

On the Influence of Emotional Valence Shifts on the Spread of Information in Social Networks

Ema Kušen¹ and Mark Strembeck^{1,2,3}

¹Vienna University of Economics and Business

²Secure Business Austria Research Center (SBA)

³Complexity Science Hub Vienna (CSH)

Email: {firstname.lastname}@wu.ac.at

Giuseppe Cascavilla

Department of Computer Science

Sapienza - Università di Roma, Italy

Email: cascavilla@di.uniroma1.it

Mauro Conti

Department of Mathematics

Università di Padova, Italy

Email: conti@math.unipd.it

Abstract—In this paper, we present a study on 4.4 million Twitter messages related to 24 systematically chosen real-world events. For each of the 4.4 million tweets, we first extracted sentiment scores based on the eight basic emotions according to Plutchik’s wheel of emotions. Subsequently, we investigated the effects of shifts in the emotional valence on the spread of information. We found that in general OSN users tend to conform to the emotional valence of the respective real-world event. However, we also found empirical evidence that prospectively negative real-world events exhibit a significant amount of shifted emotions in the corresponding tweets (i.e. positive messages). To explain this finding, we use the theory of *social connection and emotional contagion*. To the best of our knowledge, this is the first study that provides empirical evidence for the *undoing hypothesis* in online social networks (OSNs). The *undoing hypothesis* postulates that positive emotions serve as an antidote during negative events.

Keywords - sentiment analysis, diffusion, Twitter

I. INTRODUCTION

In online social networks (OSNs), rapid information diffusion comes with valuable social benefits. For example, Twitter was an important means in helping to save lives during the 2011 Tsunami disaster in Japan [1] as well as Red River floods and Oklahoma fires in 2009 [2]. However, although having a great potential to do good for society, OSNs have also been recognized as a tool to influence people. For example, a number of studies have shown that Twitter, Facebook, and YouTube have been used to spread terrorist propaganda and negatively influence users (e.g., online radicalization).

Recent studies have recognized the important role of emotions in such information diffusion processes, indicating that emotions in OSNs are contagious [3] and may lead to a high diffusion rate [4, 5].

In this paper, we investigate the effects of specific positive (e.g., joy) and negative (e.g., anger, sadness) emotions on user behavior and information diffusion patterns. To systematically address this issue, we extracted 4.4 million tweets related to 24 real-world events that induce positive, negative, or polarizing emotional reactions. In our analysis, we distinguish between *expected emotions* (a positive event is expected to trigger predominantly positive emotions) and those of a *shifted emotional valence* (negative emotions that are related to positive events).

The paper is organized as follows. In Section II, we discuss related work followed by our data analysis procedure in

Section III. We report on our results in Section IV and discuss them in Section V. Section VI concludes the paper.

II. RELATED WORK

Some prior studies did not find significant differences in the effects of positive and negative messages. For example, [6] indicated that both spread wider than messages with a neutral sentiment score. Similar findings have been reported in [5], which suggested that emotionally-charged tweets in general tend to be re-tweeted more often than neutral ones.

In contrast, other studies found empirical evidence that supports the *Pollyanna hypothesis*, which refers to the human preference to like positive messages more than negative and neutral ones [7]. For example [8] found that positive emotions boost the transmission of messages, while negative emotions induce the opposite effect. Similar findings have been reported in [9], where the authors showed that positive messages spread wider (to more users) but also slower than negative and neutral ones. Tsugawa and Ohsaki [10], on the other hand, suggested that negative messages spread faster and wider than positive or neutral messages.

Even though a number of effects related to sentiment polarities (positive, negative, and neutral) have been studied in the related work, aspects beyond the effects of polarities have rarely been investigated. For example, Berger [4] discusses the effects of psychological arousal on information sharing. In particular, he found that *affective arousal* increases the likelihood for sharing information, regardless of whether the respective information conveys a positive or a negative sentiment. In addition, a limited number of studies examines the effects of specific emotions. For example, [11, 12] consider anger, anxiety, awe, and sadness, as annotated by human encoders. The results of both studies showed that anger and awe increase the content sharing behavior, while sadness and anxiety were negatively associated with the content diffusion.

III. DATA ANALYSIS PROCEDURE

1) Data extraction. In order to study the impact of emotions on information diffusion, we systematically identified 24 events belonging to 5 different domains (as shown in Table I). We collected 4,418,655 million publicly available tweets by using the Twitter search API and a list of predefined hashtags.

The events have been selected such that they fall in one of the following categories: 1) events that trigger positive emotions (e.g., festivities), 2) events that trigger negative emotions (e.g., death), 3) emotionally polarizing events (e.g., presidential elections). In Table I, N indicates the actual number of tweets in each category, followed by the overall size of each category in our data-set (in percent), and RT is the percentage of re-tweets.

Domain	Event	Nr. tweets
Negative	($N=1,490,495$; 34%, $RT=76.38\%$)	
Politics	1) Erdogan's threats to EU	804
	2) US anti-Trump protests	381,982
Pop culture	3) Death of Leonard Cohen	89,619
	4) Death of Colonel Abrams	1,253
War & terrorism	5) Aleppo bombings	995,561
	6) Seattle shooting	73
Other	7) Lufthansa strike	3,387
	8) Ransomware in Seattle	2,564
	9) Yellowstone incident	15
	10) Earthquake in central Italy	15,237
Positive	($N=1,115,587$; 25%, $RT=68.88\%$)	
Sports	11) Rosberg winning Formula 1	215,703
	12) Murray winning ATP	62,184
	13) Rosberg retirement message	34,201
Pop culture	14) "Beauty and the Beast" trailer release	138,979
	15) "Fantastic beasts" trailer release	64,264
	16) ComiCon Vienna	704
	17) Miley Cyrus birthday	76,270
	18) New Pentatonix album released	9,341
	19) Ellen Degeneres medal of freedom	73,854
Other	20) Thanksgiving	440,087
Polarizing	($N=1,812,573$; 41%, $RT=73.90\%$)	
Politics	21) Death of Fidel Castro	720,548
	22) 2016 Austrian presidential elections	2,558
	23) 2016 US presidential elections	891,425
Pop culture	24) The Walking Dead season 7 premiere	198,042

TABLE I

LIST OF EVENTS HAPPENED BETWEEN OCTOBER AND DECEMBER 2016.

In particular, we extracted the data for each of the 24 events on a daily basis, starting the extraction when the event was first announced and stopped 7 days after. We recorded the text of each tweet, its re-tweet count, like count, account name that posted the tweet, and the date when the tweet was published. The data extraction was restricted to English language tweets.

2) Data pre-processing. After the data was collected, we cleaned the raw data-set by removing entries that contained uninformative content with respect to emotion extraction (such as URLs).

3) Emotion extraction. After pre-processing, we lemmatized the data-set, tagged parts-of-speech, and then extracted the corresponding emotions via an R script. Our script identifies the presence of Plutchik's eight basic emotions [13] by relying on the NRC word-emotion lexicon [14]. In particular, we considered negation (e.g., "I am *not* happy.") and adverbs of degree (e.g., *very*, *hardly*, *absolutely*), as well as emoticons, which we categorized as positive (e.g., happy face ":)") or negative (e.g., sad face ":("). In addition, we extended the NRC dictionary with a list of common acronyms used in social media (such as LOL or YOLO).

For processing our data-set, we used five machines: three running on Windows 7 with 16 GB RAM and Intel Core i5-

3470 CPU @3.20 GHz, and two running on Linux - one with 32 GB RAM and Intel Xeon E3-1240 v5 CPU @3.5GHz and the other with 16 GB RAM and Intel Xeon CPU E5-2620 v3 @2.40GHz. On these 5 machines, the emotion extraction procedure took approximately a week to complete.

4) Data analysis and research questions. In our analysis, we first observed how Twitter users express specific emotions during positive, negative, and polarizing events. We separate each of the 24 event-related data-sets into a subset that conveys *expected emotions* and a subset with a *shifted emotion* in terms of its valence. We then observe how user behavior in each subset is influenced by expected and shifted emotions.

Our analysis is guided by the following research questions.

RQ1: Which emotions are expressed during positive, negative, and polarizing events?

RQ2: Which tweeting behavior¹ do users exhibit during positive, negative, and polarizing events?

RQ3: Are there differences in the tweeting behavior when users are faced with tweets that convey expected emotions and those with a shifted emotional valence?

IV. RESULTS

A) Emotion intensity in different events. Figure 1 shows the intensity of the eight emotions in each of the three categories of events (polarizing, positive, and negative, respectively). In the figure, negative emotions are colored red (anger, sadness, disgust, fear), positive green (joy, trust), and conditional (i.e. context-dependent) emotions in beige (surprise, anticipation). To mitigate bias in the results, we present the scores of each emotion averaged over the sentence count. Finally, to show the relative presence of each emotion in the data-set, we divide the averaged emotion scores e (based on the sentence count S) by the tweet count (N)

$$\frac{\sum_{i=1}^n \frac{e_i}{S_i}}{N}.$$

In polarizing events there was no tendency of a particular group of emotions to dominate over the other (the difference between the negative and positive group is 0.015, see Figure 1a). In contrast, positive events showed a higher count of positive emotions (joy, trust) (difference between positive and negative is 0.15, see Figure 1b). As expected, negative events showed a comparatively higher count of negative emotions. However, and interestingly, with a noticeable presence of positive emotions (difference between the negative and positive group is only 0.044, see Figure 1c).

B) User behavior during the analyzed events. Table II summarizes the user behavior during positive, negative, and polarizing events. In particular, positive events trigger the highest number of re-tweets and likes. This shows a tendency of users to prefer engaging in positive discussions rather than negative (which confirms the *Pollyanna hypothesis*, see also

¹We study how users react to the three types of events in terms of re-tweets, number of likes, tweeting rate, tweeting count per user, and one-to-one communication.

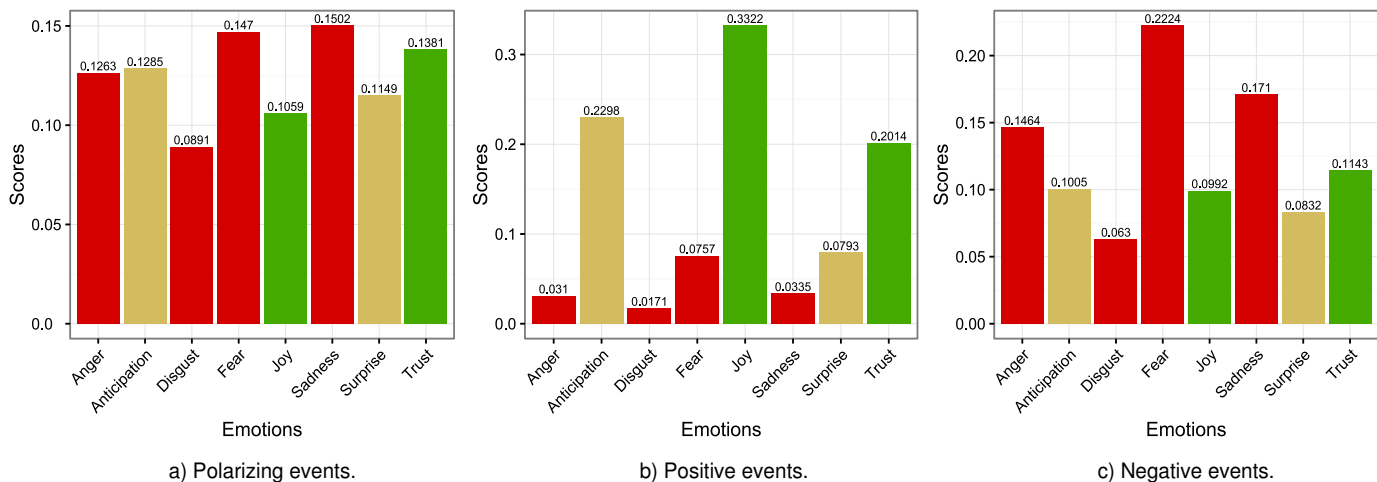


Fig. 1. Emotions in prospectively polarizing, positive, and negative events.

[7, 9]). Interestingly, users also tend to engage in a one-to-one communication (via @username) more frequently during positive events than during negative or polarizing ones. However, the results indicate that users tend to send comparatively more tweets during negative events (2.86 tweets per user) than during polarizing (1.92 tweets per user) or positive (1.62 tweets per user) events. Interestingly, in our analysis polarizing events exhibited the highest tweeting rate per minute (49.42 tweets/minute).

To examine the differences in how users respond to the tweets conveying expected emotions and those containing shifted emotions, we rely on the t-test. The results indicate a significant difference in user responses. Apart from one exception in the like count ($p > 0.05$), all other tweeting behaviors exhibit a clear pattern. In particular, expected emotions receive more re-tweets, are distributed faster, and are dominant in a one-to-one communication (@-count). Another interesting insight can be observed while examining the differences between negative and positive events. While negative events tend to exhibit a higher number of negative tweets per user, positive events dominate in the retweet count of positive tweets and one-to-one social sharing of predominantly positive emotions.

	Negative	Positive	Polarizing
Retweet count	1600.10 $t=1.98, p<0.05$	5677.63 $t=243.511, p<0.05$	4821.90
Like count	0.98 $t=-0.97, p>0.05$	1.49 $t=0.98, p>0.05$	1.22
Time rate	37.58 $t=25.94, p<0.05$	42.48 $t=57.47, p<0.05$	49.42
Tweet per user	2.86 $t=13.33, p<0.05$	1.62 $t=26.04, p<0.05$	1.92
@-count	1.02 $t=25.77, p<0.05$	1.19 $t=69.93, p<0.05$	1.02

TABLE II

USER BEHAVIOR IN POSITIVE, NEGATIVE, AND POLARIZING EVENTS.

Figure 2a) shows the time series of tweets divided into the ones that carry a predominantly positive emotion and those that carry a negative emotion. During positive events,

there is a noticeable smaller number of tweets that convey negative emotions. However, the data shows that, although small in size, negative tweets occur consistently throughout the data-set ($mean(set\ difference)=36668.11, sd(set\ difference)=45844.64$)².

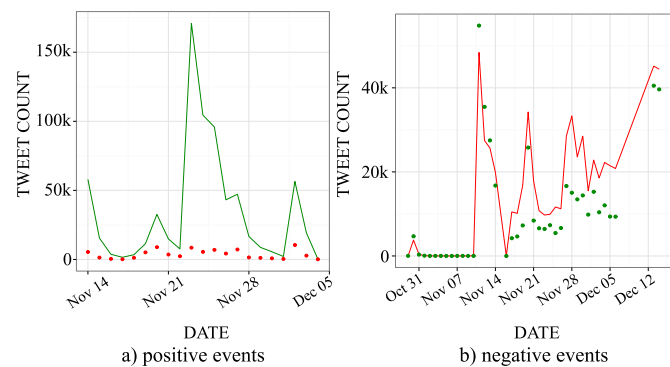


Fig. 2. Negative tweets during positive events and vice versa. Positive emotions are colored green and negative emotions red.

In contrast, our analysis indicates that this is not the case during negative events. As shown in Figure 2b), during negative events positive tweets consistently follow the tweets conveying expected negative emotions ($mean(set\ difference)=4231.67, sd(set\ difference)=5603.162$). Interestingly, our data-set also revealed unexpected cases where (shifted) positive tweets even exceed the (expected) negative tweets (the largest difference between the two subsets is 6403 tweets).

V. DISCUSSION

Our results bring forth interesting insights into how OSN users behave during positive, negative, and polarizing events when faced with shifted emotions. Delving into the field of psychology, we found that positive emotions occur also during negative events. An explanation for the observed phenomenon

²Set difference refers to the difference between the count of the expected emotions and shifted emotions.

can be attributed to *social connection* [15] as one of the fundamental human needs. In our data-set we found examples of people explicitly calling for social bonding during emotionally tough events (e.g., after the earthquake in central Italy, people tweeted: “*please join us as we #PrayforItaly*”) and a public and explicit expression of vulnerability that triggers compassion (e.g., “*Oh dear world, I am crying tonight*”, during Aleppo bombings). In particular, our data-set indicates that people tend to show appreciation and love for the person they care about or admire (e.g., a deceased singer, such as Leonard Cohen) or even comfort each other and send messages of hope during natural disasters (e.g., earthquakes in Italy) or war (e.g., Aleppo bombing). Thus, we found empirical evidence that supports the *undoing hypothesis* [16], which states that people tend to use positive emotions as an antidote to undo the effects of negative emotions.

Our results further show that expected emotions result in more re-tweets. We thereby confirm findings from [17], which indicated that people prefer to share messages that correspond to the emotional valence of the domain. This might be attributed to the human tendency to conform to the situation. Other studies bring an additional interesting insight into interpersonal interactions over social media, which might explain our observations. According to [3], emotional messages tend to influence the emotions conveyed in other users’ messages. This phenomenon, called *emotional contagion*, relies on studying the social connections of the users (or their position in the network). Even though studying the network structure of the users participating in the Twitter discourse during positive, negative, and polarizing events is out of the scope of this study, we were able to observe that messages sent by “fans” (we follow an assumption that fans follow their idols on Twitter with a high probability) tend to be congruent with the messages sent by their idols. For example, a tweet posted by Pentatonix (announcement of their new album) triggered positive reactions from their fans (in terms of retweeting and original positive tweets).

VI. CONCLUSION AND FUTURE WORK

In this paper, we studied the influence of emotional valence shifts on information diffusion in Twitter. We began our study by extracting over 4.4 million tweets belonging to positive, negative, or polarizing events. For each of the three categories, we analyzed the intensity of Plutchik’s eight basic emotions and identified shifts in emotional valence. We found that users generally prefer sharing messages that correspond to the emotional valence of the respective event. While conducting a time-series analysis, we found a clear distinction between positive and negative events, with respect to shifted emotions. We found that positive events trigger a comparatively small number of negative messages. However, while negative events exhibit predominantly negative messages, they are accompanied by a surprisingly high number of positive messages (in our data-set, this occurred in 17.5% of the cases). To the best of our knowledge, this is the first study which found empirical

evidence that supports the theory of the *undoing hypothesis* in OSNs.

In our future work, we plan to extend our analysis to studying messages written in languages other than English. Moreover, we are currently studying the effects of shifts in emotional valence over a Facebook and YouTube data-set. Another interesting direction would be to examine the impact of bots on the valence shifts.

ACKNOWLEDGMENT

Mauro Conti is supported by a Marie Curie Fellowship funded by the European Commission (PCIG11-GA-2012-321980), partially supported by the TENACE PRIN Project 20103P34XC funded by the Italian MIUR and by the Project “Tackling Mobile Malware with Innovative Machine Learning Techniques” funded by the University of Padua, and by the EU TagItSmart! Project (H2020-ICT30-2015-688061).

REFERENCES

- [1] C. St Louis and G. Zorlu, “Can Twitter predict disease outbreaks?” *BMJ*, vol. 344, 2012.
- [2] S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen, “Microblogging during two natural hazards events: What Twitter may contribute to situational awareness,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’10, 2010.
- [3] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, “Experimental evidence of massive-scale emotional contagion through social networks,” *Proc. of the National Academy of Sciences*, vol. 111, no. 24, 2014.
- [4] J. Berger, “Arousal increases social transmission of information,” *Psychological Science*, vol. 22, no. 7, 2011.
- [5] S. Stieglitz and D.-X. Linh, “Emotions and information diffusion in social media- sentiment of microblogs and sharing behavior,” *Journal of management information systems*, vol. 29, no. 4, 2013.
- [6] D. N. Trung, T. T. Nguyen, J. J. Jung, and D. Choi, *Understanding Effect of Sentiment Content Toward Information Diffusion Pattern in Online Social Networks: A Case Study on TweetScope*, 2014.
- [7] J. Boucher and C. E. Osgood, “The pollyanna hypothesis,” *Journal of Verbal Learning and Verbal Behavior*, vol. 8, no. 1, 1969.
- [8] H. S. Kim, S. Lee, J. N. Cappella, L. Vera, and S. Emery, “Content characteristics driving the diffusion of antismoking messages: Implications for cancer prevention in the emerging public communication environment,” *Journal of National cancer institute. Monographs*, vol. 47, 2013.
- [9] E. Ferrara and Z. Yang, “Quantifying the effect of sentiment on information diffusion in social media,” *PeerJ Computer Science*, vol. 1, 2015.
- [10] S. Tsugawa and H. Ohsaki, “Negative messages spread rapidly and widely on social media,” in *Proceedings of the 2015 ACM on Conference on Online Social Networks*, 2015.
- [11] B. Suh, L. Hong, P. Pirolli, and E. H. Chi, “Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network,” in *Proc. of the 2010 IEEE Second International Conference on Social Computing*, 2010.
- [12] I. Heimbach and O. Hinz, “The impact of content sentiment and emotionality on content virality,” *International Journal of Research in Marketing*, vol. 33, no. 3, 2016.
- [13] R. Plutchik, “The nature of emotions,” *American Scientist*, vol. 89, no. 4, 2001.
- [14] S. M. Mohammad and P. D. Turney, “Crowdsourcing a word-emotion association lexicon,” *Computational Intelligence*, vol. 29, no. 3, 2013.
- [15] R. F. Baumeister and M. R. Leary, “The need to belong: Desire for interpersonal attachments as a fundamental human motivation,” *Psychological Bulletin*, vol. 117, 1995.
- [16] B. L. Fredrickson, “The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions,” *The American Psychologist*, vol. 56, 2001.
- [17] “Do people prefer to pass along good or bad news? valence and relevance of news as predictors of transmission propensity,” *Organizational behavior and human decision processes*, vol. 68, no. 2.