Proceedings of the Ninth International AAAI Conference on Web and Social Media

It Is Not Only About Grievances: Emotional Dynamics in Social Media During the Brazilian Protests

Jean M. R. Costa *
Information Science
Cornell University
jmd487@cornell.edu

Rahmtin Rotabi *
Computer Science
Cornell University
rahmtin@cs.cornell.edu

Elizabeth L. Murnane
Information Science
Cornell University
elm236@cornell.edu

Tanzeem Choudhury
Information Science
Cornell University
tanzeem.choudhury@cornell.edu

Abstract

In the summer of 2013, Brazil experienced a period of conflict triggered by a series of protests. While the popular press covered the events, little empirical work has investigated how first-hand reporting of the protests occurred and evolved over social media and how such exposure in turn impacted the demonstrations themselves. In this study we examine over 42 million tweets shared during the three months of conflict in order to uncover patterns in online and offline protest-related activity as well as to explore relationships between language-use in tweets and the emotions and underlying motivations of protesters. Our findings show that peaks in Twitter activity coincide with days in which heavy protesting took place, that the words in tweets reflect emotional characteristics of protest-related events, and less expectedly, that these emotions convey both positive as well as negative sentiment.

Introduction

Social media has emerged as a powerful resource during periods of collective political action — both for the individuals involved in the movements as well as for observers aiming to better understand their dynamics. For instance, websites including Facebook and Twitter have been extensively used to facilitate mobilizations against autocratic regimes (e.g., in Egypt, Iran, and Tunisia) as well as during demonstrations in democratic countries (e.g., Austria, United Kingdom, United States). Such usage of social media to organize and express viewpoints additionally leaves behind a record of information that enables researchers to explore questions surrounding human nature in newfound ways, at broader scales, and in real-time.

The question of what drives people to protest has long intrigued social scientists, and emotional dynamics in particular are a crucial facet in understanding participation. Throughout the lifecycle of a protest, emotions are a key driver behind why individuals join, sustain, and ultimately abandon the movement (Jasper 1998). Anger, indignation, and outrage may be the emotions most obviously associated with protesters (Gamson 1992), especially since a main reason people participate in protests is to express grievances and frustrations stemming from perceived injustices or other forms of affliction and hardship (Van Stekelenburg and Klandermans 2013). However, positive emotions can also play a role in motivating involvement. For instance, pride in participating or hope for a better future can encourage people to engage in protest movements, and it is argued that the interplay between both negative and positive emotions has a pivotal triggering effect on spurring people to action and sustaining their participation over time (Jasper 1998).

Researchers have recently started to use social media to examine the emotional expressions of individuals during times of social turmoil and tension. Studies have found correlations between negative moods expressed on Twitter with rioting activity and economic downturns (Lansdall-Welfare, Lampos, and Cristianini 2012) as well as with levels of violence in the streets (De Choudhury, Monroy-Hernández, and Mark 2014). However, such work typically considers only negative emotions in relation to protest-related activities and reactions rather than the wider emotional profile of protesters' experiences (Jasper 2011). We attempt to bridge this gap by analyzing the emotions expressed by Twitter users during the Brazilian protests that occurred in 2013 and involved millions of individuals across more than 100 cities. Unlike other movements in which the demands of protesters were reasonably explicit, the turmoil in Brazil did not have a single motivation and witnessed demonstrations of diverse grievances as the conflict grew (Monroy-Hernández and Spiro 2013). Overall, the Brazilian case study provides a compelling scenario to study how emotions change over time as protest events unfold. Our findings offer new perspectives on the role that both positive and negative emotions play during times of societal conflict and mass action.

Method

Data Collection

For this study, we used Twitter's Firehose API to collect 42,280,539 tweets shared in Brazil from May 1, 2013 to July 31, 2013; the tweets' geo-coordinates were used to identify their originating location. We included all tweets, including retweets, since retweets can be considered endorsements of the message being shared and typically express a similar sentiment (Calais Guerra et al. 2011).

*Both authors contributed equally to this work.

Protest classification

To determine whether a tweet is protest-relevant, we mined hashtags used during the Brazilian protests as an initial indicator of tweet relevance and built a co-occurrence graph between hashtags that occurred more than 8 times in the dataset. We used the Jaccard measure to weight edges, normalized the outgoing edges from each node, and ran the random walk algorithm with 5 seeds having equal probability. The following hashtags were used for seeds and are all highly correlated; they were very popular during the protests and mentioned by many news articles (Monroy-Hernández and Spiro 2013): *#ogiganteacordou (the giant woke up)*, *#vemprarua (come to the street)*, *#verasqueumfilhoteunaofogealuta (you will see that your son does not run away from the fight)*, *#protesto (protest)*, and *#protestosp (protest São Paulo)*.

Using random walk with a constant depth of 5, we considered the top 150 hashtags with the highest probability for depths 1 through 5 to be protest-relevant. We trained an external annotator who is a native Portuguese speaker to manually remove all irrelevant hashtags, ultimately resulting in a set of 478 protest-relevant hashtags. We then used these hashtags to obtain 72,083 protest-related tweets. We also used an SVM classifier with unigram features to determine whether a tweet is protest-relevant or not, regardless of the presence of a protest hashtag. After training the classifier with a random sample of 800 protest-relevant tweets along with 4000 randomly sampled non-relevant tweets, our classifier achieved 98.8% accuracy and returned a total of 199,039 protest-relevant tweets.

Sentiment classification

Given this study's main goal, our next step was to determine the sentiment expressed in the tweets of our dataset. To do so, we collected a random subset of these tweets and again trained an external annotator to label the sentiment of over 3100 as negative, neutral, or positive. Taking a random sample of 1000 tweets for each class (negative, neutral, positive), we applied both SVM and Naive Bayes for classification, with words in tweets from the training set as classifier features. The Naive Bayes classifier performed better, so we opted for a multinomial Naive Bayes classifier with 9003 binary presence/absence features — specifically unigrams occurring in the training set. The data was not fully balanced, so we randomly sampled and labeled the sentiment of 300 tweets from the entire protest-tweet corpus. We then defined the priors of each class based on the ratios of labeled emotions. Trained on 900 tweets from each class and tested on 100 tweets from each class, the classifier achieved an overall accuracy of 83%. Table 1 shows the confusion matrix.

	Negative	Neutral	Positive
Negative	81	10	9
Neutral	7	83	10
Positive	6	9	85

Table 1: Confusion matrix of Naive Bayes Classifier for sentiment classification

Results

To explore how people used Twitter throughout the protests, we studied temporal trends in posting levels, hashtag usage, tweet sentiment, and active users.

First analyzing the total number of tweets posted daily in Brazil, we found more tweets were shared in June and July than in May, with the number of tweets increasing month to month — the average number of tweets in May was 429,498, in June 465,485, and in July 483,919. June 30 saw more tweets than any other day, and while protests did occur on this day, the increased posting activity is likely more largely due to Brazil's championship win in the Confederations Cup international soccer league final (Fonseca 2013).

We next looked to protest-related Twitter activity (using the hashtag and classification based methods described previously to determine whether a tweet is protest-related or not). Figure 1 illustrates how both methods reveal consistent activity trends: the number of tweets about protests is very low in May, achieves its peak mid-June (especially from June 17 to June 21, during which time users tweeted a total of 121,652 messages about protests — 61% of all tweets), and finally begins to decrease by the end of June.

Following up on protest-related hashtag use, we found that most hashtags were used during June, especially after June 13. We suspect this relates to the fact that even though the protests began on June 6, the movement became more violent after June 13, a day known as ”bloody Thursday” due to the brutality that police displayed against protesters (Hershaw 2014). This violence fueled mobilizations in additional cities and recruited more supporters to the movement — both in the streets and on social media (Tatagiba 2013). Relatedly, we found that most tweets containing protest-related hashtags were posted from June 17-20, one of the most intense periods of the movement — over 100 thousand people protested in Rio de Janeiro on June 17, and over 1 million individuals from various cities protested on June 20 (Romero 2013). The higher, blue line in Figure 2 illustrates these trends. The lower, green line shows the number of new protest-related hashtags created over time and similarly reveals a peak in posting around June 17. This implies

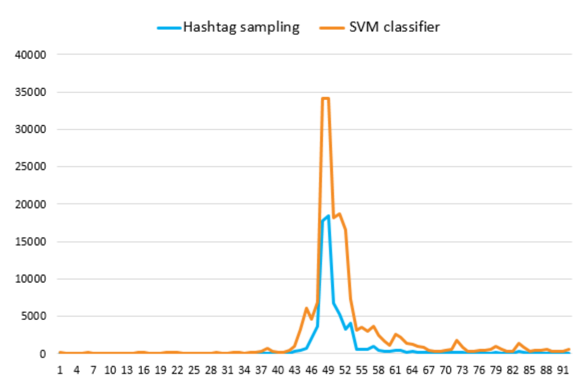

Figure 1: Number of protest-related tweets per day using hashtag-based sampling (blue) and using classifier (orange)

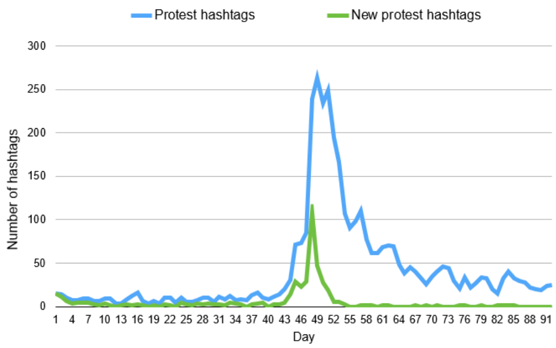

Figure 2: Number of hashtags related to protests per day

that though several protests happened after this date, people mostly continued to use the same hashtags rather than adopt new hashtags. Indeed, 93% of hashtags related to protests appeared for the first time before June 22.

Seeing how protest-related communication rose and fell over time with a peak in the middle of June, we next examined how the *emotions* expressed by Twitter users fluctuated over time. Table 2 presents the ratio of the number of tweets expressing sentiment (negative, neutral, and positive) to the number of all tweets shared that day, for all Brazilian tweets and for just protest-related tweets. During the three months of protesting, we see mostly neutral sentiment in the case of all Brazilian tweets. Low standard deviations (negative=0.013, neutral=0.015, positive=0.008) indicate those emotions remain relatively stable, though there are a few peaks and drops as seen in Figure 3. In particular, we see an increase in the ratio of negative emotions and a drop of neutral emotions in the middle of June, which coincides with the period of greater intensity of protest activity. Further, there is an increase in the ratio of neutral emotions and a drop of positive and negative emotions by the end of June.

To study sentiment in protest-related tweets specifically, we again utilized our sample of 199,039 tweets classified as protest-relevant. As seen in Table 2, we found a higher proportion of tweets express negative sentiment, not surprising given that protests are generally associated with negative emotions, especially anger (Van Stekelenburg and Klandermans 2013). Our inspection of tweets suggested that the neutral tweets may derive from news media articles that report on the protests in relatively impartial prose. The finding that 23% of protest-related tweets express *positive* emotions, however, is more surprising. Analyzing how emotional expression in protest-related tweets varied over time as seen in Figure 3, we found that the number of negative tweets is higher than the number of positive tweets on most days except one, June 17 — a similarly unexpected result given that this was one of the most intense days of protest (Romero 2013). To determine whether a small number of users sharing a large amount of positive content might be responsible for the unexpected observation of positive sentiment (O'Connor, Krieger, and Ahn 2010), we looked at the change in the number of authors expressing each emotion over time, as seen in Figure 4. Overall, the slopes of the two lines indicate that many users tweeted about protests during the time period analyzed; and regarding each emotion, more users express negative sentiment on most days, though on two days (June 17 and 18) the number of users expressing positive emotions actually was higher.

	Mean	Min	Max
Negative	(0.19, 0.45)	(0.16, 0.31)	(0.26, 0.68)
Neutral	(0.50, 0.30)	(0.44, 0.11)	(0.56, 0.47)
Positive	(0.29, 0.23)	(0.27, 0.13)	(0.32, 0.47)

Table 2: Mean, minimum, and maximum levels of negative, neutral, and positive sentiment expressed per day in (all Brazilian tweets, protest-related tweets)

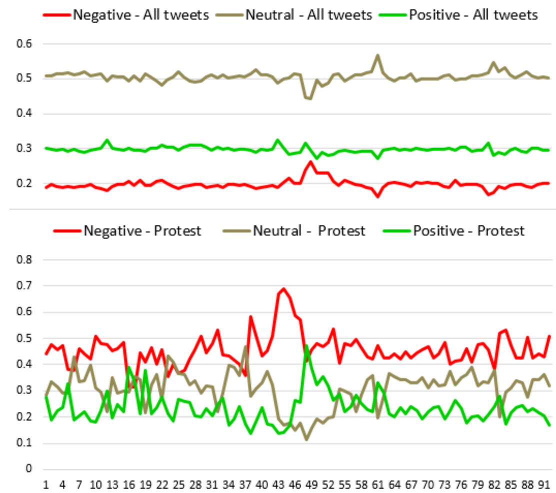

Figure 3: Negative, neutral, and positive sentiment of all Brazilian (top) and protest-related (bottom) tweets over time

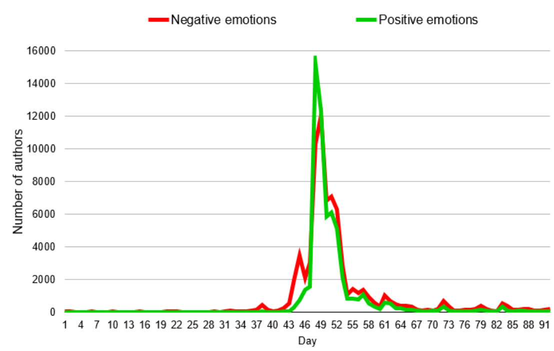

Figure 4: Number of users over time that express either negative or positive emotions in their tweets

To delve deeper for explanations, we analyzed linguistic features of positive and negative tweets. We used the log-odds-ratio to identify the most representative words and hashtags and Informative Dirichlet priors to polarize words between the two sentiment categories. This method assigns scores to each word, where negative scores correspond to negative sentiment and positive scores to positive sentiment; (Monroe, Colaresi, and Quinn 2008) provide further details about the approach. We found users tending to refer to turmoil in negative tweets, for example: *"Police threw tear gas without warning and without even tense atmosphere between protesters and cops. Aggression is systematic."* In particular, many negative words referenced the police's handling of events, such as: *atacou (attacked)*, *atirou (shot)*, *incitar (incite)*, and *inaceitável (unacceptable)*. Remaining words noted violent and illegal actions of some protesters, for instance setting fire to vehicles or causing vandalism.

Regarding the sizable number of protest tweets with positive sentiment, one potential explanation is the use of humor when referring to the movement, for example, *"Facebook stopped working. I think it is Dilma trying to censor the protests hahaha."* Usage of humor was similarly observed in studies of the 2013 protests in Turkey (Dağtas 2013). We also saw positive tweets expressing pride (*orgulho*) and satisfaction (*gratificante*) as well as patriotism, e.g., *esplendido (magnificent)*, *verasqueumfilhoteunaofogealuta (you will see that your son does not run away from the fight)*, and *patriaamadabrasil (loved country Brazil)*, all of which refer to parts of Brazil's national anthem. Many words convey hope for a better future such as *vamosmudar (let's change)* and *acreditamos (we believe)* or for peace (*paz*). The term *vemprajanela (come to the window)* was used to invite people not protesting into the streets to see what was taking place. Overall, this sort of analysis of representative emotional words appearing in tweets shed light on factors that drive participation in protests and how such motives may impact people's emotional states during these events.

Discussion and Conclusion

In this paper we undertook an analysis of how Twitter reflects protest dynamics in Brazil throughout May-July of 2013, a period in which major demonstrations happened across the country. While prior studies have analyzed sentiment on social media during periods of conflict and protest, our research explored how societal mood changes over time and in the context of the Brazilian protests. Further, we studied both positive and negative emotions, and we investigated possible explanations underlying observed emotional trends.

For researchers aiming to gain insights into collective behavior, specifically in times of social movement and political unrest, our work demonstrates the utility of computational approaches that leverage social media data. Our research also bears practical implications for government leaders, who can use social media both to make sense of motives that drive protesters' participation as well as to assess the efficacy of measures with which authorities could respond.

That being said, our data and findings are based on people who purposefully shared messages publically on social media, and opinions and emotions of individuals on Twitter may be different from those of protesters in the streets. We emphasize the need to address our approach's potential sampling bias going forward. For instance, a key next step is analyzing correlations between online Twitter activity and offline protest-related information such as reports from authorities and news media — e.g., analyzing the extent to which the negativity we see in tweets actually relates to violence happening in the streets. Future directions also involve comparing our findings from the Brazilian scenario with protests from other countries, with an aim to evaluate the generalizability both of our methodology as well as the observed social phenomena these methods enable researchers to study.

Acknowledgements

This work is supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D12PC000337, and partially supported by a Simons Investigator Award and a MURI grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Calais Guerra, P. H.; Veloso, A.; Meira Jr, W.; and Almeida, V. 2011. From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In *KDD*, 150–158. ACM.

Dağtas, S. 2013. The politics of humor and humor as politics during Turkey's gezi park protests.

De Choudhury, M.; Monroy-Hernández, A.; and Mark, G. 2014. Narco emotions: affect and desensitization in social media during the Mexican drug war. In *CHI*, 3563–3572. ACM.

Fonseca, P. 2013. Partiers outnumber protesters in Rio at Confederations Cup final. Accessed January 4, 2015.

Gamson, W. A. 1992. *Talking politics*. Cambridge Univ. Press.

Hershaw, E. 2014. Scars of police brutality in Brazilian protests haunt world cup kickoff. Accessed January 4, 2015.

Jasper, J. M. 1998. The emotions of protest: Affective and reactive emotions in and around social movements. In *Sociological forum*, volume 13, 397–424. Springer.

Jasper, J. M. 2011. Emotions and social movements: Twenty years of theory and research. *Annual Review of Sociology* 37:285–303.

Lansdall-Welfare, T.; Lampos, V.; and Cristianini, N. 2012. Effects of the recession on public mood in the UK. In *WWW*, 1221–1226.

Monroe, B.; Colaresi, M.; and Quinn, K. 2008. Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis* 16(4):372–403.

Monroy-Hernández, A., and Spiro, E. 2013. How is the Brazilian uprising using Twitter. Accessed December 10, 2014.

O'Connor, B.; Krieger, M.; and Ahn, D. 2010. Tweetmotif: Exploratory search and topic summarization for Twitter. In *ICWSM*.

Romero, S. 2013. Thousands gather for protests in Brazil's largest cities. Accessed January 7, 2015.

Tatagiba, Luciana; Blikstad, K. 2013. The Left and the June protests in Brazil. Accessed December 20, 2014.

Van Stekelenburg, J., and Klandermans, B. 2013. The social psychology of protest. *Current Sociology* 0011392113479314.