

Measuring Changes in Twitter User Happiness in Response to Catastrophic Events

Eujiin Kang
Department Of
Computer Science
University of Iowa
Grinnell, Iowa
euijkang@uiowa.edu

Mingi Lee
Department Of
Computer Science
University of Iowa
Grinnell, Iowa
mingi-lee@uiowa.edu

Dylan Laurianti
Department Of
Computer Science
University of Iowa
Grinnell, Iowa
dlaurianti@uiowa.edu

Sergio Martelo
Department Of Computer
Science
University of Iowa
Grinnell, Iowa
sergio-martelo@uiowa.edu

Introduction

In this age of social media, posts have become an off-the-cuff and immediate way for people all around the globe to share opinions, ideas, and more. The advent of social media as mass opinion-sharing platforms has meant that every minute of every day, tens of thousands of posts are being uploaded by people all over the world. The combination of the nature of posts and this world-wide propagation of social media has given rise to an interesting opportunity: mass sentiment analysis of posts has now become a viable way to gauge a population's emotions around important topics or events. In this paper, we use social media and happiness analysis as tools to analyze how populations react to catastrophic events, like terrorist attacks or natural disasters. More specifically, we analyze the patterns and shifts in happiness before and after said events to analyze their impact on a population.

Due to the sheer amount of social media platforms and social media posts, we have decided to focus on posts made on Twitter. Also, we have decided to focus on catastrophic events that were highly publicized and garnered media attention; This approach ensures that we will have enough data for analysis. We further constrain our data to events that occurred from March 2009 to March 2010 in the U.S. due to data availability.

Another goal of this paper is to find whether there are differences in how a population responds to natural disasters and to man-made disasters, like terrorist attacks. To that end, our set of 4 events includes two man-made disasters and two natural disasters in th.

The events we are focusing on are the following: the flooding of Southeast U.S. in September 2009, the severe winter storm in December 2009, Appomattox shootings that killed 8 people in January 2010, and Fort Hood shooting that killed 13 people in November 2009. All of these events garnered worldwide attention when they occurred and mark some of the darkest moments of 2009 - 2010.

We are interested in performing mass happiness analysis on posts from the regions in which these events occurred and from other regions in order to see how the overall happiness of these populations was affected by these catastrophic events and how/if populations differ in their response to these catastrophic events. We strive to find recurring temporal patterns when it comes to how communities react to these events. We also expect to see a difference in the shift of happiness that arises from natural disasters and man-made disasters. We expect to see a down-turn in country-wide and local happiness, however we also strive to find any communities where the events might have had the opposite effect, like white supremacist communities, however due to the insular nature of such communities this might not be possible.

As a second objective of this paper, we analyze the impact of 4 major positive events on local and country-wide happiness. We perform this analysis in order to compare the impact of positive and negative events on local and global populations. These events are Christmas/Holiday season 2009, the October 2009 NASA launching Ares-I X Mission, the Fourth of July 2009 Statue of Liberty reopening, and the 2010 Winter Olympics. We expect to see negative events having more long-lasting and more significant effects than positive events, meaning that communities that

experience a negative event take on average a longer time to return to baseline than communities that experience a positive event.

To summarize, the questions we will attempt to answer in this work are the following: Are there significant differences between the effects of man-made disasters and natural disasters on user happiness on Twitter? How does user happiness on Twitter change in response to positive events? And, how do the effects of positive events compare to the effects of negative events on user happiness on Twitter?

Related Research

The sub-areas we will cover include data processing to extract useful information from the data that are created by web mining. The papers that we have cited give us a few examples of how sentiment was derived and then analyzed from a social media dataset. We adapt their methodology to the Twitter dataset to get a clear data-based picture of the effect of these specific traumatic news events on the Twitter user happiness.

Most methods of population happiness evaluation rely on self-reported assessments and are subject to our memory and perceptual biases [2]. These methods are costly, fundamentally invasive, and introduce misreporting. This gives us strong motivation to develop methods of happiness analysis which are able to do so remotely, or without direct engagement with members of the population. A clear route to this is to use textual information that individuals have already made public and to derive their sentiment from that freely accessible data. The most basic method for obtaining happiness from text sources is to use scores assigned to individual words, for example this paper uses scores of good-bad, active-passive, and strong-weak on a scale of 1-9 by 0.5 increments, then computing a happiness valence for each word and using a weighted average. This paper overall gives a simple, intuitive and effective method for computing the sentiment of larger texts, but we will need to expand this method with those from other texts that allow us to get more information from smaller amounts of text data per post by recognizing context.

Our research investigates whether a certain event can change the happiness of social media users. In this investigation, knowing whether happiness can spread among groups of people can give insight into the research and enable us to make predictions about the result. The clusters of people who are happy are clearly visible in the social network [4]. Moreover, they state that people who are surrounded by happy people tend to become happier in the future. Based upon the result of this paper, we expect that the catastrophic

events will have an impact on the overall happiness of groups on Twitter as the users are connected to each other.

The idea of relating happiness, or sentiment in general, on a social media platform to real world events is not a new one. Analyzing Twitter sentiment using Profile of Mood States (which measures depression, anger, vigor, fatigue, and confusion) and comparing it against a timeline of major events, showed that there is indeed some correlation to major events and the overall Twitter sentiment - even if the sentiment shift tended to be delayed [5]. However, their study looked at a short time frame of 5 months, which limited them to observing many less impactful events; furthermore, their research was conducted at the infancy of social media, 2009, and the landscape of social media has changed greatly over the years.

Happiness analysis relates to disasters more generally than just the textual analysis/social media approach which we have focused on [8]. Focusing on deep-learning using images, we can see that a lot of the markers for sentiment and whether or not it is related to the given disaster will translate to textual analysis. This paper talks about how specific activities can be associated with different emotional states and we can use this to enhance the depth of our textual analysis of Twitter posts.

Methods

Dataset and Data Cleaning

We are using the dataset contained at [1]. The authors of this dataset were focused on investigating the possibility of using Twitter user's tweets and machine learning to predict a user's location. So, they collected U.S. user tweets and location data for about 9 million tweets and about 120 thousand unique users. The dataset spans from 03/01/2009 to 03/01/2010 and covers all of the continental United States. The results of their paper can be found at [9]. Their data was separated into 4 different datasets; A dataset of all of the tweets in their training set, a dataset of all of the tweets in their test set, a dataset of all of the users in the training set and their locations, and a dataset of all of the users in the test set and their locations.

Our first step was to join these four datasets into one dataset containing all of the tweets, their dates, and their locations. This proved more difficult than it sounds due to the strange formatting of the original datasets and due to the fact that some users had their location as a city name and some users had their locations as coordinates. We chose to convert all of the city names into their corresponding coordinates using the U.S. cities dataset found at [10]. Then, we

normalized the dates into the same time zone (UTC) to aid in our analysis.

Next, we focused on creating a pipeline to clean the tweets. The pipeline had the following steps: (1) remove any mentions of other Twitter users, (2) remove any links in the tweet, (3) extract and save hashtags, (4) remove extra whitespace and/or special characters, (5) turn all words into lowercase, and (6) remove any words not in the lexicon at [2]. After cleaning, we were left with a dataset containing a tweet's text, date, hashtags, cleaned text, and coordinates which we utilized for our analysis.

Scoring Tweets' Happiness

In order to score tweets based on their happiness, we used the hedonometer methodology and the lexicon created by Dodds et al. for [1]. This method basically involves two different ways of measuring happiness. The first is taking a simple weighted average of the words in a tweet, while ignoring all words that are not in the lexicon. The reason we ignore the non-lexicon words is that the lexicon contains happiness rankings of over 10,000 words that have been found to be a robust way of measuring happiness by the authors of [1], however we have no information for words not in the lexicon and, thus, we ignore them. The second method of measuring happiness described by Dodds et al. is "Ambient Happiness". This method involves finding tweets with keywords related to the event whose ambient happiness you are measuring, removing these keywords from the tweet, and then taking the weighted average of the happiness as before. In the analysis that follows, we used both methods of measuring happiness to garner insights from the data.

Analysis

We are mainly focused on analyzing the changes in U.S. population happiness before and after a positive or negative event. We calculate the happiness of a group by calculating the mean happiness of that group's tweets. To analyze the impact of an event on the population, we first partition our set of tweets into two groups; (1) A pre-group that includes all tweets posted before the date of the event in question and (2) a post-group that includes all the tweets posted after the date of the event in question. Then, we further filter these groups into three subgroups which overlap. These subgroups are (1) all of the tweets that happened in the 14-day period before/after the event, (2) all of the tweets that occurred in the 7-day period before/after the event, and (3) all of the tweets that occurred in the 3-day period before and after the event.

The rationale for subdividing the pre and post groups is that this allows us to analyze the impact of the event in three different time-frames. We keep our time-frames relatively small to mitigate the effects of the given event overlapping with other events that might affect the happiness of a population as well.

Then, we aggregate the happiness of each group by taking the mean happiness of all of the tweets it contains and calculate the difference in mean happiness of each group with its counterpart. So, for example, if we were to calculate the mean happiness of the 14-day pre-group, we would then calculate the difference between that group and the 14-day post group in order to quantify the effect of the event on population happiness in this 14-day period. To assess whether the difference in mean happiness between any group pair is significant we perform a Welch's t-Test and a common t-Test on the groups with a significance level of 5%. The reason we perform both a Welch's and common t-test is because, although our sample sizes are large enough to perform a common t-test without having to worry about problems with the normality of our distributions, we still want to make our analysis more robust by performing a Welch's t-test. This test does not make assumptions about the normality of our data or the equality of variances in the two groups, so it is more robust to potential quirks of our dataset.

We repeat this analysis for ambient happiness by filtering our dataset for tweets that include keywords related to the event in question as described in the section above.

Furthermore, we also analyzed location specific changes in happiness for negative events as these events are more localized than positive events and, thus, we expect to see differences in how specific locations react to the events. We also investigate possible differences in impact between man-made and natural disasters and overall time-trends in the happiness of our population.

Results

Data Overview

Data was collected from September 9th, 2006 to March 17th, 2010 (Figure 5a). The number posts per day contained in the dataset generally increased heavily overtime and was at the max rate of data collection at the end of the collection period. We observed an average hedonometer score of 4.10 across the data set of 1029647 posts, with standard deviation of 0.23, minimum of 2.54 and maximum of 5.92. One anomaly we observed in our data was strong clustering of tweets around certain scores. This is a result of twitter posts tending to be short, and

tweets containing only a few extremely common words and some urls or uncommon or misspelled words which were not frequently used enough to be given a rating (ex. tweet: <http://twitpic.com/r4qz0> I spyyyy @adammoyer). These posts were computed as identical by our hedonometer even though they contained different words, the words that were rated were identical. The score data per time period from the hedonometer drastically decreases in variance as expected with the increase in tweet collection (Figures 5b, 5c). Notably we see a significant shift downward for the final quarter of 2009, which may be driven by lasting international unrest relating to the worsening of the Financial Crisis and the terrorism conflict.

Effect of Negative Events on Population Happiness

We investigate the 4 negative events that we selected: the flooding of Southeast U.S, the North American Blizzard, Appomattox shooting, and Fort Hood shooting. Figure 6.a, 7.a, 8.a, and 9.a show the average happiness for tweets containing the keywords, tweets not containing the keywords, and all tweets. For all the plots, the average happiness for tweets containing the keywords is much lower than that of the other two groups, meaning that the tweets that contain keywords relevant to negative events have lower average happiness score. Figure 6.b, 7.b, 8.b, and 9.b tell us that the frequencies of the key words that we selected significantly increased after the event.

Analysis of the Effect of 2009 Southeastern United States Floods on Population Happiness

Table 5 shows the results of our analysis on the effect of the 2009 Southeastern United States floods, on the happiness of the population. The tables tell us that there is not a significant pattern in the difference in average happiness before and after the event for all three groups of timespans. We believe that these results come from the existence of confounding events. The third table shows the statistics for regions in the Southeastern United States. We observe that the difference in average happiness is the smallest in the 6-days period, which means Tweets generated in the Southeastern regions might have an impact on the change in the average happiness in the short time span. However, the sample size for groups is too small and the p-value is too high to draw a conclusion.

Analysis of the Effect of the 2009 North American Blizzard on Population Happiness

Table 6 shows the results of our analysis on the effect of 2009 Southeastern United States floods, on the

happiness of the population. The first and second table show an increase in the average happiness for the 6-day period. We believe that this is mainly because of confounding events like Christmas. The impact of Christmas on the user's happiness is highly significant so even the tweets showed higher happiness scores. The third table tells us that among the regions where the blizzard had a severe impact, we can see that the difference in average happiness decreased even with the 6-day period statistics. We believe that those regions had severe damage from the storm so even with the significant impact of Christmas, the average happiness had decreased. However, due to the small sample size and high p-value, this cannot be generalized to draw a significant statistical conclusion.

Analysis of the Effect of the 2009 Appomattox Shooting on Population Happiness

Table 7 shows the results of our analysis on the effect of 2009 Appomattox Shooting on the happiness of the population. The tables tell us that in the 6-day period, there are decreases in the average happiness for all three tables. However, the small sample size and p-value tell us that the decrease might not be stemming from this event.

Analysis of the Effect of the 2010 Fort Hood Shooting on Population Happiness

Table 8 shows the results of our analysis on the effect of 2009 Southeastern United States floods, on the happiness of the population. The tables tell us that there is an overall increase in the difference in average happiness before and after the event for all three groups. We believe that these results come from the existence of confounding events like Halloween.

Differences Between Man-Made and Natural Disasters

Finally, we analyze whether there were differences in how Twitter behaved when exposed to a Natural disaster vs a Man-Made Disaster. To analyze the differences, the sets of tweets filtered by location and keywords were utilized as they had higher relevance to the relation between happiness level changes and events occurring in specific areas. As stated in the previous sections, there is a lack of data for certain time frames of the disasters, which may have resulted in some miscalculations.

From the analysis done by a two sample t-Test over the aggregate pre/post happiness differences of man-made vs natural events, we saw that there were no real significant differences in the behavior of the users before and after man-made/natural disasters respectively. For all the time

frames we looked at (14 days before and after, 7 days before and after, and 3 days before and after), all of them had p-values ranging in the .7 ~ .8. As happiness' behavior seemed to have no real patterns in the previous analyses, the results were not unexpected.

Final Remarks on the Effects of Negative Events

From the analysis, we observe that the negative events seem to not have a huge impact on the overall happiness of Twitter Users. Even though the tweets that contain relevant keywords show relatively lower average happiness scores, the events do not seem to change the overall happiness dramatically. The ambient happiness and local ambient happiness data seem to be more affected by the event, but we cannot draw any conclusion due to the small sample size and high p-value. We believe that this is mainly because of the confounding events. For example, the 2010 Fort Hood shooting and 2009 North American Blizzard include big events like Halloween and Christmas, which makes the impact of those negative events relatively negligible.

Effect of Positive Events on Population Happiness

To study the effect of positive events on population happiness, we now turn the focus of our analysis to 4 clearly positive events that occurred during the time period of our data. Again, we use the hedonometer method developed by the authors of [2] to measure the happiness of a given tweet and we calculate the happiness of a group by calculating the mean happiness of that group's tweets. Figures 1, 2, and 3 above show time series highlighting the average happiness of tweets containing keywords related to the events whose effects we are studying, tweets not containing these keywords, and all tweets over the 14-day periods before and after the event occurred. These figures also include keyword term frequency data over the whole time-span of our dataset. One important thing to note for all these figures is that the variance of the average happiness for tweets containing the keywords is always much higher than that of the other two groups. This should not be taken to have any significant meaning, as we believe that this difference stems from the smaller sample sizes of the keyword groups, which is unavoidable given the nature of the sampling. What follows are our results for each positive event and our conclusions about the impact of positive events.

Analysis of the Effect of Christmas Day 2009 on Population Happiness

Tables 1.a and 1.b show the results of our analysis on the effect of Christmas Day 2009, i.e. December 25th, on the happiness of the population. As you can see, our analysis of all tweets yielded significant differences in mean happiness for all groups. However, we actually see decreases in mean happiness in the 14-day and 7-day periods, while mean happiness increases in the 3-day period after the event. Also, we observe that the decrease in happiness in the 7-day period is smaller than that of the 14-day period, meaning that as time goes on we expect to see mean happiness decrease. It seems like the underlying dynamic of this event is that it is correlated with a small boost in happiness in the overall population that is followed by a dip in happiness that actually takes mean happiness to below its pre-event levels.

In regards to the second table, we can see that ambient happiness tells a different story. As you can see, ambient happiness increases after the event for every pair of groups. This increase in ambient happiness is way larger in magnitude than any decrease or increase seen in the overall population. We also see that the ambient happiness decreases as we get further from the event, with the 3-day period group having a higher mean ambient happiness than the 7-day period group and so on. Again, differences in ambient happiness for this event are significant across all the groups.

Analysis of the Effect of Independence Day 2009 on Population Happiness

Tables 2.a and 2.b show the results of our analysis on the effect of Independence Day 2009, i.e. July 4th, on the happiness of the population. This analysis did not yield significant results. There seems to have been statistically insignificant decreases in ambient and overall happiness for almost all time periods, however overall it does not seem like the fourth of July 2009 had a measurable impact on happiness in the population.

Analysis of the Effect of October 2009 NASA Rocket Space Launch on Population Happiness

Tables 3.a and 3.b show the results of our analysis on the effect of Independence Day 2009, i.e. July 4th, on the happiness of the population. This analysis yielded interesting results, however we believe that some of the results, specifically the 14-day and 7-day post-groups, are confounded by the fact that this event occurred in late October right before Halloween. So, we can only yield insights from the 3-day group pair and the ambient happiness analysis.

From this analysis, we can see that there are significant increases in both ambient and overall happiness in the 3 days after the event when compared

to the 3 days before the event. However, we can also see that this event did not yield any significant changes in ambient happiness in the other groups. We believe that this is in part due to the nature of the event. NASA space launches are big events for enthusiasts but not as much for most people. Even so, an enthusiast's happiness might not be increased for too long by a launch due to the fact they occur relatively frequently and so they might not have as lasting of an effect as we would expect.

Analysis of the Effect of 2010 Winter Olympics Start on Population Happiness

Tables 4.a and 4.b show the results of our analysis on the effect of the start of the 2010 Winter Olympics, i.e. February 12th, on the happiness of the population. This analysis yielded similar results to our Christmas analysis. For overall happiness, we see an increase in the 3-day period groups and an increase in both of the other groups. Similarly, the decreases are of larger magnitude as the time periods get larger. One important thing to note is that, although all of the changes in overall happiness are significant, the increases or decreases in overall happiness do not seem to be as large as those of Christmas. That seems to indicate that Christmas is a more significant event than the Winter Olympics, which we believe is a true reflection of the reality of the preferences of the U.S. population.

For ambient happiness, we also see significant, and much larger, changes in mean happiness across the different group pairs. The increases in happiness also seem to get smaller and smaller as time goes on, which we speculate is just a reflection of interest and excitement around the event waning as time goes on. It is also important to note that this is the only non-US based event of all of our positive events, so it is interesting to see it have a greater effect on the U.S. population than some key U.S. national holidays.

Final Remarks on the Effects of Positive Events

From our analysis, we can draw a few broad conclusions about positive events that could be investigated further. First, it seems like single positive events can have an impact on the overall happiness of a population, however this impact is at best small and at worst insignificant. Also, we found that those changes in happiness that were significant seem to only affect the population for a short amount of time and sometimes even led to decreases in happiness in the long run. The changes in ambient happiness before and after a positive event do seem to be more pronounced than those of overall happiness, which is

intuitive in that we expect people who tweet about an event to be more invested in it and, thus, gain more happiness from it. These changes, although a little longer lasting, also seemed to be short-lived. Our conclusion on this topic is then that positive events, even significant ones, seem to have at most small short-lived effects on the U.S. population, whether this is due to humanity's tendency to focus on the negative and quickly forget the positive or due to the short attention span of the American populace is left up to interpretation.

Conclusions/Limitations

Overall, in this paper, we found out that negative events seem to not significantly contribute to changing the overall happiness levels on Twitter while positive events had minor effects. While the conclusion to draw from this seems to be that events in the world seem to have little impact on Twitter's overall happiness levels, when looking at the data we used in this paper, a different answer can be found.

Firstly, while we had more data than Bollen et al. [5], we were still considerably limited in what events we could analyze. Due to the constraints of the dataset, we were unable to escape the confines of the geographical region of the United States and the time frame of 2009-2010, leading us to observing less impactful events (though not insignificant either). In all four of our disasters, the number killed did not exceed 10s of people. For instance, if we had some more data, we would've been able to analyze the huge earthquake in Haiti in 2009 or the 9/11 terrorist attack's effects on Twitter's overall happiness. While the shootings and the floods were indeed disasters, they were small enough in scale where many people could not have heard of such events.

Thus, with this in mind, the conclusion we can now draw from our research becomes: small-scaled and/or non-catastrophic events seem to often be buried under the deluge of information that Twitter offers.

In the future, we would like to try and obtain a larger dataset so that we can analyze the differences between larger and smaller disasters as well as looking to see if large scaled disasters are able to have a larger impact on social media.

In terms of difficulties we've had during the project, we had significant troubles in correctly formatting the data as there were a variety of things we had to consider when tokenizing such corpora, such as misspellings, emojis, and more. Besides this, we also had some trouble when trying to determine keywords as the keywords could very easily become too powerful. For instance, if 'damage' was used as one of the keywords for the flood or the winter storm, many unrelated events

may end up skewing our average happiness rating, especially as ‘damage’ often has negative connotations.

Contributions

Mingi: Analyzed negative events and their impacts. Helped with finding the dataset and cleaning.

Sergio: Analyzed positive events and their impacts. Helped clean and wrangle data and score tweets.

Euijin: Analyzed differences between man-made and natural disasters in terms of impact. Helped score tweets and write conclusions.

Dylan: Analyzed overall trends in data. Helped write related research and conclusions.

REFERENCES

- [1] Z. Cheng, J. Caverlee, and K. Lee, “Twitter CIKM 2010,” Internet Archive, 11-May-2011. [Online]. Available: https://archive.org/details/twitter_cikm_2010. [Accessed: 26-Apr-2023].
- [2] Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM (2011) Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. PLoS ONE 6(12): e26752. doi:10.1371/journal.pone.0026752

- [3] Dodds PS, Danforth CM (2009) Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. Journal of Happiness Studies.
- [4] Fowler JH, Christakis NA (2008) Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. BMJ 337: article #2338.
- [5] Bollen J, Pepe A, Mao H (2011) Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM). Barcelona, Spain.
- [6] Mishne G, de Rijke M (2006) MoodViews: Tools for blog mood analysis. AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs.
- [7] B. Pang and L. Lee, “Opinion mining and sentiment analysis.” [Online]. Available: <https://www.cs.cornell.edu/home/llee/omsa/omsa-published.pdf>. [Accessed: 29-Mar-2023].
- [8] A. M. Sadiq, H. Ahn, and Y. B. Choi, “Human Sentiment and Activity Recognition in Disaster Situations Using Social Media Images Based on Deep Learning,” Sensors, 2022. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7763261/pdf/sensors-20-07115.pdf>. [Accessed: 29-Mar-2023].
- [9] Z. Cheng, J. Caverlee, and K. Lee, ‘You Are Where You Tweet: A Content-Based Approach to Geo-Locating Twitter Users’, in Proceedings of the 19th ACM International Conference on Information and Knowledge Management, Toronto, ON, Canada, 2010, pp. 759–7
- [10] “United States Cities Database,” *simplemaps*. [Online]. Available: <https://simplemaps.com/data/us-cities>. [Accessed: 26-Apr-2023].