets', tend to be in-the-moment expressions of the user's current experiences.

Using the Twitter Search API, we collected approximately 484 million tweets over a period of 31 months between 1^{st} July 2009 and 19^{th} January 2012 by 9,812,618 users. For each of the 54 most populous urban centres in the UK, we periodically retrieved the 100 most recent tweets every $3-5$ minutes, geo-located to within a 10km range of an urban centre, without specifying any keywords or hashtags. The text of each tweet was stemmed by applying the Porter's Algorithm [8] before using a text analysis tool to assess the emotional valence. The emotional valence is calculated by counting the frequency of emotion-related words in each text published on a given day. The emotion-related words were organised into four lists: anger, fear, joy and sadness. These word lists were based on extracts taken *a priori* from the tool 'WordNet Affect' [9]. The lists were further processed via stemming and filtering in a way that only single words are kept in the end. After this pre-processing, the word lists contained 146 anger words, 92 fear words, 224 joy words and 115 sadness words.

We computed the score for each emotion by first scoring each word, in the respective list, as the fraction of tweets containing it on that day. We then averaged this quantity over all words linked to that particular emotion to obtain a score for each emotion on that day. This method is based on the assumption that the average frequency of a word indicates its importance. Words with higher frequencies will have a larger impact on the mean value of each emotion. This gave us a daily representation of each emotion in the 933 day time period.

3. PERIODIC EVENTS

Our approach is corroborated by the fact that certain periodic events, such as Christmas and Halloween, evoke similar mood patterns every year. Figures 1–4 show the emotional valence time series for each of the emotions. It can be seen that certain periodic events have a large effect on the emotional valence for that day. For example, there is always a large spike of joy on Christmas Day followed by a slightly smaller spike on New Years. This pattern can also been seen for other holidays such as Valentine's Day and Easter. Of course, this effect is not just seen for joy, but is exhibited in other emotions, for example, Halloween gives rise to a jump in the sadness felt by the public.

It should be noted that we do not expect that a high frequency of the word 'happy' necessarily signifies happier mood in the population, as this can be due to expressions of greeting. In other words, our measures of mood are not perfect, but these effects could be filtered away by a more sophisticated tool designed to ignore expressions such as 'Happy New Year'. It is however a remarkable observation that certain days have reliably similar values in different years, suggesting that we have reduced statistical errors to a very low level.

Each of the emotional time series also exhibits a high level of autocorrelation between consecutive days (anger: 0.6592, fear: 0.5449, joy: 0.4869, sadness: 0.5545), along with a weekly pattern of autocorrelation (anger: 0.5465, fear: 0.4006, joy: 0.4499, sadness: 0.6115). These results are intuitively what one would expect, that similar moods are detected on consecutive days, and days of the week. The patterns found support findings from other studies [3] using similar, but independent methods to assess the emotional content of tweets.

4. SIGNIFICANT EVENTS

Significant events can also be detected from the emotional time series, highlighting events that changed the mood in the UK for a large period of time, rather than a single spike for a particular day. These events are detected by applying change point detection to the time series, looking for abrupt changes in the mean of the emotion valence.

Abrupt changes in the mean were found by using a 100-day moving window (50 days either side) to calculate the difference in mean before and after each day. Figure 5 shows the difference in mean for each day, where peaks can be interpreted as days which had the largest increasing effect on the next 50 days (and potentially for longer) in comparison to the previous 50 days. This measure is used to show the rate of mood change, and can informally be thought of as an indication of the derivative of the overall mood level. Using this information, we were able to identify two significant events which we felt warranted further investigation and therefore in this paper we have chosen to focus on them and their affect on the public mood in the UK, namely: 20^{th} October 2010 – UK Budget Cuts Announced, and 6^{th} August 2011 – UK Summer Riots.

UK Budget Cuts Announced

On 20^{th} October 2010 the UK government announced severe cuts to the budget in order to reduce with the governmental deficit and deal with the ongoing recession. The budget cut announcement did not come as a surprise to the public, with extensive news media coverage of the upcoming event. However the lasting effect of the announcement does appear to have changed the public mood from that day onwards. Figure 6 shows a large peak in the mean difference of fear and anger indicating that there is a change point in the time series for this date. Furthermore, using a cumulative summation of the mean we can see in Figure 7 that there is a change from days consistently being below the mean to suddenly being above the mean for anger and fear.

UK Summer Riots

On 6^{th} August 2011 riots broke out in major cities across the UK. There are a mixture of reasons given for the unrest from gang culture to the government cuts [10]. Figure 6 shows a build up of anger and fear in Twitter content, beginning in early spring 2011, until just before the riots happened. Interestingly, it appears that the occurrence of an event in early May 2011, possibly the royal wedding between Prince William and Kate Middleton, temporarily paused the ill feelings for a short time, before the rise eventually resumed. This can been seen in both Figure 6 as a period of unchanging difference in mean, and in Figure 7 as a decrease in the rate at which the cumulative sum is increasing. This trend points towards the wide media coverage given certain events, such as the royal wedding, altering the national public mood, an interesting and currently under-studied effect.

5. CONCLUSIONS

The use of social media mining to access information about the general population is not new. Our group, for example, has been working on the detection of Flu Rates [5], while

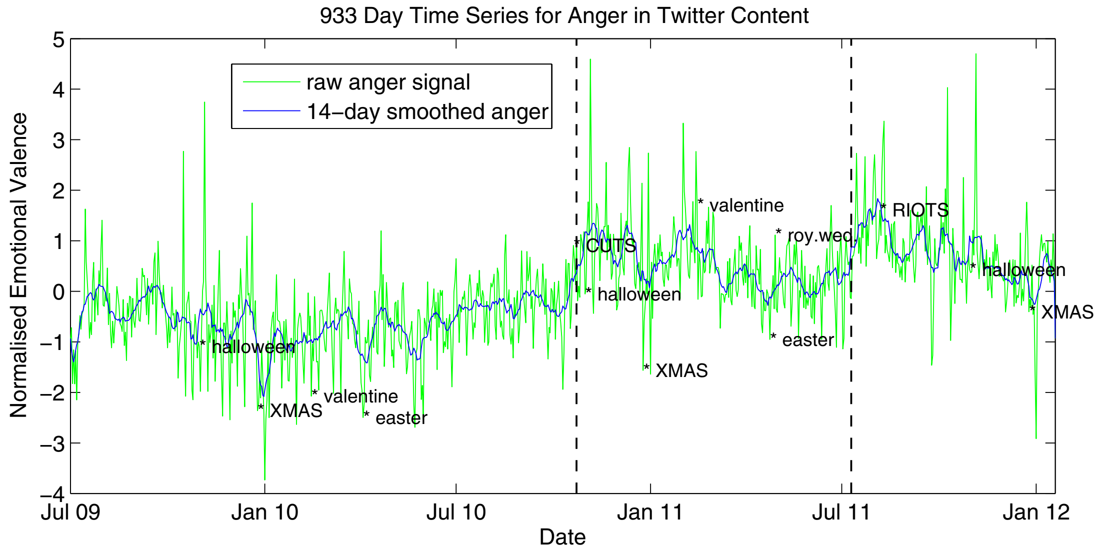

Figure 1: This plot shows the raw normalised emotional valence signal for anger along with a smoothed anger time series using a 14-day moving window for the 933 days in the time series. Dashed black lines indicate the days of highest mood change determined by difference in mean using a 100-day moving window.

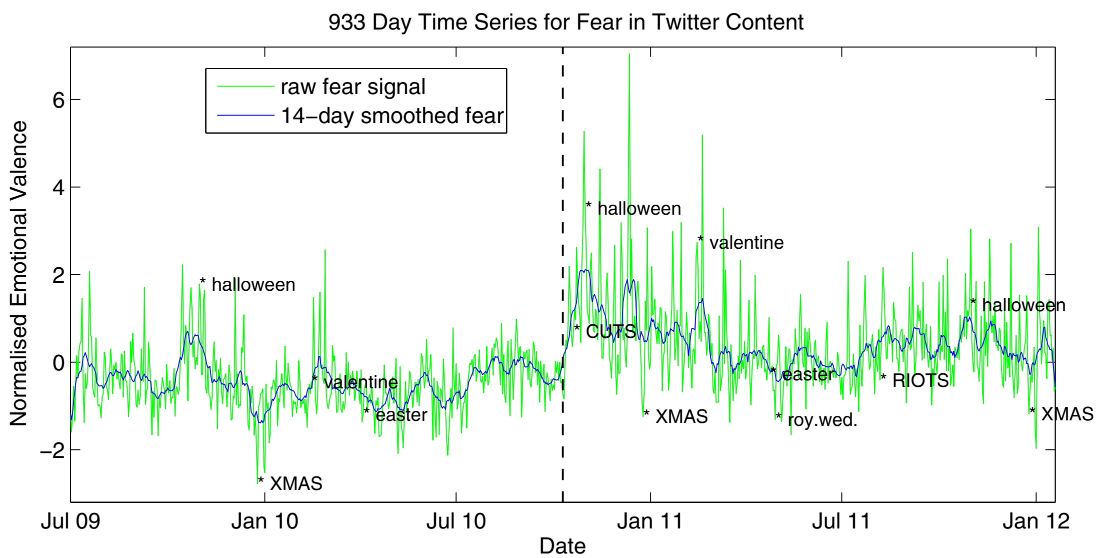

Figure 2: This plot shows the raw normalised emotional valence signal for fear along with a smoothed fear time series using a 14-day moving window for the 933 days in the time series. Dashed black line indicates the day of highest mood change determined by difference in mean using a 100-day moving window.

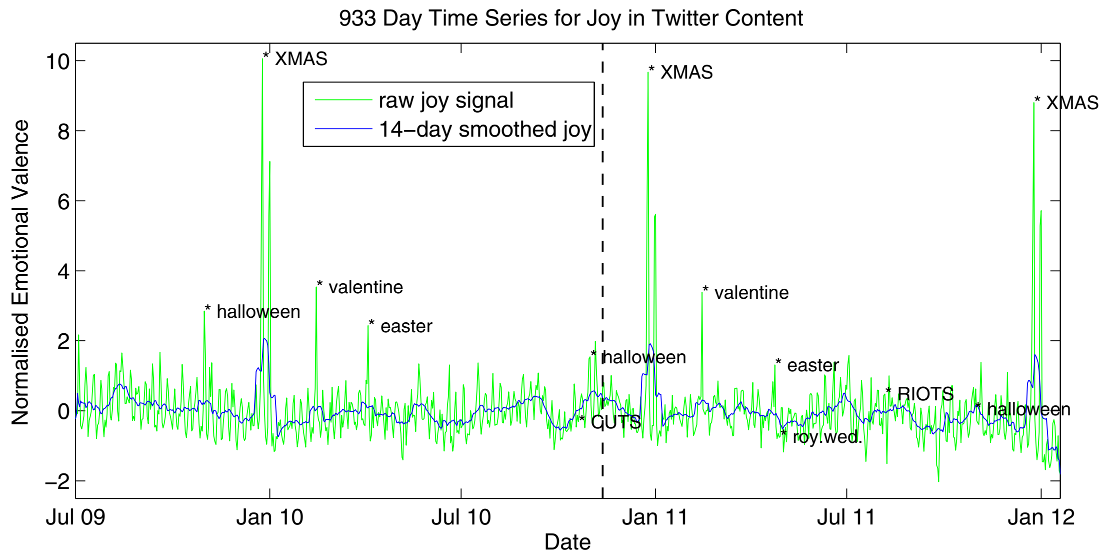

Figure 3: This plot shows the raw normalised emotional valence signal for joy along with a smoothed joy time series using a 14-day moving window for the 933 days in the time series. Dashed black line indicates the day of highest mood change determined by difference in mean using a 100-day moving window.

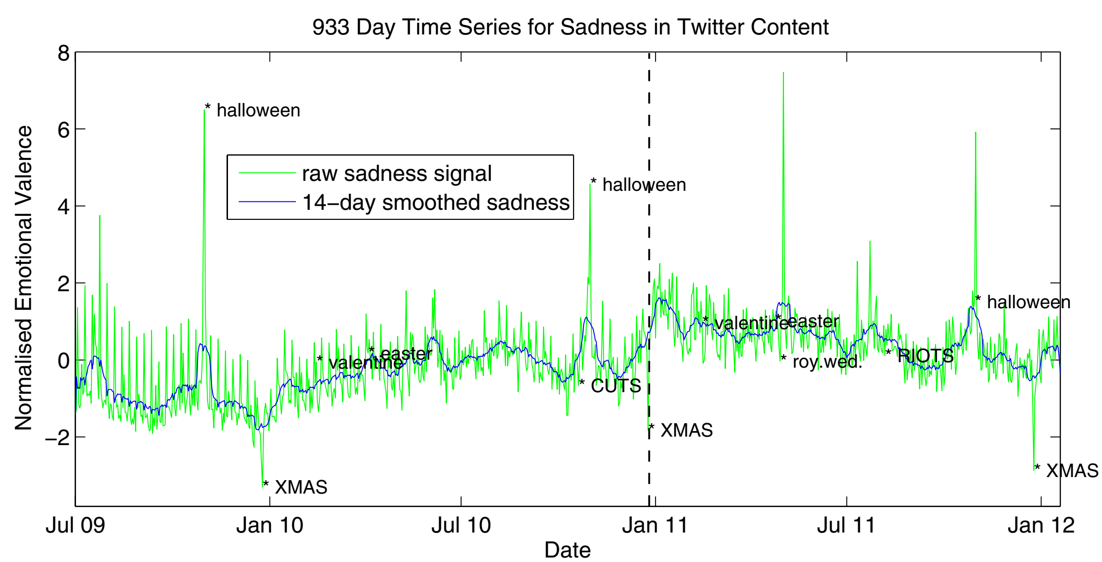

Figure 4: This plot shows the raw normalised emotional valence signal for sadness along with a smoothed sadness time series using a 14-day moving window for the 933 days in the time series. Dashed black line indicates the day of highest mood change determined by difference in mean using a 100-day moving window.

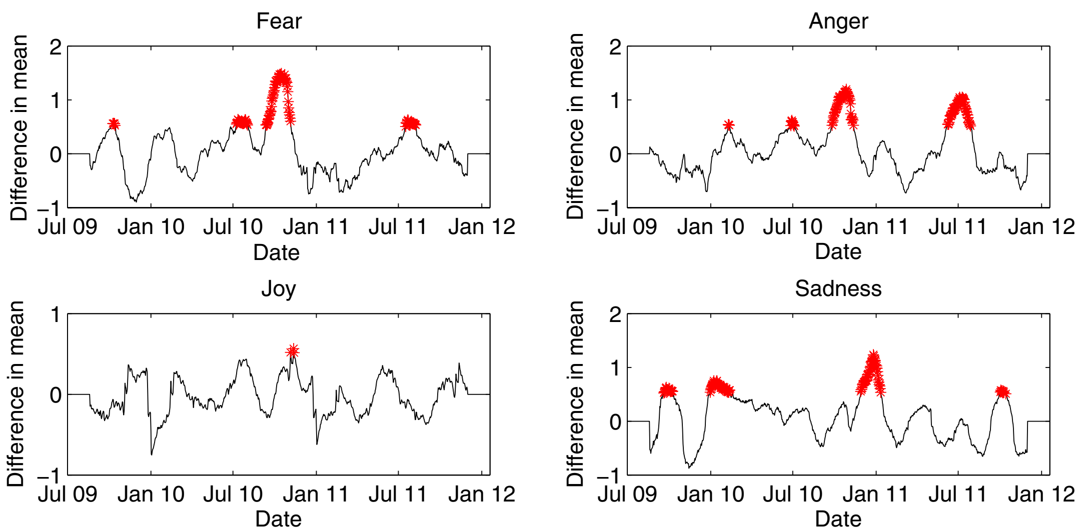

Figure 5: This plot shows the rate of mood change using the difference in mean for the 50 days before and after an event. Significant days ($p < 0.005$ for $t = 10,000$) are represented by red asterisks. Peaks in the graph indicate where there has been an increase in mood change for that emotion, while troughs indicate a decrease in mood change.

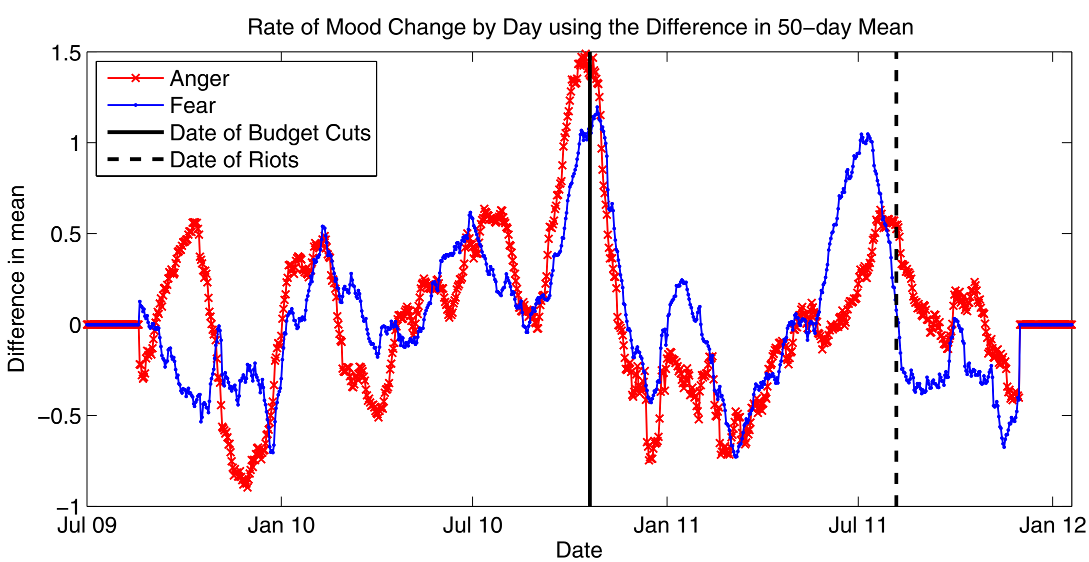

Figure 6: This plot shows the rate of mood change using the difference in mean for the 50 days before and after an event for anger and fear, with the dates for the UK Budget Cuts and the UK Summer Riots highlighted with vertical black lines. Peaks in the graph indicate where there has been an increase in mood change for that emotion, while troughs indicate a decrease in mood change.

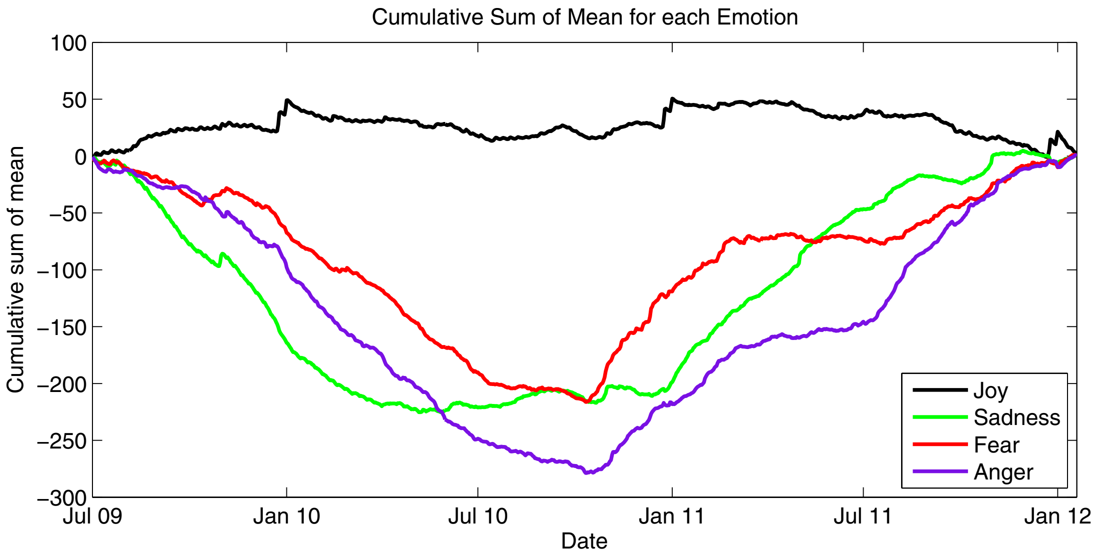

Figure 7: This plot shows the cumulative sum of the mean for each emotion over the time series. An increase indicates a period when the values are above the mean of the time series, while a decrease indicates a period when the values are below the mean.

other groups have focused on mood and happiness, based on similar data [3, 4]. While we use independent data and tools, it is reassuring that similar methods give similar results, in such a new field [1, 2, 3, 4, 6, 7]. This approach to measuring the state of a population holds great promise for social scientists, epidemiologists, perhaps also anthropologists. Here, we have demonstrated how a simple experiment on a large amount of Twitter data can reveal an important shift in public sentiment corresponding with the effects of the recession and government policy in the United Kingdom. Our group has also implemented a website showing the daily scores for each emotion for several regions in the UK[1]. Consequently, we have seen signals preceding the riots that could possibly be used as indicators. For this conclusion to be reached in more certain ways, we need more data, but we observe that a steady increase in anger was observed in the weeks preceding the riots. The involvement of social scientists would be crucial in determining causes and analysing reactions to events in a more thorough manner. Further work to investigate the effect of the traditional news media on public mood in social media could hold promise in finding how great an impact the news media has on public mood and how much is derived from personal experience.

6. ACKNOWLEDGEMENTS

Nello Cristianini is supported by the FP7 project 'Complacs'. All authors are supported by the Pascal2 Network of Excellence.

7. REFERENCES

[1] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2011.

[2] J. Bollen, A. Pepe, and H. Mao. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *CoRR, abs/0911.1583*, pages 1–10, 2009.

[3] P. Dodds, K. Harris, I. Kloumann, C. Bliss, and C. Danforth. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PLoS One*, 6(12), 2011.

[4] S. Golder and M. Macy. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051):1878–1881, 2011.

[5] V. Lampos, T. De Bie, and N. Cristianini. Flu detector-tracking epidemics on twitter. *Machine Learning and Knowledge Discovery in Databases*, pages 599–602, 2010.

[6] M. Mitrović, G. Paltoglou, and B. Tadić. Quantitative analysis of bloggers' collective behavior powered by emotions. *Journal of Statistical Mechanics: Theory and Experiment*, 2011:P02005, 2011.

[7] B. O'Connor, R. Balasubramanyan, B. Routledge, and N. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, pages 122–129, 2010.

[8] M. Porter. An algorithm for suffix stripping. *Program: electronic library and information systems*, 14(3):130–137, 1993.

[9] C. Strapparava and A. Valitutti. Wordnet-affect: an affective extension of wordnet. *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1413–1418, 2004.

[10] YouGov. Yougov/ the sun survey results. `http://cdn.yougov.com/today_uk_import/yg-archives-pol-sun-riots-100811.pdf`, August 2011.

[1] Mood of the Nation website available at http://geopatterns.enm.bris.ac.uk/mood/