Why so Emotional?
An Analysis of Emotional Bot-generated Content on Twitter

Ema Kušen[1] and Mark Strembeck[1,2,3]
[1]*Vienna University of Economics and Business (WU Vienna), Vienna, Austria*
[2]*Secure Business Austria (SBA), Vienna, Austria*
[3]*Complexity Science Hub (CSH), Vienna, Austria*
ema.kusen@wu.ac.at, mark.strembeck@wu.ac.at

Keywords: Bot Behavior, Emotion Analysis, Emotional Conformity, Temporal Patterns, Twitter.

Abstract: In this paper, we present a study on the emotions conveyed in bot-generated Twitter messages as compared to emotions conveyed in human-generated messages. Social bots are software programs that automatically produce messages and interact with human users on social media platforms. In recent years, bots have become quite complex and may mimic the behavior of human users. Prior studies have shown that emotional messages may significantly influence their readers. Therefore, it is important to study the effects that emotional bot-generated content has on the reactions of human users and on information diffusion over online social networks (OSNs). For the purposes of this paper, we analyzed 1.3 million Twitter accounts that generated 4.4 million tweets related to 24 systematically chosen real-world events. Our findings show that: 1) bots emotionally polarize during controversial events and even inject polarizing emotions into the Twitter discourse on harmless events such as Thanksgiving, 2) humans generally tend to conform to the base emotion of the respective event, while bots contribute to the higher intensity of shifted emotions (i.e. emotions that do *not* conform to the base emotion of the respective event), 3) bots tend to shift emotions to receive more attention (in terms of likes and retweets).

1 INTRODUCTION

Over the past decade, we have witnessed a fast growing interest in online social media and online social networks (OSNs). Aside from convenient role of OSNs that help us stay in touch with our friends and family, OSNs also support people while trying to organize themselves during natural disasters (St Louis and Zorlu, 2012) or in political movements (Howard et al., 2011). A recent statistic reported that currently about 2.46 billion individuals use social media, with a predicted increase to 3.02 billion till 2021[1].

In addition to ordinary social media users, researchers also significantly benefit from the large volume of data that emerges from the use of OSNs. In particular, these data allow for gaining unprecedented insights into various aspects of online user behavior, including the identification of influential users, the study of information diffusion on social media platforms, or the temporal evolution of topics and communities that emerge on OSNs. Thus, OSNs enable an improved understanding of complex micro- and macro-societal phenomena (Thai et al., 2016; Eagle and Pentland, 2006).

Since OSNs are networks of people, *human emotions* can significantly influence user behavior and information diffusion over OSNs. In particular, emotions communicated in OSN messages may boost or decrease the diffusion rate of the corresponding messages (Kim et al., 2013; Tsugawa and Ohsaki, 2015). Moreover, recent studies found empirical evidence of emotional contagion through OSNs (Kramer et al., 2014).

In this paper, we extend our previous work concerning the influence of emotions on OSN user behavior (Kušen et al., 2017b) by distinguishing between the emotions conveyed in Twitter messages sent by humans as compared to the emotions conveyed in Twitter messages sent by social bots.

A single network of Twitter bots may consist of several hundred thousand automated accounts[2]. Re-

[1]https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

[2]http://www.bbc.com/news/technology-38724082

Kušen, E. and Strembeck, M.
Why so Emotional? An Analysis of Emotional Bot-generated Content on Twitter.
In *Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS 2018)*, pages 13-22
ISBN: 978-989-758-297-4

cent studies indicated that a single bot may generate as many as 500 tweets per day (Kollanyi et al., 2016) and bot-generated content may reach up to 19% of the tweets on particular topics (Bessi and Ferrara, 2016), leading to a threat of negatively affecting the public opinion. For example, Ferrara et al. discussed a number of malicious consequences that might arise from bot activities, including perception altering, destroying user reputation, or manipulating the users' opinions (Ferrara et al., 2016). Thus, given the high volume of bot-generated content, it is important to study the potential impact of bots on emotional content disseminated through OSNs.

Previous studies mainly focused on the presence of bot accounts in the social media discussion on a single real-world event, thus making the findings difficult to generalize (Dickerson et al., 2014; Bessi and Ferrara, 2016).

For this paper, we analyzed a data-set consisting of 4.4 million Twitter messages related to 24 systematically chosen events. In particular, we analyzed the data-set for the presence and intensity of the eight basic emotions identified by Robert Plutchik (Plutchik, 2001). The messages in our data-set have been sent from 1.3 million distinct Twitter accounts, 35.2 thousand of which we identified as bots.

Our analysis shows that human and bot accounts exhibit distinct behavioral patterns with respect to the emotions they spread. While humans tend to conform to the base emotion of an event (e.g., express sadness or fear during negative events), bot accounts disseminate a more heterogeneous set of emotions. The distinction between bot and human behavior is especially evident during polarizing events, where bot accounts purposefully pick sides, i.e. they follow a strategic agenda to influence human users.

The remainder of the paper is organized as follows. In Section 2 we summarize related work. Section 3 outlines our research procedure, followed by a report on our results in Section 4. We further discuss our findings in Section 5 and conclude the paper in Section 6.

2 RELATED WORK

As discussed in (Abokhodair et al., 2015; Chu et al., 2010; Kollanyi et al., 2016; Mascaro et al., 2016), bot accounts differ from human accounts in their tweeting frequency. However, Chavoshi et al. suggested that particular bots may also delete some of their tweets to reach a tweet generation rate that is comparable to human accounts (Chavoshi et al., 2017b). The difference between bots and human accounts is especially evident when observing temporal tweeting patterns, as human accounts tend to predominantly tweet during weekdays, whereas bots are equally active during weekdays and in the weekends (Chu et al., 2010).

Another distinction between Twitter bots and humans lies in the follower-followee ratio. In a study involving 500,000 Twitter accounts, Chu et al. found that bot accounts tend to follow a lot of users, but have only a few followers themselves (Chu et al., 2010).

Moreover, it has been reported that Twitter bots often have an agenda, e.g., by trying to persuade or manipulate with Twitter users. To achieve this, bots typically use content-related features, such as URLs and hashtags to promote their messages. For example, they may boost the perceived importance of a specific topic by (re)tweeting a certain URL (Chu et al., 2010; Gilani et al., 2017), use hashtags, or even identify and mention potentially interested target users (@username) to mobilize people for action. In fact, (Savage et al., 2016) indicated that bot accounts can capture the attention of human Twitter users who subsequently engage in a discussion on the bot-generated topic and further promote bot-generated content.

Multiple studies indicated that bot accounts may endanger democratic elections by swaying the voters' opinions, spreading misinformation, or even amplifying the perceived influence of a specific political candidate (Ferrara, 2017). Ratkiewicz et al. studied bot activities related to US politics and found that bots are responsible for generating thousands of tweets that contain links and strategically mention a few popular users (Ratkiewicz et al., 2011). These users in turn receive tweets sent by bot accounts and spread the respective tweets to their followers. Thereby, Ratkiewicz et al. found empirical evidence which confirms that bots may generate information cascades. Moreover, Kollanyi et al. found that bot accounts indeed have an impact on the political discourse over Twitter (Kollanyi et al., 2016). By analyzing the 2016 US Presidential Elections, Kollanyi et al. found that bots systematically combined pro-Trump hashtags with neutral and pro-Clinton hashtags such that by the time of the election, 81.9% of the bot-generated content involved some pro-Trump messaging.

The importance of studying sentiments communicated by bots and human accounts has been addressed by Dickerson et al. , who suggested that humans tend to express positive opinions with a higher intensity, as compared to bots (Dickerson et al., 2014). Moreover, the authors found that humans tend to disagree more with the base sentiment of the event they studied (the 2014 Indian election). Based on this finding, (Everett et al., 2016) indicated that bot-generated messages which disagree with the opinion of a crowd are deceptive, reaching a high likelihood of 78% to trick people into believing a particular bot message was actually generated by a human account. However, some of the prior findings mentioned above could not be confirmed in our analysis (see Section 5).

3 RESEARCH PROCEDURE

The analysis we conducted for this paper included five phases (see Figure 1).

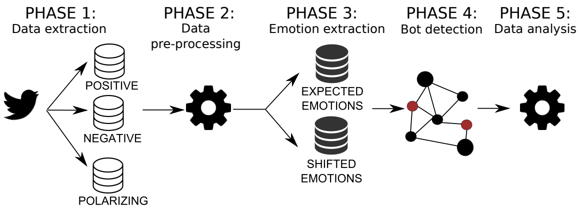

Figure 1: Research procedure.

Phase 1. We systematically collected 4,418,655 tweets related to 24 events that can be classified either as positive (e.g., release of an acclaimed movie), negative (e.g., a natural disaster), or polarizing (e.g., political campaigning) (see Table 1). For data extraction, we used Twitter's Search API[3] and a list of carefully selected hashtags for each of the 24 events (Kušen et al., 2017b). For each event, we extracted tweets published within one week since the event's announcement[4] and restricted the extraction to tweets written in English language only. In total, it took three months to systematically collect tweets related to the 24 events in our study (October 2016 - December 2016).

Phase 2. After obtaining the data-set, we conducted several pre-processing steps, e.g., by removing duplicate entries and information irrelevant with respect to emotion extraction, such as URLs (Van den Broeck et al., 2005).

Phase 3. Next, we applied our emotion extraction procedure to the pre-processed data-set. In particular, the emotion extraction relies on a number of heuristics used to assess emotions in written texts (such as negation, emoticons, or adverbs of degree, see (Kušen et al., 2017a) for details) and results in an own vector of emotional intensities for each of the 4.4 million tweets, i.e. for each tweet in the data-set we identified the presence and the intensity for each of the eight basic emotions found in the Plutchik's wheel of emotions (anger, disgust, fear, sadness, joy, trust, surprise, anticipation).

Table 1: List of events analyzed in our study.

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

For processing the 4.4 million tweets, we used five machines to run a corresponding R script: three machines 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.

Phase 4. Next, we extracted the list of unique screen names (i.e. Twitter user names) from the tweets in our data-set. This list of screen names has then been analyzed via DeBot (Chavoshi et al., 2016a; Chavoshi et al., 2016b) to obtain bot scores for each of the corresponding Twitter accounts. In our analysis we used DeBot, because it reaches a higher precision as compared to other bot detection approaches (Chavoshi et al., 2016a). DeBot correlates account activities of millions of users in near real-time, it can detect bot accounts within just two hours since they

[3]https://developer.twitter.com/en/docs

[4]Note that some events we considered for our study actually started several weeks before we collected our data, such as the bombings in Aleppo or the announcement of the US and Austrian presidential elections. For such events, we extracted tweets related to an important episode of the respective event. For example, we extracted tweets related to the 2016 Austrian presidential elections published one week before the actual election date.

started their activities (Chavoshi et al., 2016a), and is able to identify synchronized bot behavior (Chavoshi et al., 2017a).

In total, we used DeBot to analyze 1,317,555 distinct Twitter user accounts, 35,247 of which have been identified as bots – giving us an overall percentage of 2.67% of bot accounts in our data-set.

Phase 5. In the final step, we analyzed our data-set.

Since it includes events related to three different base emotions (i.e. events that are either positive, negative, or polarizing, see also Table 1), the goal of this paper is to study how emotions conveyed by Twitter bots compare to emotions spread by human accounts. More specifically, we are interested in the impact of bots on the diffusion of emotional content.

To this end, Section 4 reports our findings in three parts: 1) relative intensities for each of the eight basic emotions as conveyed by bots and human accounts, 2) temporal patterns in tweeting of emotional content by bots and human accounts, and 3) user reactions on emotional tweets sent by bots and human accounts.

4 ANALYSIS RESULTS

4.1 Intensities of Emotions Conveyed in Tweets Authored by Human and Bot Accounts

Figures 2-4 show the relative presence and intensity for each of the eight emotions during positive, negative, and polarizing events. In the figures, we visualize positive emotions (trust, joy, anticipation[5]) in green, negative emotions in red (anger, disgust, sadness, and fear), and the conditional emotion *surprise*, which can, by default, neither be classified as positive nor negative, in yellow. In Figures 2-4, the scores for each emotion e are averaged over the sentence count S and divided by the tweet count (N)

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

[5]We classify *anticipation* as a positive emotion because Spearman's correlation coefficient ρ has shown that anticipation correlates strongly with positive emotions (joy, trust) and only weakly with negative emotions (anger, fear, sadness, disgust). For example, for tweets related to positive events, anticipation strongly correlated with trust ρ=0.69, but only weakly with fear ρ=0.31. We observed the same pattern for negative and polarizing events. In contrast, *surprise* did not exhibit a strong correlation with either positive or negative emotions. Therefore, we treat *surprise* as a separate category.

Our results show that humans in general tend to conform to the base emotion of the respective event (e.g., predominantly positive emotions are sent during positive events, and predominantly negative emotions during negative events).

Twitter bots, however, exhibit a different behavioral pattern with respect to the emotions they convey in their messages. In particular, during negative events humans exhibit a larger difference (d) between positive and negative emotions (d_{n-p}=0.192) as compared to bot accounts (d_{n-p}=0.005) (see Figure 2). The same observation can be made for tweets sent during positive events (see Figure 3), where the difference between positive and negative emotions sent by human accounts is d_{p-n}=0.666, while the difference drops to d_{p-n} =0.282 for bot accounts. During polarizing events (e.g., political campaigning or other controversial topics) one can expect a mixture of emotions. Therefore, as expected, humans and bots alike express positive as well as negative emotions in polarizing events (see Figure 4). As an interesting finding we observed that human accounts are more negatively inclined during polarizing events (d_{n-p}=0.0189), while bot accounts tend to send more tweets receiving a positive emotion score during polarizing events (d_{n-p}=-0.102).

However, note that especially in polarizing events, a positive emotion score does not necessarily convey a positive message but is usually biased towards one of the polarizing opinions (see Section 5). Thus, the positive emotion score often results from a bot that purposefully "picked a side" to promote. An example of such a biased bot-generated tweet from our data-set reads: "*#ObamaFail I'll be so happy to see this joke move out of the White House!! #VoteTrumpPence16*".

To further examine how well emotions communicated by bot accounts correlate with those expressed by human accounts, we converted the intensities of each emotion into a ranked list for each emotion category (positive, negative, polarizing) and obtained Kendall's rank coefficient τ. The results indicate that in general human and bot accounts tend to spread comparative emotions during positive events (Kendall's τ is a strong positive 0.85) as well as during negative events (Kendall's τ is a moderate positive 0.5). However, we found a larger distinction between humans and bots during polarizing events (Kendall's τ is a weak positive 0.14).

In order to provide more insight into this observation, we examine the impact of retweets on the overall emotion scores in our data-sets. In general, we found that humans as well as bots predominantly sent retweets. During positive events, 68.80% of the tweets generated by human accounts are retweets, while 80.37% of the messages sent by bots consist of retweets. A similar, though weaker effect holds for tweets sent during negative events: 76% of the content generated by human accounts are retweets, while bots generated 79.2% retweets. During polarizing events, human accounts sent 73.46% retweets, while 87.9% of the messages sent from bot accounts consist of retweets.

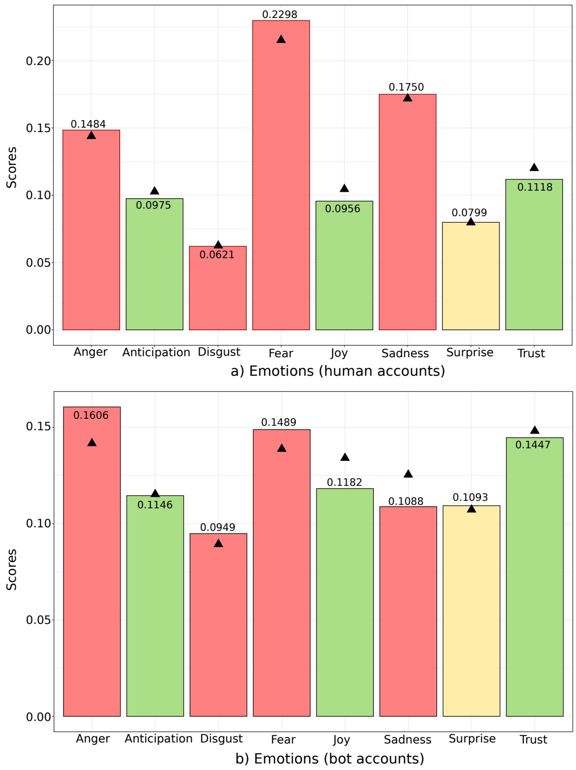

Figure 2: Emotions expressed by human and bot accounts during negative events. The effects of retweets are depicted with a black triangle.

When adjusted for the retweets, unique occurrences of tweets on negative events (see Figure 2) show that positive emotions (joy, trust, and anticipation) are amplified by the effects of retweets by human as well as bot accounts. Moreover, Figure 2b) indicates that bots especially tend to amplify *sadness* during negative events.

During positive events (see Figure 3), retweets disseminated by human accounts amplify *joy*, but also *anger* and *sadness*, whereas bot accounts amplify *anger*, *sadness*, as well as two positive emotions (*joy*, *anticipation*).

During polarizing events, *joy*, *anticipation*, and *trust* are amplified by the retweets generated by human as well as bot accounts, while *surprise* is boosted by the retweets generated by bot accounts only (see Figure 4).

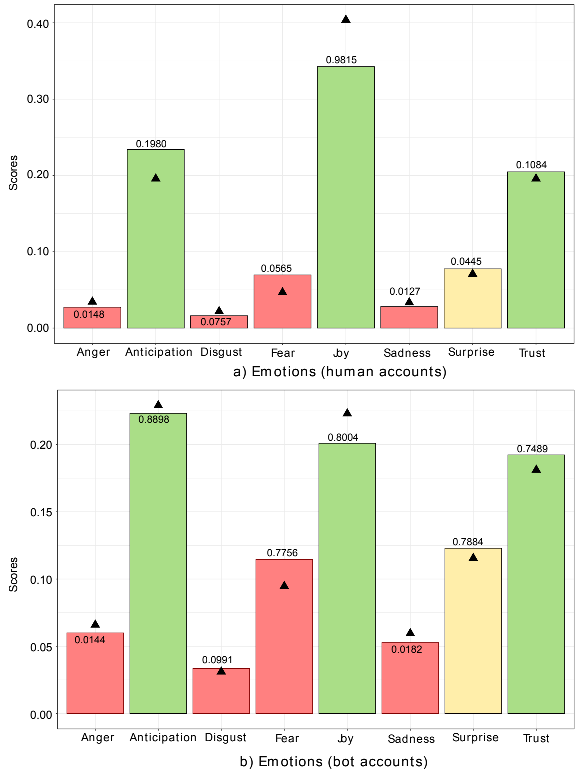

Figure 3: Emotions expressed by human and bot accounts during positive events. The effects of retweets are depicted with a black triangle.

Next, we examine whether the differences between bot accounts and human accounts are statistically significant. Therefore, we first define the following null hypothesis:

H_0: *There is no difference between the mean scores of the respective emotions sent by bot and human accounts.*

We use Welch's two sample t-test with a 95% confidence level where we contrast the emotion scores in the subsets containing tweets sent by bots and human accounts, respectively. The t-test results (see Table 2) indicate that there is a statistically significant difference in the intensities of emotions spread by bot accounts and human accounts in all three event types (positive, negative, and polarizing). In particular, bots sent on average more negative emotions (anger, disgust, sadness, and fear) during positive events as compared to human accounts. Thus, we reject the null hypothesis. Moreover, the t-test results indicate that bots do not tend to comply with the positive base emotion expected during positive events – a trait which significantly distinguishes bots from human emotional reactions to positive events (Heath, 1996). Our results also indicate that bot accounts tend to send more positive messages containing joy, trust, and anticipation dur-

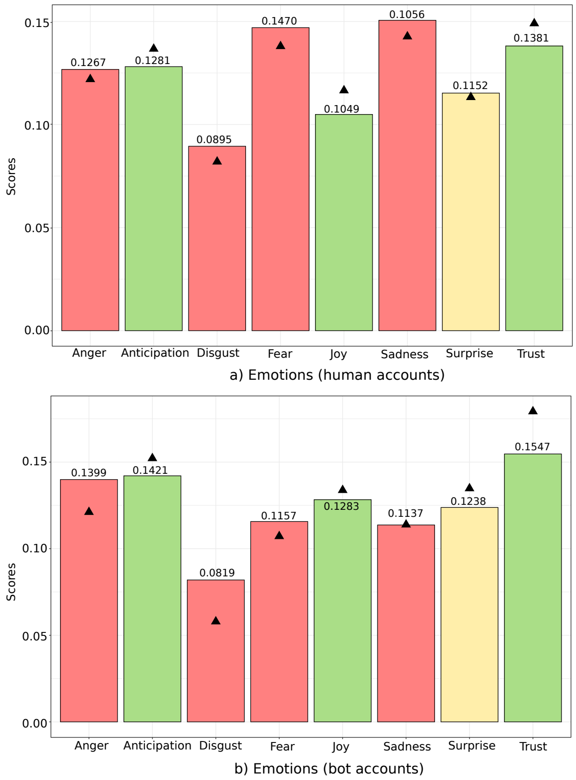

Figure 4: Emotions expressed by human and bot accounts during polarizing events. The effects of retweets are depicted with a black triangle.

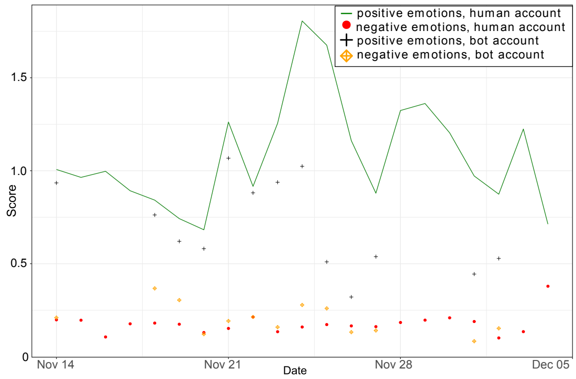

Figure 5: Temporal patterns during positive events.

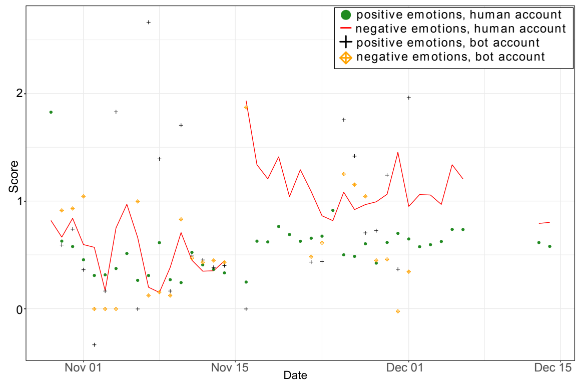

Figure 6: Temporal patterns during negative events.

ing polarizing events compared to human accounts. However, as mentioned above, in polarizing events, a positive emotion score does not necessarily convey a positive message but is usually biased towards one of the polarizing opinions (see also Section 5).

When adjusted for the effects of retweets (i.e. by considering unique tweets only), we found that no particular emotion is more intensely expressed in unique tweets generated by bot accounts (see Table 2 without retweet (RT) entries). This confirms that bots especially tend to amplify certain emotions by retweeting.

4.2 Temporal Patterns

We now examine whether distinctive temporal patterns exist for human and bot accounts. Thus, we compared the intensities of positive and negative emotions averaged over each day of data extraction for both account types (human and bot).

Figure 5 shows the temporal development of emotion scores during positive events and indicates that bot as well as human behavioral pattern are comparative in positive events – though the average difference between the intensities of positive and negative emotions is smaller for bot accounts (d_{p-n}=0.282). In this context, it is worth mentioning that the bot accounts identified in our study only marginally contributed to the Twitter discourse related to positive events (only 0.63% of the corresponding messages was generated by bot accounts).

However, for negative and polarizing events this finding does not hold. Figure 6 shows that during negative events we can also observe a considerable number of positive messages. Previous studies have shown that people tend to comfort each other during negative events, such as natural disasters and terror attacks (Kušen et al., 2017b). In particular, this observation can be explained by the *undoing hypothesis* (Fredrickson, 2001) which refers to the human tendency to remain positive in order to undo the effects of negative emotions. As an interesting finding, Figure 6 shows that not only human accounts but also bots tend to disseminate positive emotions during negative events (we refer the reader to the "black cross" symbols in Figure 6). In fact, positive emotions sent by bot accounts tend to dominate on certain dates over the negative emotions – which, again, distinguishes bot tweeting-behavior from human behavior.

Figure 7 shows the temporal development of emotion scores during polarizing events. As expected,

Table 2: Results of Welch's two sample t-test with a 95% confidence level of two samples (bots and humans respectively). Numbers in brackets indicate degrees of freedom. Statistically significant results are shown in bold.

	Polarizing ($N_{human} = 1757663$; $N_{bot} = 54910$)	Positive ($N_{human} = 1108577$; $N_{bot} = 7010$)	Negative ($N_{human} = 1453591$; $N_{bot} = 36904$)
Anger			
t (with RT)	**t(57858)=-4.3941, p<0.05**	**t(1115600)=5.6, p<0.05**	**t(1490500)=-26.82, p<0.05**
t (without RT)	**t(121030)=-6.88, p<0.05**	t(69787)=0.752, p>0.05	**t(100350)=-14.30, p<0.05**
Disgust			
t (with RT)	**t(1490500)=-16.35, p<0.05**	**t(1115600)=2.34, p<0.05**	t(38559)=-0.86, p>0.05
t (without RT)	**t(121030)=-13.26, p<0.05**	t(69787)=-1.05, p>0.05	t(9791)=-1.45, p>0.05
Sadness			
t (with RT)	**t(1812600)=-37.86, p<0.05**	**t(1115600)=3.27, p<0.05**	**t(1490500)=-60.42, p<0.05**
t (without RT)	**t(121030)=-15.37, p<0.05**	t(69787)=-0.15, p>0.05	**t(100350)=-24.37, p<0.05**
Fear			
t (with RT)	**t(1812600)=-34.08, p<0.05**	**t(7083.3)=3.5, p<0.05**	**t(1490500)=-70.48, p<0.05**
t (without RT)	**t(121030)=-16.42, p<0.05**	t(69787)=1.95, p>0.05	**t(100350)=-29.63, p<0.05**
Trust			
t (with RT)	**t(58244)=4.63, p<0.05**	**t(1115600)=-17.03, p<0.05**	**t(1490500)=-17.1, p<0.05**
t (without RT)	t(100350)=0.32, p>0.05	**t(69787)=-11.92, p<0.05**	**t(100350)=-10.15, p<0.05**
Joy			
t (with RT)	**t(1812600)=10.87, p<0.05**	**t(1115600)=-35.4, p<0.05**	**t(1490500)=-17.97, p<0.05**
t (without RT)	t(100350)=-1.02, p>0.05	**t(69787)=-24.41, p<0.05**	**t(100350)=-8.16, p<0.05**
Anticipation			
t (with RT)	**t(58354)=1.09, p<0.05**	**t(1115600)=-14.83, p<0.05**	**t(1490500)=-23.73, p<0.05**
t (without RT)	**t(100350)=-3.13, p<0.05**	**t(69787)=-7.58, p<0.05**	**t(100350)=-12.74, p<0.05**
Surprise			
t (with RT)	**t(58337)=-1.98, p<0.05**	**t(1115600)=3.78, p<0.05**	**t(1490500)=-10.91, p<0.05**
t (without RT)	t(100350)=-0.32, p>0.05	t(69787)=0.14, p>0.05	**t(100350)=-6.37, p<0.05**

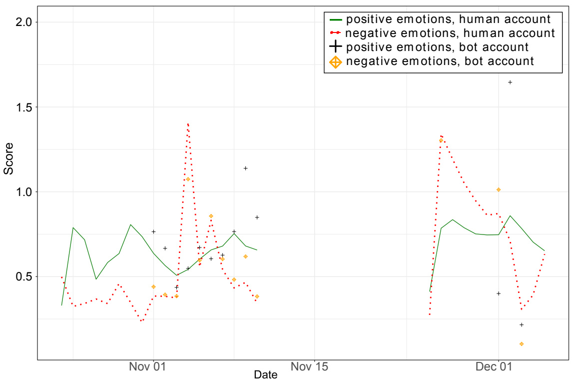

Figure 7: Temporal patterns during polarizing events.

both bot and human accounts tend to exhibit mixed emotions about polarizing events. However, we still found a distinctive behavioral pattern in our data-set. Interestingly, bots incline towards positive emotions during polarizing events, as compared to humans who generated more tweets that convey negative emotions. However, we also found that in spite of a positive emotion score, the corresponding tweets often do not convey a positive message but are biased towards one of the polarizing opinions (see Section 5).

Moreover, the larger difference between the intensities of positive and negative emotions conveyed in bot-generated tweets (d_{n-p}=-0.102) shows a tendency of bot accounts to pick sides, i.e. they polarize and/or amplify a certain sentiment (positive or negative). Compared to the positive and negative events in our data-set (see Table 1), bot accounts contributed more tweets during polarizing events (54910 tweets, 3%).

4.3 Effects of Emotions on User Reactions

Finally, we study user reactions on tweets generated by human and bot accounts. Table 3 summarizes the means and standard deviations of two distinct user reactions to tweets – the retweet count and the like count. The results show that human-generated tweets which carry positive emotions receive on average more retweets compared to tweets generated by bot accounts. In contrast, positive bot-generated tweets receive on average more likes during polarizing and negative events. The same holds for bot-generated tweets conveying negative emotions – such tweets receive on average more likes if they belong to polarizing or negative events. Moreover, negative bot-generated tweets during positive events tend to receive more retweets, as compared to those published by human accounts. In terms of emotionally neutral content, tweets generated by human accounts tend to receive more retweets and likes.

Our findings (as summarized in Table 3) bring

Table 3: Summary of user reactions (mean and standard deviation) on emotional content disseminated by bot and human accounts. Bot-related table entries which received more attention in terms of liking or retweeting as compared to human-generated tweets are printed in bold.

	Polarizing	Positive	Negative
Positive			
RT_{human}	6142.62±18921.07	5727.02±17087.09	1502.73±4890.3
$Like_{human}$	1.19±76.18	1.48±77.44	1.06±40.57
RT_{bot}	2629.78±11332.41	1389.35±5674.46	404.8±1209.07
$Like_{bot}$	**1.47±28.41**	0.815±13.18	**1.49±34.11**
Negative			
RT_{human}	2910.25±7452.78	407.41±1677.12	1703.89±5157.2
$Like_{human}$	1.17±94.23	1.3±29.19	0.94±35.47
RT_{bot}	1991.02±5654	**608.41±2269.43**	659.69±2213.14
$Like_{bot}$	**2.26±37.94**	0.678±14.91	**1.38±27.84**
Neutral			
RT_{human}	5546.71±17686.4	7857.25±11441.16	1625.84±4787.55
$Like_{human}$	1.29±84.38	1.61±130.12	0.94±34.61
RT_{bot}	2943.51±6763.51	1375.88±3724.72	1003.24±3455.20
$Like_{bot}$	0.87±14.59	0.73±12.33	0.89±20.84

forth an interesting insight into two distinct types of user reactions on emotional bot-generated content – thereby also indicating a potential for bot accounts to emotionally manipulate Twitter users (note that in our data-set no neutral bot-generated tweet received more attention than emotionally-neutral tweets generated by human accounts).

5 DISCUSSION

Our findings suggest that people, unlike bots, tend to conform to the base emotion of an event. This finding is in line with "offline" social studies (Heath, 1996) which showed that people tend to pass along (word-of-mouth) positive messages during positive events and negative during negative events. By comparing our findings to the ones in the related work (see Section 2), we cannot confirm the findings indicated in (Dickerson et al., 2014), according to which people disagree more with the base sentiment of a particular event. One possible explanation for this result is that for our study we analyzed tweets from 24 different events while the authors of (Dickerson et al., 2014) studied a single event (the 2014 Indian election) only.

As noted in (Everett et al., 2016), bots may try to deceive human users by diverging the sentiment conveyed in their tweets from the base emotion of the respective event. Thus, the observation that bots differ from human accounts by not complying to the base sentiment of the event may be explained by a bot's agenda to deceive human users. In fact, we also observed a deviation from the base sentiment in positive and polarizing events. More specifically, we found that bots tend to be more positive during polarizing events and more negative during positive events.

Identifying emotionally polarizing bots during polarizing events can serve as an indicator for an attempt of opinion swaying, i.e. the bots pick a side that they promote in their tweets. In our study, we have shown that bot-generated retweets have an impact on the perceived emotionality in the Twitter discourse. In particular, our results indicate that in negative events positive emotions are amplified by the effects of retweets, similar to an amplification of positive emotions during polarizing events.

Below we show examples of bot messages with positive emotion scores during polarizing events. Related to the 2016 US presidential elections, the bots in our data-set disseminated messages such as:

- "*#ObamaFail I'll be so happy to see this joke move out of the White House!! #VoteTrumpPence16*",
- "*So proud of my daughter! She just voted for @realDonaldTrump #Millennials4Trump #Women4Trump #VoteTrumpPence16 #America*".

While related to the 2016 Austrian presidential elections, bots disseminated messages such as:

- "*Save your country, take back control and stop Islamisation. We support Austria's Hofer in tomorrow's election. #bpw16*".

Since messages that support one particular candidate make up the vast majority (99.37%) of bot-generated tweets in the subset containing positive messages (87.8% of those messages are retweets), we can conclude that bots clearly have a strategic agenda during polarizing events.

Our findings further indicate that bots tend to spread (retweet) more negative emotions during positive events (86.41% retweets) as compared to humans. Observing the bot-generated tweets in our data-set, the respective messages tend to either 1) express a negative opinion about a specific topic, such as:

- "*This was not as good as the last one. It's hard to ink when there is a lot of black #FantasticBeasts*",
- "*That explains the retarded haircut. I hate his mother even more. #FantasticBeasts*",
- "*Nico Rosberg articulates the F1 season and his resignation but offers no real clues as to why #NicoRosberg*",

or 2) surprise the prospective readers by injecting negative content about a particular subject. For example, the human-generated Thanksgiving tweets in our data-set are predominantly positive, while bots injected topic-wise unrelated negative tweets carrying a Thanksgiving hashtag, e.g.:

- "*Sissy Mitt Romney signed Massachusetts gun ban #thanksgiving #Trump #MAGA*".

6 CONCLUSION

We systematically collected 4.4 million tweets related to 24 real-world events that are either positive (e.g., birthday of a celebrity), negative (e.g., terror attacks), or polarizing (e.g., political campaigning). For each of the 4.4 million tweets we applied an emotion-extraction procedure (Kušen et al., 2017a) that provides intensity scores for the eight basic emotions according to Plutchik's wheel of emotions. In total, the tweets in our data-set have been generated by 1.3 million unique user accounts, 35.2 thousand of which were identified as bots via DeBot (Chavoshi et al., 2016b; Chavoshi et al., 2016a; Chavoshi et al., 2017b).

In order to examine how bots and human accounts compare to each other in terms of the emotions they spread, we examined the relative presence of emotions, as communicated by both types of accounts. Moreover, since previous studies have shown that bots tend to exhibit a higher retweet frequency than human accounts, we adjusted the scores for the effects of retweets. Our findings suggest that, in general, humans conform to the base emotion of the respective event, while bots contributed to the higher intensity of shifted emotions (e.g., negative emotions during positive events). Our study shows that bots tend to shift emotions in order to receive more attention (in terms of likes or retweets) and that they often follow a specific agenda. In particular, we showed that bots tend to emotionally polarize during controversial events (such as presidential elections). Furthermore, we found that bots inject shifted emotions into topic-wise unrelated Twitter discussions (e.g., messages related to the 2016 US presidential election that include Thanksgiving hashtags). Given such observations, emotions sent by an OSN account may serve as a valuable indicator of automated accounts.

We also performed a time-series analysis and identified temporal patterns which distinguish human and bot accounts in terms of their emotionality. Finally, we showed that emotional bot-generated tweets tend to get more likes/retweets if a bot spreads a tweet conveying shifted emotions (i.e. emotions that do not comply with the base emotion of the respective event). Interestingly, emotionally neutral tweets authored by bots did not receive much attention in terms of likes and retweets.

In our future work, we plan to further investigate the effects of basic as well as derived emotions on the diffusion of information in OSNs. Moreover, we plan to investigate whether the same patterns found on Twitter hold for other OSN platforms, such as Facebook.

ACKNOWLEDGEMENTS

We thank Nikan Chavoshi from the Department of Computer Science, University of New Mexico, for her kind assistance with the bot detection via DeBot.

REFERENCES

Abokhodair, N., Yoo, D., and McDonald, D. W. (2015). Dissecting a social botnet: Growth, content and influence in Twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW'15, pages 839–851, New York, NY, USA. ACM.

Bessi, A. and Ferrara, E. (2016). Social bots distort the 2016 US presidential election online discussion. *First Monday*, 21(11).

Chavoshi, N., Hamooni, H., and Mueen, A. (2016a). Debot: Twitter bot detection via warped correlation. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 817–822.

Chavoshi, N., Hamooni, H., and Mueen, A. (2016b). Identifying correlated bots in Twitter. In Spiro, E. and Ahn, Y.-Y., editors, *Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part II*, pages 14–21, Cham. Springer International Publishing.

Chavoshi, N., Hamooni, H., and Mueen, A. (2017a). On-demand bot detection and archival system. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, pages 183–187, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Chavoshi, N., Hamooni, H., and Mueen, A. (2017b). Temporal patterns in bot activities. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, pages 1601–1606, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. (2010). Who is tweeting on Twitter: Human, bot, or cyborg? In *Proceedings of the 26th Annual Computer Security Applications Conference*, ACSAC '10, pages 21–30, New York, NY, USA. ACM.

Dickerson, J. P., Kagan, V., and Subrahmanian, V. S. (2014). Using sentiment to detect bots on Twitter: Are humans more opinionated than bots? In *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pages 620–627.

Eagle, N. and Pentland, A. (2006). Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268.

Everett, R. M., Nurse, J. R. C., and Erola, A. (2016). The anatomy of online deception: What makes automated text convincing? In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, SAC '16, pages 1115–1120, New York, NY, USA. ACM.

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22(8):1–33.

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. (2016). The rise of social bots. *Commun. ACM*, 59(7):96–104.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *The American Psychologist*, 56:218–226.

Gilani, Z., Farahbakhsh, R., and Crowcroft, J. (2017). Do bots impact Twitter activity? In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, pages 781–782, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Heath, C. (1996). 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*, 68(2):79–94.

Howard, P., Duffy, A., Freelon, D., Hussain, M., Mari, W., and Mazaid, M. (2011). Opening closed regimes, what was the role of social media during the Arab Spring? *Project on Information Technology and Political Islam*, pages 1–30.

Kim, H. S., Lee, S., Cappella, J. N., Vera, L., and Emery, S. (2013). Content characteristics driving the diffusion of antismoking messages: Implications for cancer prevention in the emerging public communication environment. *Journal of National cancer institute. Monographs*, 47:182–187.

Kollanyi, N., Howard, P., and Woolley, S. (2016). Bots and automation over Twitter during the U.S. election. *Data Memo. Oxford, UK: Project on Computational Propaganda*, 2016(4):1–5.

Kramer, A. D. I., Guillory, J. E., and Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790.

Kušen, E., Cascavilla, G., Figl, K., Conti, M., and Strembeck, M. (2017a). Identifying emotions in social media: Comparison of word-emotion lexicons. In *Proc. of the 4th International Symposium on Social Networks Analysis, Management and Security (SNAMS) (co-located with IEEE FiCloud 2017)*. IEEE.

Kušen, E., Strembeck, M., Cascavilla, G., and Conti, M. (2017b). On the influence of emotional valence shifts on the spread of information in social networks. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 321–324. ACM.

Mascaro, C., Agosto, D., and Goggins, S. P. (2016). One-sided conversations: The 2012 presidential election on Twitter. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research*, dg.o '16, pages 112–121, New York, NY, USA. ACM.

Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89(4).

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., and Menczer, F. (2011). Detecting and tracking political abuse in social media. In *Proc. 5th International AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 297–304.

Savage, S., Monroy-Hernandez, A., and Höllerer, T. (2016). Botivist: Calling volunteers to action using online bots. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 813–822, New York, NY, USA. ACM.

St Louis, C. and Zorlu, G. (2012). Can Twitter predict disease outbreaks? *BMJ*, 344.

Thai, M. T., Wu, W., and Xiong, H., editors (2016). *Big Data in Complex and Social Networks*. CRC Press, Taylor & Francis Group.

Tsugawa, S. and Ohsaki, H. (2015). Negative messages spread rapidly and widely on social media. In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, pages 151–160.

Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R., and Herbst, K. (2005). Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. *PLoS Medicine*, 2(10).