CS 8803 Data Analytics for Well-being: Data Modeling I

Munmun De Choudhury
munmund@gatech.edu
Week 4 | February 1, 2016
Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter
Main Idea

- The article says the following about the web: “a collective, open recording of an enormous number of transactions, interactions, and expressions, marking a clear transition in our ability to quantitatively characterize, and thereby potentially understand, previously hidden as well as novel microscale mechanisms underlying sociotechnical systems”
- Method (a hedonometer): use of word frequency distributions combined with independently assessed numerical estimates of the ‘happiness’ of over 10,000 words obtained using Amazon’s Mechanical Turk
Contributions and Findings

- The paper explores happiness as a function of time, space, demographics, and network structure.
- Examine temporal variations in happiness including: the overall time series; regular cycles at the scale of days and weeks; time series for subsets of tweets containing specific keywords; and detailed comparisons between texts at the level of individual words.

![Graph A](image1)

![Graph B](image2)
How word coverage declines with increasing word rank is shown in Figs. 2D, 2E, and 2F, which together show the mutual agreement of our choice of a specific metric to pick up variations across texts (requiring higher frequency). We find that when we see a maximum coverage of approximately 50%. However, at the same time, we lose coverage of texts. The number of individual words for which we have evaluations decreases as word rank increases, with a 50% coverage at word rank 1. We choose a pairwise comparison of all time series using Pearson’s correlation coefficient, which shows a strong mutual agreement. We choose a specific happiness score, Eq. (1), as a weighted average of subsets of 10,222 individual word happiness assessments on a 1 to 9 scale, obtained from labMT 1.0 word list (Data Set S1) as a function of word rank.

We support the robustness of our choice with evidence that provides a test for robustness and tunability of our text-based hedonometer. Figure 2 demonstrates the distribution of average happiness of individual words. For different metrics, each time series is generated by omitting words with different word ranks, from 0 to 9, resulting in a suitable compromise. For 0 ≤ havg < 1, we are also safely above the transition, and for 0 ≤ havg > 1, we see that as the number of stop words increases, so does the variability of the time series.

In plot C, we see the percentage of the Twitter data set covered by each word (which uses plot B). In plot E, we show the number of unique words left in our list, accounting for word frequency; for labMT 1.0 word list (Data Set S1) as a function of havg, we have analysed such as blogs, books, and State of the Union Addresses [23]. This discrepancy in total coverage remains. The fraction of the Twitter corpus covered by these 3,686 word is approximately 23% (Fig. 2E). By comparison, the ANEW study’s 1,034 words collective-ly cover only 3.7% of the corpus, typical of other texts. For 0 ≤ havg = 1, our metric uses 22.7% of all words. Lastly, in plot F, we show time series of average happiness for Twitter, binned by day, produced by di...
Bailout of the U.S. financial system: 

\[ T_{\text{ref}}: \] 7 days before and after \( (h_{\text{avg}}=6.00) \)

\[ T_{\text{comp}}: \] Monday, 2008/09/29 \( (h_{\text{avg}}=5.95) \)

Royal Wedding of Prince William & Catherine Middleton

\[ T_{\text{ref}}: \] 7 days before and after \( (h_{\text{avg}}=5.98) \)

\[ T_{\text{comp}}: \] Friday, 2011/04/29 \( (h_{\text{avg}}=6.04) \)

Death of Osama Bin Laden:

\[ T_{\text{ref}}: \] 7 days before and after \( (h_{\text{avg}}=5.98) \)

\[ T_{\text{comp}}: \] Monday, 2011/05/02 \( (h_{\text{avg}}=5.89) \)
**B. Word Shift Analysis**

To give a deeper sense of the underlying moods reflecting the happiness cycle, we performed a word shift analysis on the diaries. Words that move against the general trend on Saturdays (average happiness $h_{avg}$=6.03) are:

- love +
- no −
- haha +↑
- party +↑
- new +↑
- weekend +↑
- not →
- happy +↑
- dont −
- bored −
- drunk −
- fight −
- miss −
- hangover +↑
- free −
- court −
- nice +↑
- won +↑
- school +↑
- movies +↑

The period 5–6 am marks 'biological midnight' with the lowest happiness. An evening low is consistent with the下班 (5–6 pm). An afternoon low is consistent with the observed happiness cycle. Fig. 11 shows the normalized frequencies for five example probe words.

**VI. DAILY CYCLE**

The balance plot (bottom right inset) shows that 5 to 10 am is the happiest hour of the day. As shown in Fig. 10, the happiest hour of the day is 5 to 6 am, after which we see a steep decline until midday. The happiness peak is followed by a more gradual descent to the on-average level of 10 to 11 pm, and then a return to the daily peak at 12 midnight.

The insets of Fig. 8 provide further insight and information due to excessive drinking. Although Saturdays may be on average happier than other days, this is not always the case. Thus while Saturdays may be on average happier than other days, this is not always the case. Days given equal weighting with outlier dates removed. See Fig. S5 in Supplementary Information for word shifts based on alternate methods of creating word frequency distributions of users' expressed happiness.

The insets of Fig. 8 provide further insight and information due to excessive drinking. Although Saturdays may be on average happier than other days, this is not always the case. Days given equal weighting with outlier dates removed. See Fig. S5 in Supplementary Information for word shifts based on alternate methods of creating word frequency distributions of users' expressed happiness.
Modeling Public Mood and Emotion: Twitter Sentiment and Socioeconomic Phenomena – (Bollen, Pepe, Mao, 2010)

- Examine how Twitter moods reflect social, political, and economic events
- Use POMS (profile of mood states) for detecting mood-indicative twitter posts.
  - POMS dimensions: tension, depression, anger, vigor, fatigue and confusion
- Investigate how a six vector representation of moods deviates during different big scope events.
- High stress/tension during elections; excitement/vigor during thanksgiving.
Our investigation of the produced public mood time series reveals a high level of sentiment fluctuation during the following events:

1. a timeline of the most important social, cultural, political and economic events;
2. Variance normalized: a 153 day, 6-dimensional time series for each of the POMS mood dimensions, z-score normalization (right);
3. the time series extracted from our collection of tweets for the period under study:
   - DJIA-I: August 1 to 24
   - DJIA-II: August 25 to September 1
   - DJIA-III: September 2 to 24
   - DJIA-IV: September 25 to October 17
4. General correlation drivers versus public sentiment:
   - positive spike in Vigour of the entire period we study, i.e. 0 deviation for Vigour and -2 to +2 standard deviations for the other dimensions.
   - Against the backdrop of the week- or month-long patterns in socio-economic indicators.

Figure 4: Raw POMS Confusion scores (left) vs. sparklines shown in Fig. 6 bear this out. All mood dimensions are expressed in standard deviation units: +2 standard deviations for Vigour and -2 to +2 standard deviations for the other dimensions.

Against the backdrop of the week- or month-long patterns in socio-economic indicators:

- DJIA underwent significant changes in value. We examine the DJIA and WTI trend lines; notable examples:
  - DJIA-I: August 1 to 24 - a drop in Fatigue. November 4 is characterized by a drop in Fatigue.
  - DJIA-II: August 25 to September 1 - a drop in Fatigue.
  - DJIA-III: September 2 to 24 - a drop in Fatigue.
  - DJIA-IV: September 25 to October 17 - a drop in Fatigue.
- WTI over the same period of time, namely August 1, 2008 to December 20, 2008.

Figure 5: The mood curves shown in Fig. 5 provide a fine-grained view of public mood changes in the three-day period surrounding the US presidential election on November 4, 2008. For each of the POMS mood dimensions, z-score normalization (right). The y-axis corresponds to mood z-scores, expressed in standard deviation units: +2 standard deviations for Vigour and -2 to +2 standard deviations for the other dimensions.

Figure 6: The mood curves shown in Fig. 6 provide a fine-grained view of public mood changes in the three-day period surrounding the US presidential election on November 4, 2008. For each of the POMS mood dimensions, z-score normalization (right). The y-axis corresponds to mood z-scores, expressed in standard deviation units: +2 standard deviations for Vigour and -2 to +2 standard deviations for the other dimensions.

Figure 7: The mood curves shown in Fig. 7 provide a fine-grained view of public mood changes in the three-day period surrounding the US presidential election on November 4, 2008. For each of the POMS mood dimensions, z-score normalization (right). The y-axis corresponds to mood z-scores, expressed in standard deviation units: +2 standard deviations for Vigour and -2 to +2 standard deviations for the other dimensions.

Figure 8: The mood curves shown in Fig. 8 provide a fine-grained view of public mood changes in the three-day period surrounding the US presidential election on November 4, 2008. For each of the POMS mood dimensions, z-score normalization (right). The y-axis corresponds to mood z-scores, expressed in standard deviation units: +2 standard deviations for Vigour and -2 to +2 standard deviations for the other dimensions.

2008 Presidential elections

Thanksgiving
Diurnal and Seasonal Mood Vary with Work, Sleep, and Day Length Across Diverse Cultures

• One of the early works examining relationship between social media mood and behavior and psychological theories.
• Identify daily and seasonal mood variations and relate it to work, sleep and daylight.
• Validate circadian rhythms in humans.
  • PA spike in the morning, NA increases as the day progresses
• Measure positive affect and negative affect based on the lexicon LIWC.
• PA and NA are not mirror images of each other.
poral granularity and is vulnerable to memory er-

trospective self-reports, a method that limits tem-

ey they sleep. Further, these studies typically rely on

9

ror and experimenter demand effects. Researchers

8

poral granularity and is vulnerable to memory er-

necessarily representative of the larger population

have also found that NA is not subject to diurnal

Afternoon (B) and evening (E) peaks at noon and evening (A), who tend to be active at different times

cycles), who tend to be active at different times

Students are exposed to varying academic

E

A

R

G

U

L

P

M

ving site that records brief, time-stamped public

online sources such as Twitter, a microblog-

argeting patterns.

LIWC, a prom-

guistic Inquiry and Word Count (LIWC), a prom-

of Twitter users worldwide, allowing cross-societal

al in scale (I), making it possible to obtain pre-

transcribed daily speech.

change, the LIWC lexicon was designed to analyze diverse genres

4.

petic Inquiry and Word Count (LIWC), a prom-

the globe. We measured PA and NA using Lin-

11

specific issues and political candidates varied

That is now changing. Data from increasingly

2.

using data from Twitter, we found that for all emotional expression words, LIWC

an in situ real-time hourly ob-

7.

that in previous research that has consistently shown

1979
Is measurement of positive and negative emotion sufficient for assessing the well-being of populations?
The hedenometer algorithm uses ratings from Amazon’s Mechanical Turk on words obtained from music lyrics, Twitter, NY Times and Google Books. Are there limitations to this rating gathering approach?
Twitter is used by millions, but could it also have bias?
Dictionary approach of mood detection: what is its limitation?
True emotion versus displayed emotion on social media: how would you tackle this issue?
People use social media for all kinds of reasons and purposes. Would that affect the moods they express?
Would “self-presentation”, “social comparison” or identity impact the kinds of moods shared?
Can social media manifested emotion have a cultural, demographic, or geographical bias?
Could the moods of certain Twitter users be more “important” than others? (Hint: influencers and contagion)
What are some of the other aspects, not considered in the papers, they may impact mood? (Hint: Aristotle said: “man is a social animal”)

One possible application is to study Twitter moods during important events, and how they impact each other. However, can public displays of mood from others impact our opinions?