Online Social Networks and Media 10–11 (2019) 1–17

Contents lists available at ScienceDirect

Online Social Networks and Media

journal homepage: www.elsevier.com/locate/osnem

Something draws near, I can feel it: An analysis of human and bot emotion-exchange motifs on Twitter

Ema Kušen[a], Mark Strembeck[a,b,c,*]

[a] Vienna University of Economics and Business (WU Vienna), Austria
[b] Secure Business Austria Research Center (SBA), Austria
[c] Complexity Science Hub Vienna (CSH), Austria

ARTICLE INFO

Article history:
Received 16 January 2019
Revised 4 April 2019
Accepted 17 April 2019

Keywords:
Emotion analysis
Emotion-exchange motifs
Network motifs
Social bots
Social network analysis
Temporal patterns
Triadic patterns
Twitter

ABSTRACT

Social bots are software programs that automatically produce messages and interact with human users on social media platforms. In this paper, we provide an analysis of the emotion-exchange patterns that arise from bot- and human-generated Twitter messages. In particular, we analyzed 1.3 million Twitter accounts that generated 4.4 million tweets related to 24 systematically chosen real-world events. To this end, we first identified the intensities of the eight basic emotions (according to Plutchik's wheel of emotions) that are conveyed in bot- and human-generated messages. We then performed a temporal analysis of the emotions that have been sent during positive, negative, and polarizing events. Furthermore, we investigated the effects on user reactions as well as on the message exchange behavior between bots and humans. In addition, we performed an analysis of the *emotion-exchange motifs* that occur when bots communicate with humans. For this purpose, we performed a systematic structural analysis of the multiplex communication network that we derived from the 4.4 million tweets in our data-set. Among other things, we found that 1) in contrast to humans, bots do not conform to the base mood of an event, 2) bots often emotionally polarize during controversial events and even inject polarizing emotions into the Twitter discourse on harmless events such as Thanksgiving, 3) when bots directly exchange messages with human accounts they are, however, indistinguishable from humans with respect to the emotions they send, 4) direct message exchanges between bots and humans result in characteristic and statistically significant *emotion-exchange motifs*.

1. Introduction

Currently about 2.46 billion individuals use social media, with a predicted increase to 3.02 billion until 2021[1]. In general, social media users can post their own content (such as text messages, pictures, or videos) and also react to posts sent by other users (e.g. by liking, retweeting, or replying). In recent years, online social networks (OSNs) have become a valuable source of data for studying human behavior and detecting patterns in message dissemination during various real-world events [28]. For example, Colleoni et al. [15] analyzed political orientation of the US voters while Procter et al. [59] studied the 2011 riots in England. A recent study by Warner-Soderholm et al. [73] found that the users of OSNs in general trust the news that spread among their OSN contacts and that they often disclose own private information to their contacts (such as hobbies, friendship relations, opinions, or political affiliations) [72]. Moreover, Yan and Jiang [76] showed that group influence is an important factor for deciding which message receives more attention within an OSN community. In addition to helping people stay in touch with friends and family, OSNs also provide support during natural disasters [66] or when organizing political movements [33]. Moreover, numerous studies have also pointed to the use of OSNs for political campaigning, such as the use of Twitter during the 2014 elections to the European Parliament [54], the 2016 Austrian presidential elections [44], or the 2016 US presidential elections [48].

In this context, social bots may pose a threat to human users [26] because of their potential to manipulate and steer the opinions of human users. For example, Kollanyi et al. [37] indicated that even a single bot may flood Twitter users with messages. Aside from an increased message load, a high volume of bot-generated content also has the potential to negatively affect public

* Corresponding author at: Vienna University of Economics and Business (WU Vienna), Austria.
E-mail address: mark.strembeck@wu.ac.at (M. Strembeck).

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

https://doi.org/10.1016/j.osnem.2019.04.001
2468-6964/

opinion [8]. Furthermore, Ferrara et al. [26] argued that bots may be responsible for altering a user's perception in the Twitter-sphere or even for destroying a user's reputation. In recent years, bots became more sophisticated [75] and may convincingly mimic human behavior [24,26,77]. Thus, studying the behavior of bot accounts as well as their influence on information dissemination in OSNs has become an important research topic.

Kim et al. [36] and Tsugawa and Ohsaki [71] pointed to the importance of studying human emotions conveyed in OSN messages and showed that emotions affect the diffusion rate of messages. In this context, Kramer et al. [38] empirically showed that emotions are transferable among OSN users and may lead to emotional contagion. Thus, the study provided empirical evidence for the importance of written cues (OSN posts) for influencing one's emotional state. This finding supplements the traditional assumption that, predominantly, physical non-verbal cues (such as one's body language and facial expressions) and inter-personal interactions are responsible for the emotion transfer among people.

This paper is an extended version of [45] and is a contribution to our ongoing work concerning the influence of emotions on OSN user behavior [see, e.g., [40–44,46]]. In this paper, we focus on the commonalities and differences of emotional messages that have been sent by bots and human users. In particular, we analyze a data-set consisting of 4.4 million Twitter messages annotated with respect to the presence and intensity of the eight basic emotions identified by Plutchik [58]. The messages in our data-set have been sent by 1.3 million distinct Twitter accounts, 35.2 thousand of which have been identified as bots via DeBot [12].

In our analysis, we show that human and bot accounts differ in their behavioral patterns with respect to the emotions they communicate. In this context, we show that humans conform to the base emotion of an event (e.g., express sadness or fear during negative events, or joy during positive events), while bot accounts spread a more heterogeneous set of emotions. The distinction between bot and human behavior is especially evident during polarizing events (such as political elections), where bot accounts seem to intentionally choose sides and strategically try to influence human users.

However, we also found that when bots directly exchange messages with human accounts they are indistinguishable from humans with respect to the emotions conveyed in their messages. However, our findings also suggest that bots and humans still behave differently with respect to the *emotion-exchange motifs* that are formed when bots and humans exchange emotional messages. We use the term *emotion-exchange motif* [see 39] for statistically significant and over-represented communication patterns that arise when OSN users (humans or bots) exchange emotional OSN messages.

The remainder of this paper is organized as follows. Section 2 first provides an overview of related work. Next, Section 3 outlines our research approach and Section 4 presents the results of our analyses. Section 5 then provides a discussion of our findings and Section 6 concludes the paper.

2. Related work

Current bot detection techniques predominantly rely on a set of carefully chosen features. For example, Botometer [17] uses Twitter's REST API to fetch a user's recent tweeting behavior and applies a specifically trained classifier to compute the likelihood of an account being a bot. The classifier uses more than 1000 features such as language used by an account, a user's geographic location, account creation date, number of followers and followees, temporal content generation, and the type of content that has been posted. Another tool, called DeBot [12], examines the synchronicity of tweeting behavior among Twitter users for a certain duration of time. If users exhibit a synchronous behavior, they are more likely to be bots.

Given such tools, research has provided valuable insights into bot-like behavior. One of the most distinctive features between Twitter bots and human accounts results from the tendency of bots to send a comparatively higher number of tweets per unit of time [2,13,37,50,56]. Recently, however, Chavoshi et al. [12] showed that Twitter bots may mimic a human-like tweet generation rate and may also delete some of their tweets. Twitter bots usually differ from humans in their follower-followee ratio. For example, while analyzing a sample of 500,000 Twitter accounts, Chu et al. [13] showed that bots attract generally only a few followers but tend to follow a large number of users themselves.

As for their role on Twitter, multiple studies have shown that bots actively try to influence Twitter users. For example, Chu et al. [13] and Gilani et al. [29] have shown that bots tend to (re)tweet specific URLs, use particular hashtags, or even directly mention target users in their tweets (via @screenname) to attract more attention to the bot-disseminated content. Savage et al. [63] subsequently indicated that bot accounts may trigger a discussion among human users on the topics that before have been injected by these bots. As pointed by Stefanie et al. [67], bot accounts may deceive human users by imitating human-like behavioral patterns and building trust relationships with human users.

Especially alarming is a finding by Edwards et al. [22], who indicated that bot accounts may even be perceived as credible sources in OSNs. In this context, multiple studies discussed the potential dangers that bot account might pose to real-world events such as democratic elections. As shown in [25,67], bots have a potential to sway voters' opinions, spread misinformation, or even amplify the influence of a particular political candidate in OSNs. Ratkiewicz et al. [60] and Kollanyi et al. [37] examined the role of bots with respect to US politics and found that bots may strategically boost the dissemination of particular messages (e.g., that favor a particular political candidate). Moreover, Llewellyn et al. [49] showed that bots actively used trending hashtags (such as #Brexit) to promote specific political messages.

Since OSNs are predominantly designed to be online social networks of people, human emotions and sentiments conveyed in OSN messages play a vital role in inspiring and motivating various interactions among human users (e.g., liking, commenting, resharing a post). However, a limited number of studies has thus far focused on the differences between human and bot accounts with respect to emotions conveyed in their messages. While studying the 2014 Indian elections, Dickerson et al. [19] found that humans disagree more with the base sentiment of the event, as compared to bots. Everett et al. [24] further investigated the impact of such shifted ("disagreeing") sentiments conveyed in bot-generated messages on the formation of a crowd opinion. The authors indicated that such messages are deceitful and can be used by bots to mimic human-like behavior.

Therefore, a topic of particular interest is the analysis of emotion-exchange patterns that discriminate bots from humans. In recent years, the concept of network motifs (see Section 3) has become a prominent approach for identifying patterns in networks. Thus far, network motifs [51] have predominantly been used to study structural patterns in various types of biomedical networks, such as the structure of metabolic networks [6] or gene regulation networks [3]. Lately, network motifs have also found their application in studying the structure of online social networks. In particular, such studies predominantly focus on detecting communication patterns [see, e.g., [4,14,31,57]] and the formation of friendship networks [see, e.g., [21,62,70]].

Regarding the motifs that emerge as social media users communicate with each other, Coletto et al. [14] studied dyadic motifs that arise when a pair of users discusses controversial as well as

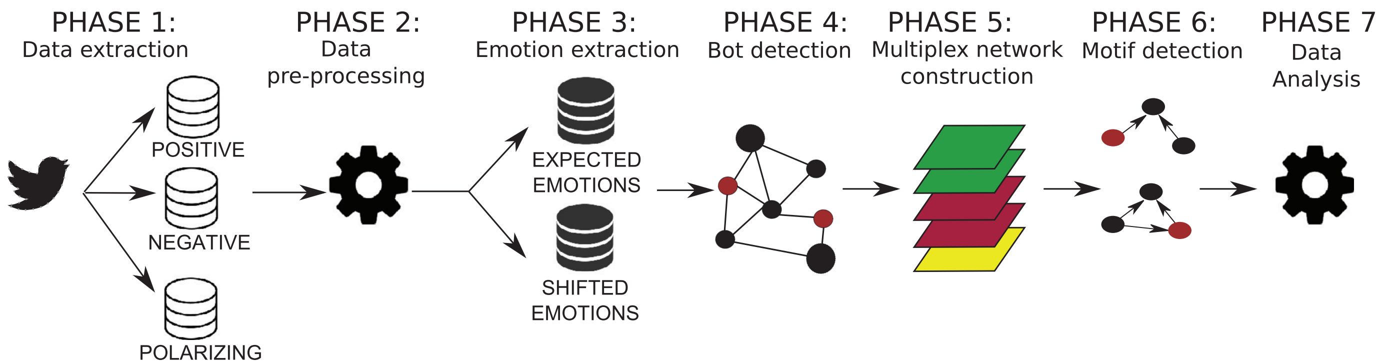


Fig. 1. Research procedure.

non-controversial topics. Interestingly, the discussion networks derived from controversial and non-controversial data-sets exhibited distinctive motifs which actually differentiate controversial from non-controversial discussion patterns. Barash et al. [4] used motifs to study network structures that support rumor dissemination on Twitter. Some studies also investigated the motifs that are characteristic for online social media communication as compared to e-mail communication. For example, Paranjape et al. [57] study so-called blocking motifs which represent communication patterns where a node is blocked until it receives a response to a previous message. They found that such blocking motifs are more characteristic for the communication patterns on Facebook compared to e-mail communication.

Zhao et al. [78] investigated motifs arising from a city's phone call records and compared them with those found in Facebook postings. As a result, they identified chain motifs as representative patterns for phone call records, while Facebook postings showed a comparatively higher presence of star-like motifs. Finally, Gurukar et al. [31] examined temporal aspects of message sending behavior on Twitter and Facebook and found motifs that are characteristic for both social media platforms. In particular, they found that Twitter exhibits a high number of motifs where one user, such as a celebrity, is mentioned frequently for a short period of time (the phenomenon of *burstiness* on Twitter [34]). In contrast, a longer and more frequent message-sending behavior between the same set of users seems to be more characteristic for communication via Facebook.

Some other studies indicate that motifs can be also used to classify various online social networking platforms (such as Facebook, Twitter, and Google Plus) according to the distributions of the motifs that are found within the respective platform's friendship networks [see, e.g., 21,70]. Moreover, Rotabi et al. [62] used motifs to examine strong ties in an undirected communication network derived from Twitter.

Even though a couple of papers study motifs in an OSN context (see above), the application of network motifs for studying behavioral patterns that result from interactions of different types of users/agents is still considerably understudied. In our previous work on emotion-exchange motifs, we especially investigated motifs that emerge as bots and humans exchange messages during riot events, i.e. a class of events that predominantly produces negative emotions (see [41,42]).

3. Research procedure

Our research procedure included seven main phases (see Fig. 1).

Phase 1: Data extraction. We systematically collected 4,418,655 tweets related to 24 events that can be classified either as positive (e.g., public holiday, release of a movie, birthday of a pop star), negative (e.g., a natural disaster, acts of war), or polarizing (e.g., political elections, death of a controversial political figure) (see Table 1).

Table 1
Events analyzed in our study.

Domain	Event	Number of tweets
Polarizing	(N=1,812,573; 41%, RT=73.90%)	
Politics	1) Death of Fidel Castro	720,548
	2) 2016 Austrian presidential elections	2558
	3) 2016 US presidential elections	891,425
Pop culture	4) The Walking Dead season 7 premiere	198,042
Positive	(N=1,115,587; 25%, RT=68.88%)	
Sports	5) Rosberg winning Formula 1	215,703
	6) Murray winning ATP	62,184
	7) Rosberg retirement message	34,201
Pop culture	8) "Beauty and the Beast" trailer release	138,979
	9) "Fantastic beasts" trailer release	64,264
	10) ComiCon Vienna	704
	11) Miley Cyrus birthday	76,270
	12) New Pentatonix album released	9341
	13) Ellen Degeneres medal of freedom	73,854
Other	14) Thanksgiving	440,087
Negative	(N=1,490,495; 34%, RT=76.38%)	
Politics	15) Erdogan's threats to EU	804
	16) US anti-Trump protests	381,982
Pop culture	17) Death of Leonard Cohen	89,619
	18) Death of Colonel Abrams	1253
War & terrorism	19) Aleppo bombings	995,561
	20) Seattle shooting	73
Other	21) Lufthansa strike	3387
	22) Ransomware in Seattle	2564
	23) Yellowstone incident	15
	24) Earthquake in central Italy	15,237

For our data extraction, we used Twitter's Search API[2] which returns a number of tweets based on a pre-defined search query that includes one or more search (the full list of search terms that we used is provided in Table A.13 in Appendix). In particular, we queried Twitter by using carefully selected hashtags for each of the 24 events [see also 46]. For each event, we extracted tweets published within one week since the event's announcement/occurrence[3] and restricted the extraction to tweets written in English language only. To this end we used the "*lang*" language parameter which is provided by Twitter's API and takes an ISO 639-1 code to restrict the extracted tweets to a particular language. In total, it took three months to collect the tweets related to the 24 events in our study (October 2016 - December 2016).

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

[3] Note that some events we considered for our study 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 related to the corresponding event. For example, we extracted tweets related to the 2016 Austrian presidential elections published one week before the actual election date.

Phase 2: Data pre-processing. After obtaining the data-set, we conducted several pre-processing steps. For example, we removed duplicate entries and information irrelevant with respect to emotion extraction, such as URLs and HTML tags.

Phase 3: Emotion extraction. In this phase, 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, for details see Algorithm 1 and [40]) and results in an own vector of emotion 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). In our previous work [see 40], we tested the accuracy of our procedure on a sample of 7691 short texts that contain features which are characteristic for online social media texts (such as typos, abbreviations, and smileys), achieving the accuracy of 0.85 (F-measure).

Compared to related approaches [see, e.g., 1,69] that detect emotions belonging to Plutchik's wheel, our approach provides an intensity score for each of the eight basic emotions for each tweet. In contrast, Abdul-Mageed and Ungar [1] identified the presence of a dominant emotion in a tweet by restricting their tweets to only those that include at least five words and an emotional hashtag at the end of a tweet. After deploying a Gated Recurrent Neural Network (GRNN), they achieved an overall accuracy of 0.96 (F-score). Suttles and Ide [69] reduced the classification complexity by drawing from Plutchik's theory of *emotional polar opposites* which postulates that a person can experience one of two polar emotions. According to Plutchik's wheel, polar opposite emotions are anger or fear, joy or sadness, anticipation or surprise, and disgust or trust. Thus, the authors deploy four independent binary classifiers to make binary decisions which opposing emotion would be most probable (e.g. sadness or joy). Similarly to [1], they consider hashtags and in addition emoticons and emojis for their emotion prediction task, achieving an overall accuracy of 0.75-0.91[4].

Phase 4: Bot detection. Next, we extracted the list of unique screennames (Twitter user names) from the tweets in our data-set. This list of screennames has then been analyzed via DeBot [9,10] to obtain a bot score for each Twitter account. DeBot examines correlated account activities in near real-time [9] and searches for synchronized behavior [11] to distinguish between bot and human accounts. DeBot has been validated for its accuracy in a number of empirical evaluations [see, 11]. In a human evaluation, the bots identified by DeBot achieved a 94% agreement with the assessment of human judges [see, 9]. Moreover, 45% of the accounts flagged as bots by DeBot were subsequently suspended by Twitter.

In total, we processed 1,317,555 Twitter accounts, 35,247 of which were identified as bots by DeBot. This gave us an overall percentage of 2.67% bot accounts in our data-set.

Phase 5: Multiplex network construction. After bot detection, we reconstructed the corresponding communication network by following the @-traces. On Twitter, two users can directly communicate with each other by mentioning their respective screennames preceded by the @ symbol (e.g., @TwitterUser). To characterize the communication patterns of bot and human accounts, we subset the communication network with respect to edges that are formed when a bot node communicates with a human node. We call this subset the *bot-focused network*. For our analysis, we re-constructed three bot-focused networks from our positive, negative, and polarizing event data-sets, respectively. Table 2 shows some basic structural information for each of the three bot-focused networks.

Algorithm 1: Emotion extraction.

```
Data: C, Lexicon_dict, Secondary_dicts
Result: Emotion vector for each c ∈ C

C_clean = [ ];
Emoticons = [ ];
Lemmatized = [ ];
foreach c_i ∈ C do
    /* clean unnecessary HTML, tags */
    C_clean.append(RemoveCode(c_i));
end
foreach c_i ∈ C_clean do
    /* Extract emoticons */
    Emoticons.append(CodeSmiley(c_i));
end
foreach c_i ∈ Emoticons do
    /* Lemmatize the short texts */
    Lemmatized.append(Lemmatizer(c_i));
end
dict_emot_v = { };
foreach c_i ∈ Lemmatized do
    foreach sent_q ∈ sentiments do
        dict_emot_v[i][sent_q] = 0
    end
    /* Split text to sentences */
    dict_emot_v[i] = { };
    sentences = SplitToSentence(c_i);
    foreach s_j ∈ sentences do
        /* Split sentence to words */
        words = SplitToWord(s_j);
        foreach w_k ∈ words do
            /* Find lexicon match */
            if w_k ∈ Lexicon_dict then
                /* Update emotion score corresponding to w_k */
                dict_emot_v[i][Lexicon_dict[w_k][0]]+ = [Lexicon_dict[w_k][1];
                /* Find negation and intensifiers */
                foreach w_r ∈ [w_{k-3} : w_{k-1}, w_{k+1} : w_{k+3}] do
                    foreach d_t ∈ Secondary_dicts do
                        /* Updating emotions related to w_k */
                        /* using negation and intensifiers */
                        if w_r ∈ d_t then
                            dict_emot_v[i][Lexicon_dict[w_k]]+ = d_t[w_r];
                        end
                    end
                end
            end
        end
    end
end
return dict_emot_v ;
```

[4] Note, however, that the design of a user study with human annotators considerably impacts the perceived overall precision of the emotion detection task. In [1], human annotators had to either agree or disagree that an automatically assigned emotion is relevant for a tweet, while Suttles and Ide [69] limit the annotators to two pre-selected polar emotions, "neutral" and "don't know" as possible answers. In contrast, the evaluation of our approach relied on Schimmack's method for affect measurement [see, 64] according to which we asked annotators of different ages, different genders, and cultures (Western and Eastern Europeans, Americans, Australians and Asians) to first identify whether an emotion is present in a tweet and then assign its intensity without having an insight into the scores assigned by our algorithm.

Table 2
Information about the three bot-focused networks: number of nodes and edges, network density, average degree μ and its variance *var*, as well as the number of weakly connected components including the average number μ and the standard deviation *sd* of the members per components.

Domain	Nodes	Edges	Density	Degree	Connected components
Polarizing	7164	8735	0.0002	$\mu = 1.22,\ var = 276.71$	752, $\mu = 9.53,\ sd = 190.10$
Positive	2250	2448	0.0004	$\mu = 1.09,\ var = 33.67$	449, $\mu = 5.01,\ sd = 34.79$
Negative	6789	10538	0.0002	$\mu = 1.55,\ var = 198.21$	559, $\mu = 12.14,\ sd = 229.39$

Next, we constructed a multiplex network consisting of eight layers, each of which representing a sub-network of nodes communicating a specific emotion. A *multilayer social network* is formally defined as a quadruple $M = (\mathcal{U}, \mathcal{L}, N, E)$ where $\mathcal{U}$ is a set of social network users, $\mathcal{L}$ a set of layers, E a set of edges, and $N \subseteq \mathcal{U} \times \mathcal{L}$ [20]. A multiplex network is then a multilayer network in which all layers include the same set of vertices (users) but on each layer the vertices are connected via a different type of edges (i.e. edges representing a particular type of relationship, such as emotions). Our multiplex network counts three positive layers (joy, trust, anticipation[5]), four negative layers (anger, fear, disgust, sadness), and an additional layer representing surprise, which cannot be classified as neither positive or negative. By utilizing a multiplex network model, we are able to capture multiple types of relations (i.e. different types of emotions communicated between pairs of users) and embed such relations via emotion-annotated edges in different layers. Subsequently, this representation of emotion-annotated edges organized in layers brings advantages for constructing a number of isolated simplex networks. This property allows us to examine patterns that emerge as Twitter users communicate emotions that belong to a specific combination of layers (as described in *Phase 6: Motif detection*).

Phase 6: Motif detection. A *network motif* is a statistically significant subgraph pattern that is over-represented in a real-world network as compared to the subgraph patterns that emerge in synthetically generated networks which have similar characteristics [51]. The edges in the network we analyze carry a semantic meaning. In particular, each edge represents a message which conveys one of the eight emotions considered in this study (anger, fear, sadness, disgust, joy, trust, anticipation, surprise). Therefore, we use the term *emotion-exchange motifs* to refer to the corresponding subgraph patterns [see also 39].

In order to detect emotion-exchange motifs, we applied a subgraph enumeration procedure so that we can later analyze the extent to which bot nodes are involved in a motif. In particular, we applied the ESU enumeration algorithm [74] to enumerate all possible subgraphs of size k (see Algorithm 2). For the analyses conducted in this paper we set k to 3, i.e. we investigate all subgraphs consisting of three nodes. Next, we performed an isomorphism test on all enumerated subgraphs by using the VF2 algorithm [16]. To limit the test space in our subgraph pool, we categorized the subgraphs into smaller groups based on their degree sequence (i.e., we pre-identified possible isomorphic candidates). Afterwards, we applied the VF2 algorithm to assign all enumerated subgraphs to an isomorphism class.

Moreover, in order to determine if a subgraph is statistically significant and over-represented in a real-world network, we have to ensure that it does not emerge by chance. A common approach to ensure this property is to generate a set of *null models*[6], and test if the subgraphs found in the null models appear significantly less frequent than in the respective real-world network [23,51,65]. If a subgraph is statistically significant for our real-world network, we can safely assume that it emerges due to specific communication patterns among bot and human accounts rather than by chance.

We applied the motif detection procedure described above over (see also Fig. 2):

(a) each of the eight emotion layers individually,
(b) two aggregated valence layers for positive and negative emotions,

Algorithm 2: Motif detection.

```
Input: input_network;
Output: list_of_motifs = [];
Initialize: i = 0;
# ENUMERATE AND CLASSIFY subgraphs
def procedure: esu_vf2(list_layers)
foreach l in list_layers do
    subgraphs = esu(l)
    dss = degree_set(subgraphs)
    for dss do in parallel
        foreach s in subgraphs do
            subgraphs' = subgraphs \ s
            foreach s' in subgraphs' do
                if vf2(s, s') then
                    assign_common_isomorphism_class
                    subgraphs' = subgraphs' \ s'
                    subgraphs = subgraphs \ s'
                end
            end
        end
    end
end
end procedure
# GENERATE LAYERS AND INTER-LAYERS
detect layers in input_network
layer_negative.add_edges_from(layer_anger, layer_sadness,
layer_disgust, layer_fear)
layer_positive.add_edges_from(layer_joy, layer_anticipation,
layer_trust)
foreach i in range(length(V(input_network))) do
    if v_i ∈ V(layer_negative) & v_i ∈ V(layer_positive) then
        inter_layer.add_edges_from(layer_negative.edge_containing(v_i),
        layer_positive.edge_containing(v_i))
    end
end
list_layers = [layer_anger, layer_joy,..., layer_surprise, layer_negative,
layer_positive, interlayer, input_network]
esu_vf2(list_layers)
# GENERATE NULL MODELS
while i < 1000 do
    foreach l ∈ list_layers do
        null[l] = matching(l.in_degree(), l.outdegree())
    end
    esu_vf2(null)
    i = i+1
end
```

[5] We classify *anticipation* as a positive emotion because Spearman's correlation coefficient ρ indicated a strong correlation between anticipation and positive emotions (joy and trust); and only a weak correlation with negative emotions (anger, fear, sadness, and disgust). For instance, for tweets found in the positive events data-set, anticipation strongly correlated with trust $\rho = 0.69$, but only weakly with fear $\rho = 0.31$. We observed the same pattern for negative and polarizing events. *Surprise*, however, did not exhibit a strong correlation with either positive or negative emotions. Therefore, *surprise* is placed in a separate category in this paper.

[6] A *null model* is a synthetically generated random network which resembles the real-world network that we are analyzing.

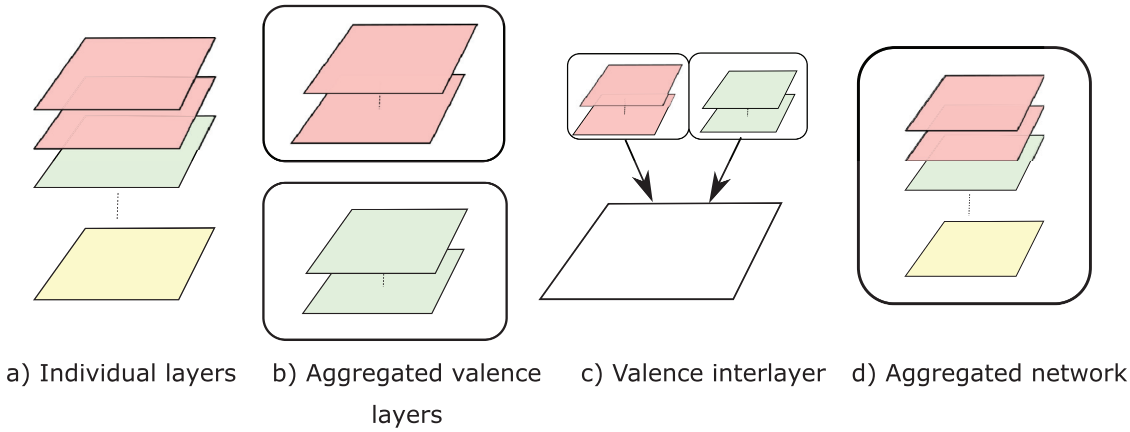


Fig. 2. Layers used in our analyses (green layers represent positive emotion layers, red represent negative emotion layers, and yellow represents surprise). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(c) a valence interlayer (i.e. the edges shared between the set of common nodes in the two aggregated valence layers for positive and negative emotions),
(d) an aggregated network over all eight emotion layers.

To this end, we generated 1000 null models for each layer. For the null model generation we used the stub-matching algorithm [7,53], and then enumerated and classified all 3-node subgraphs in each of the null models. We therefore analyzed 12,000 null models for the positive, negative, and polarizing events respectively, resulting in 36,000 null models in total. Finally, we identified motifs by applying the Z-score measure and a corresponding p-value ($p < 0.05$) [23].

Since we are dealing with communication networks, the number of messages (i.e. edges) exchanged between a pair of vertices may also have a semantic meaning. For example, a higher number of messages exchanged between two nodes may signal a stronger personal relationship between two users [see, e.g., 68], while a large number of unanswered messages sent towards one specific user may indicate an attempt to spam the receiver [see, e.g., 2]. Thus, the subgraph enumeration procedure that we apply considers and stores all the edges between each pair of nodes.

Since all 3-node subgraphs with weighted edges can be generalized into one of the 16 possible isomorphic triadic states (see Fig. 3), we use the MAN labeling scheme [18] when reporting on our general results. Such labels consist of three digits with the first one standing for the number of mutual edges, the second one for the number of directed edges, and the third one for the number of missing edges between a pair of nodes. Furthermore, since different triads can have the same distribution of edges, the MAN labeling scheme further distinguishes between those that contain a cycle (C), are transitive (T), and whose edges have an upward (U) and downward (D) direction.

As some of the motifs we found in our analysis also include self-loops, they are not yet considered in the MAN labeling scheme. Thus, we introduce two new labels: L^R (root self-loop)[7] and L^L (leaf self-loop), and adjust the number of the directed edges in the MAN scheme accordingly, to uniquely identify tree-like motifs that contain self-loops.

The motif detection procedure described above has been performed on a machine with Intel Xeon CPU E3-1240 v5 @ 3.5GHz (4 cores/8 threads) and 32 GB RAM. On this machine the procedure took approximately four days to complete.

Phase 7: Data analysis. In the final step, we analyzed our dataset to study how emotions communicated by Twitter bots compare to emotions spread by human users. In particular, we are interested in the impact of bots on the diffusion of emotional content and especially in the role they play in emotion-exchange motifs that arise when they directly interact (communicate) with human users. Section 4 reports on our findings in four parts:

1. relative intensities for each of the eight basic emotions as conveyed by bots and human users (Section 4.1),
2. temporal patterns in tweeting of emotional content by bots and human users (Section 4.2),
3. user reactions on emotional tweets sent by bots and human users (Section 4.3), and
4. emotions conveyed in messages that are directly exchanged between human and bot accounts, as well as the functional role of bots in the corresponding emotion-exchange motifs (Section 4.4).

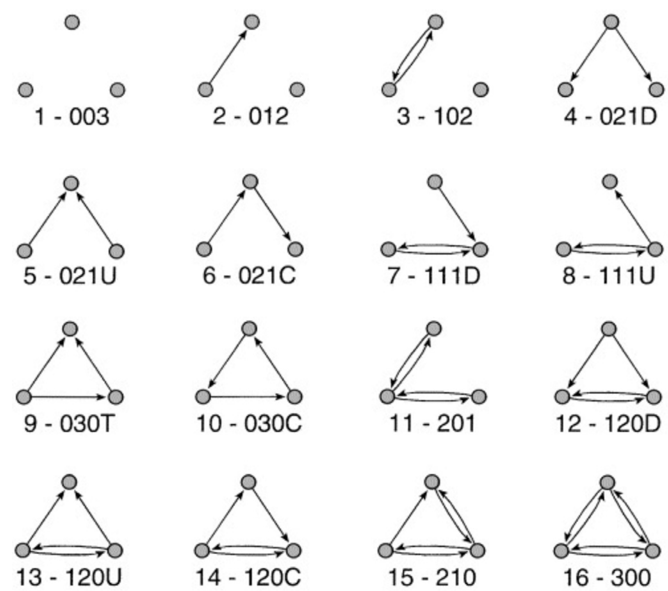


Fig. 3. 16 types of triads for a directed graph [5].

[7] In our labeling, there is no need to consider between a root of an in-tree and an out-tree. Conveniently, MAN labeling already considers the (anti-)arborescence of tree graphs by assigning the letters U and D to describe the edge direction in a tree.

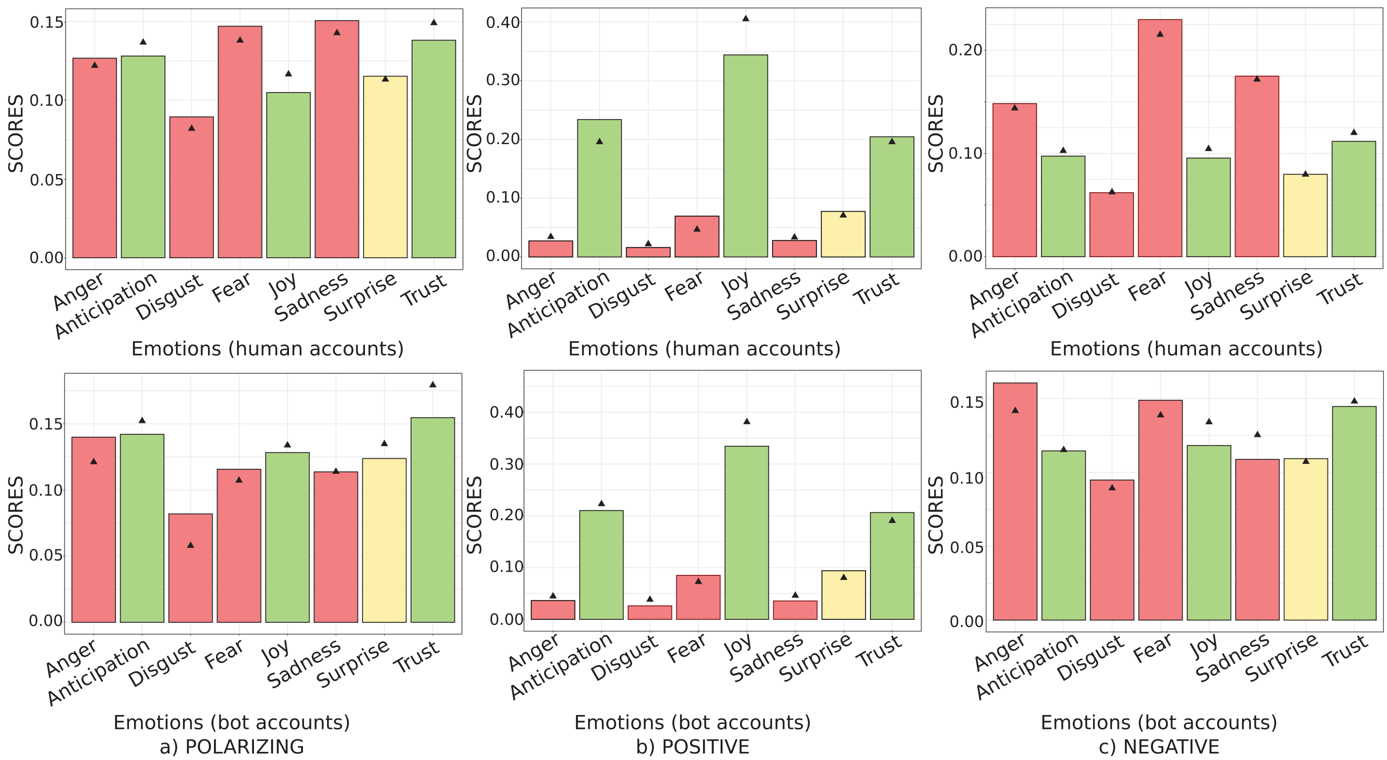


Fig. 4. Emotions expressed by human and bot accounts during polarizing, positive, and negative events. The effects of retweets are depicted via a black triangle respectively. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

4. Results

4.1. Intensities of emotions conveyed in bot- and human-authored tweets

While examining the intensities of the eight basic emotions, we found distinct differences between bot and human accounts. While humans tend to conform to the base emotion of an event, we found that bot accounts exhibit a more heterogeneous set of emotions. In particular, we quantify this difference by obtaining the difference score (d) between the sum of intensities of a particular group of emotions (according to their respective valence, i.e. positive or negative). We refer to the group of emotions that correspond to the base emotion of an event as the *expected emotions* (e.g., a message of joy that is sent during positive events), while *shifted emotions* refer to those emotions that do not comply with the base emotion of an event (e.g., a message of joy that is sent during negative events), see also [46,47]. Fig. 4 visualizes positive emotions (anticipation, joy, trust) in green and negative emotions (anger, disgust, fear, sadness) in red. *Surprise* (yellow) cannot be regarded as positive or negative emotion by default. We thus treat it as a context-dependent emotion. Fig. 4 presents relative emotion intensity scores where we averaged each specific emotion e over the sentence count S divided by the overall number of tweets in a data-set (N):

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

Our results revealed that during positive and negative events humans, in contrast to bots, show a larger difference (d) between the intensities of the expected and the shifted emotions. In particular, during negative events the difference between positive and negative emotions is $d_{n-p} = 0.192$, while bots exhibit a comparatively low score $d_{n-p} = 0.005$. A similar observation can be made during the positive events, where the difference between positive and negative emotions is $d_{p-n} = 0.666$ for human accounts and $d_{p-n} = 0.282$ for bot accounts.

As expected, during polarizing events humans and bots exhibit a comparable behavior by expressing both positive and negative emotions (see Fig. 4). Interestingly, however, though the overall expression of a set of emotions is similar, tweets generated by human accounts are more negatively inclined ($d_{n-p} = 0.0189$) compared to bots ($d_{n-p} = -0.102$). However, such a tendency of bots to express positive emotions implies that bots use positive emotions to lean towards one particular polarizing opinion. For example, a bot-generated tweet with a positive emotion score reads "*#ObamaFail I'll be so happy to see this joke move out of the White House!! #VoteTrumpPence16*", and clearly reveals a political preference of the corresponding bot account.

In addition to the difference score of expected and shifted emotions, we next correlate emotions conveyed in bot-authored tweets with those expressed by human accounts. To this end, we used Kendall's rank coefficient τ, which revealed that humans and bots disseminate comparative emotions during positive events ($\tau = 0.85$) and negative events ($\tau = 0.5$). However, polarizing events exhibit a noticeably larger distinction with the correlation falling to a weak positive ($\tau = 0.14$).

Since both humans and bots in our data-sets predominantly sent retweets (see Table 3), we further examine whether the retweets might have impacted the reported emotion scores. Thus, in the subsequent analysis, we observe only the unique occurrences of tweets (i.e., by excluding the retweets from each data-set).

Table 3
Presence of retweets in each data-set.

	Polarizing	Positive	Negative
Retweets (humans)	73.46%	68.80%	76%
Retweets (bots)	87.9%	80.37%	79.2%

Table 4
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**

When adjusted for the effect of retweets, during polarizing events, positive emotions (*joy, anticipation*, and *trust*) are amplified by retweets sent by human and bot accounts. *Surprise*, however, is only amplified by the retweets sent by bot accounts (see Fig. 4-a). During positive events (Fig. 4-b), we found that human- and bot-generated retweets intensified *joy* as well as some negative emotions (*anger, disgust*, and *sadness*), with *anticipation* being the only additional emotion intensified by bot-generated retweets. A similar observation can be made during negative events (see Fig. 4-c). Humans and bots amplify the intensities of positive emotions (*joy, trust*, and *anticipation*) via retweets. *Sadness*, however, is only intensified by bots during negative events.

We further examine whether such differences between bots and human accounts are statistically significant and define the following null hypothesis:

H_0: *There is no difference between the mean scores of the 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 conveyed in tweets authored by bots and human accounts, respectively. The t-test results are shown in Table 4 and point to a statistically significant difference (depicted in bold) in the intensities of emotions spread by bot accounts and human accounts in all three events (polarizing, positive, and negative). We therefore reject the null hypothesis.

In particular, bots communicated on average a higher intensity of negative emotions (*anger, disgust, sadness*, and *fear*) during positive events as compared to human accounts. We also found that bots do not comply with the positive base emotion during positive events – a trait which significantly distinguishes bots from human emotional reactions to positive events [32]. Our t-test results also indicate that bots tend to send more positive messages containing *joy, trust*, and *anticipation* during polarizing events compared to human accounts. However, as mentioned above, a positive emotion score does not necessarily imply that a message is positive but might also be biased towards one of the polarizing opinions (see also Section 5).

When adjusted for the effects of retweets (i.e. by considering the unique tweets only), we found that no particular emotion is more intensely expressed in unique tweets generated by bot accounts (see Table 4 without retweet (RT) entries). This confirms that bots especially tend to amplify certain emotions by retweeting.

4.2. Temporal emotion score patterns

In order to examine whether there are distinctive temporal patterns in emotion intensities conveyed in human- and bot-generated tweets, we contrasted the intensities of positive and negative emotions averaged over each day of the data extraction period.

The temporal progression of emotions sent during positive events (see Fig. 5) reveals that bots and humans exhibit comparable behavioral patterns. However, the same does not hold for negative and polarizing events. With respect to negative events, related studies have pointed to the noticeable presence of positive emotions that serve as a coping mechanism. Thus, such positive messages sent during negative events convey hope, gratitude, empathy, or comfort [see, e.g., [30,35,46,55]] and are generally explained via the so-called *undoing hypothesis* [27] (the human tendency to remain positive in order to undo the effects of negative emotions).

This effect is also evident in our negative events data-set. Positive emotions prevail over the negative ones on specific days (notice the positive emotion peaks in Fig. 6). This is especially noticeable for the positive emotions communicated by bots (see the black cross symbols in Fig. 6).

Finally, we investigate the temporal progression of emotions during polarizing events. Fig. 7 shows an expected presence of mixed emotions communicated by both bot and human accounts. Though, as previously noted, bots tend to incline towards a higher intensity of positive emotions as compared to humans.

4.3. Effects of emotional messages on user reactions

On Twitter, a user may react to a tweet by sharing it via sending a retweet or endorsing it with a like. In this section, we report on the effects of emotions conveyed in bot- and human-generated tweets on the corresponding retweets and likes. To this end, we

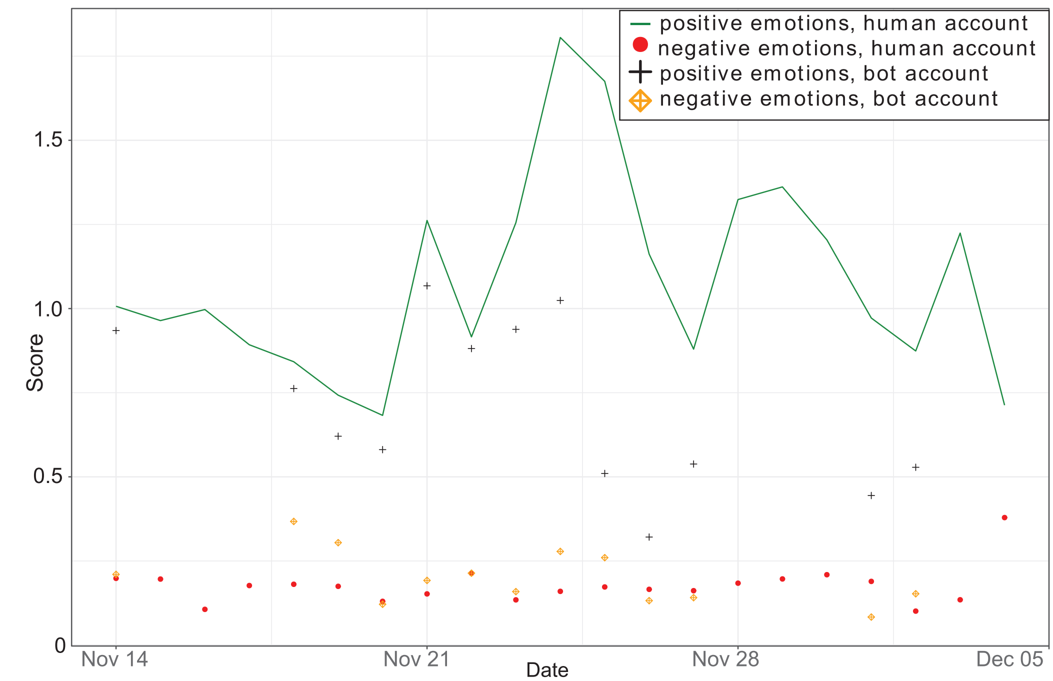


Fig. 5. Temporal emotion score patterns during positive events.

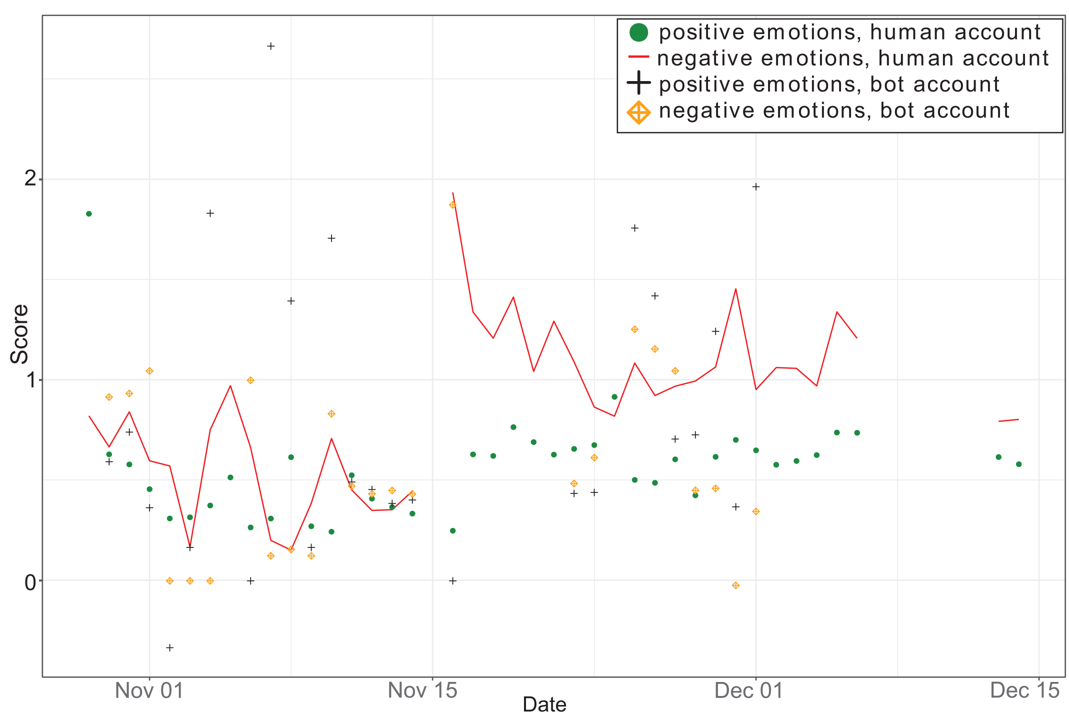


Fig. 6. Temporal emotion score patterns during negative events.

distinguish between the positive, negative, and emotionally neutral messages generated by bots and human accounts.

As shown in Table 5, human-generated tweets generally receive on average more retweets compared to bot accounts. In terms of likes, however, bots attract on average more likes during polarizing and negative events, compared to human accounts. However, when considering the effects of emotionally neutral tweets, our data reveals that human accounts attract more retweets and likes compared to bots during all three events. Thus, our findings indicate that emotions are an important aspect when studying the potential influence of bots on Twitter (note that in our data-sets no neutral bot-generated tweet received more attention than emotionally-neutral tweets generated by human accounts).

4.4. Message exchange behavior between humans and bots

In this section, we report on our findings concerning the emotions communicated via direct messaging between human accounts and bots. As noted in previous studies [see, e.g., 61], direct messaging (via @-mentioning on Twitter) typically accounts for a smaller number of messages in a Twitter discourse. Depending on the respective event, our data-sets count between 6%-9% of

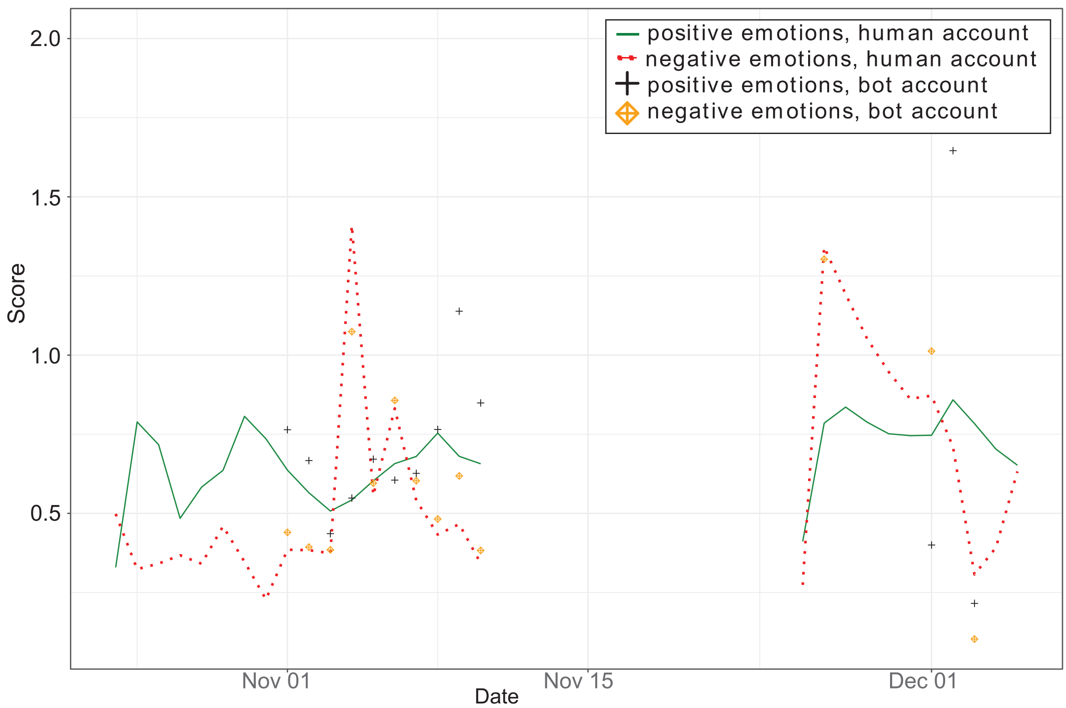


Fig. 7. Temporal emotion score patterns during polarizing events.

Table 5
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
$Like_{bot}$	$\mathbf{1.47 \pm 28.41}$	0.815 ± 13.18	$\mathbf{1.49 \pm 34.11}$
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	$\mathbf{608.41 \pm 2269.43}$	659.69 ± 2213.14
$Like_{bot}$	$\mathbf{2.26 \pm 37.94}$	0.678 ± 14.91	$\mathbf{1.38 \pm 27.84}$
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

Table 6
Direct involvement of bots in message exchanges with human accounts. Number of messages refers to the total number of messages sent in each data-set. Messages per bot/human show the mean number of messages sent per bot or human account followed by the respective standard deviation in brackets.

	Polarizing	Positive	Negative
Number of messages	119177 (6.57%)	103079 (9.24%)	99543 (6.68%)
Messages per bot	$\mu = 2.43(4.34)$	$\mu = 2.16(3.21)$	$\mu = 4.28(10.58)$
Messages per human	$\mu = 1.51(2.29)$	$\mu = 1.37(1.88)$	$\mu = 2.42(9.88)$

messages that include a direct @-mentioning. Though the subset of bots involved in direct messaging is relatively low, they still tend to send on average more messages than human accounts (see Table 6).

However, although having a relatively low involvement with respect to the overall data-set, bots seem to prefer certain emotions when sending direct messages. Fig. 8 shows that overall bots send messages of a higher emotional intensity compared to human accounts during polarizing events. In particular, when averaged and compared to the emotions communicated by human accounts, four emotions are statistically significant in bot messages (as shown in Table 7). These emotions are anger, sadness, fear, and trust, revealing that bots tend to amplify the expression of negative emotions during polarizing events. Similar findings also hold for the positive events in which bots tend to send negative emotion messages of a higher emotional intensity as compared to humans. Interestingly, we found that although bots were responsible for a substantial amount of (broadcast) messages bearing positive emotions during negative events (see Section 4.2), they still tend to conform to the base mood of the event when they send direct messages to certain users.

4.4.1. Emotion-exchange motifs

Finally, we examine the occurrence of *emotion-exchange motifs* in our data-sets. *Emotion-exchange motifs* are statistically significant communication patterns that arise when OSN users (humans or bots) exchange emotional OSN messages. Each of these communication patterns (i.e. each emotion-exchange motif) is represented via a corresponding k-node subgraph. In principle, this analysis can be done for subgraphs of any size k. In this paper, however, we focus on triadic communication patterns, i.e. we investigate 3-node subgraphs. Thus, we identify structural subgraph patterns that emerge when bot accounts exchange emotional messages with human accounts. To this end, we enumerated all possible 3-node subgraphs in each of the emotion-annotated communication networks and detected those that are significant for our real-world networks.

Our analysis shows that bots tend to form specific subgraphs when they communicate emotions to human users. In particular, we identified 243 distinct motifs that differ in their shape and the message-exchange frequency (denoted as edge weights). While some are specific for one network only, the others occur in at least two communication networks (i.e. two out of the three communication networks resulting from the positive, negative, and

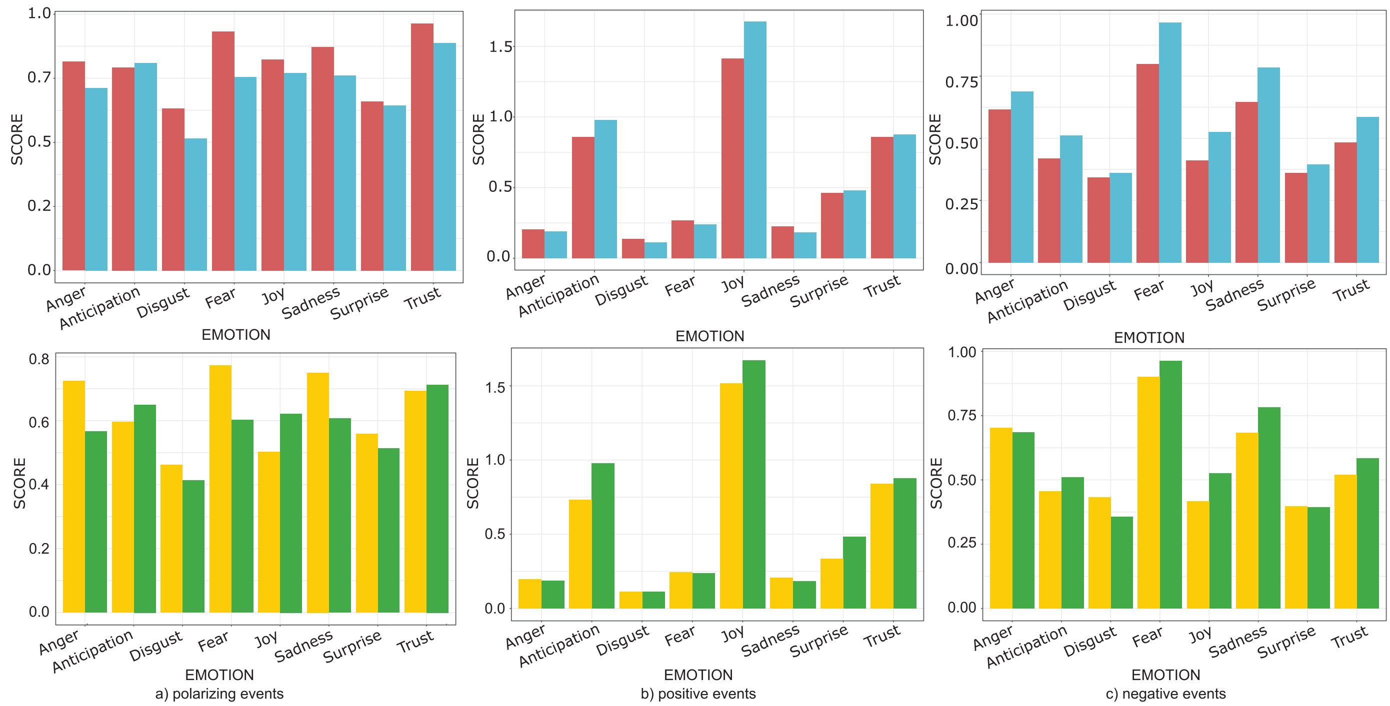


Fig. 8. Emotions sent via direct @-mentioning. Emotion intensities of messages sent by bots are presented in red and emotion intensities sent by human accounts in blue. Emotion intensities of messages received by bots are shown in yellow and emotion intensities received by humans in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
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 ($N_{human} = 76188$; $N_{bot} = 1628$)	Positive ($N_{human} = 74188$; $N_{bot} = 530$)	Negative ($N_{human} = 39315$; $N_{bot} = 1033$)
Anger	**t(4201.9)=3.78, p < 0.05**	$t(1168.3) = 0.68, p > 0.05$	**t(4862.6)=−3.18, p < 0.05**
Disgust	$t(4186.9) = 4.5593, p < 0.05$	$t(1170.2) = 1.29, p > 0.05$	$t(4839) = -1.06, p > 0.05$
Sadness	**t(4185.2)=3.816, p < 0.05**	$t(1170.1) = 1.82, p > 0.05$	**t(4882.6)=−5.74, p < 0.05**
Fear	**t(4190)=6.09, p < 0.05**	$t(1167.9) = 1.09, p > 0.05$	**t(4895.6)=−6.29, p < 0.05**
Trust	**t(4209.8)=2.919, p < 0.05**	$t(1173) = -0.44, p > 0.66$	$t(4883.4) = -5.91, p < 0.05$
Joy	$t(4217.8) = 1.88, p > 0.05$	**t(1176.2)=−4.3, p < 0.05**	**t(4927.1)=−6.35, p < 0.05**
Anticipation	$t(4265.3) = -0.77, p > 0.05$	**t(1176.7)=−2.59, p < 0.05**	**t(4915.1)=−5.71, p < 0.05**
Surprise	$t(4233) = 0.71, p > 0.05$	$t(1179.1) = -0.58, p > 0.05$	**t(4867.9)=−2.29, p < 0.05**

polarizing events in our data-set). Below, we report on a set of re-occurring motifs that emerge in at least two communication networks. In our analysis, we identified in total 29 motifs that occur in two of the communication networks we investigated and 13 that occur in all three communication networks. Subsequently, we first report on the set of re-occurring motifs and then on the network-specific motifs that emerged in one network only.

For the purposes of our analysis, we define the *bot ratio* θ_B as the observed number of bots (B_o) in relation to the maximum amount of possible bots (B_t) in a subgraph.

$$\theta_B = \frac{F(B_o)}{F(B_t)}$$

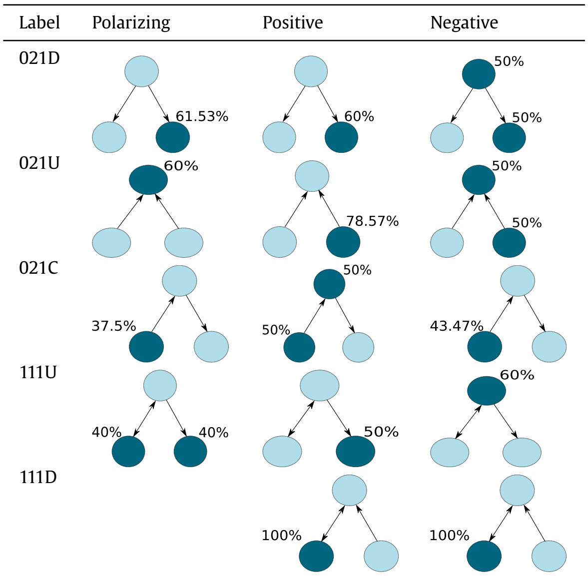
Re-occurring motifs. In general we found that bots take over different functional roles while engaging in an exchange of emotional messages with human users. Table 8 summarizes the generalized motifs[8] categorized with respect to their triadic census type and labelled according to the MAN labeling scheme. Nodes with a bot ratio larger than 33% are depicted in a darker color and labelled with the corresponding bot ratio (in %). While communicating emotions, bots and human accounts tend to form out-stars (021D), in-stars (021U), message chains (021C), and triads with reciprocal edges (111U, 111D). When observing the bot's role in the emotion-exchange motif, we distinguish between bots that send an emotional message (i.e. directly mention a human user via @screenname) and receive an emotional message (i.e. attract an @-mention by a human user).

In contrast to the belief that bots predominantly send messages (for example to spam human users [see, e.g., 67]), our results indicate that bots tend to predominantly attract messages (by being mentioned) during polarizing events (see Table 8). The same, however, does not hold for the positive and negative events where bots tend to equally send and receive messages within a triad. Though the generalization of the triads shown in Table 8 reveals information on the position and the general role of bots in the emotion-exchange motifs, the question still remains to which extent bots take over such roles (when considering the message exchange frequency) and on which emotion layers such motifs occur in our multiplex network.

For further analyses, we thus include the edge weights which reveal the number of messages that have been sent or received by

[8] The motifs in Table 8 focus on the structural characteristics of the respective 3-node subgraphs while generalizing over the edge weights.

Table 8
Bot ratio in common emotion-exchange motifs. Nodes with a bot ratio > 33% are depicted in a darker color.



each node in the respective emotion-exchange motif, as well as the emotion layer of the multiplex network on which the emotion exchange occurred. Table 9 summarizes the in-degree ($deg^{in}(v)$) and out-degree ($deg^{out}(v)$) scores for bot and human accounts averaged over the total in- and out-degree on each layer of our multiplex network. For the purpose of simplicity, Table 9 reports on the emotion-exchange motifs found on:

1. *positive emotion layers*: including the individual positive emotion layers *trust, joy, anticipation*, as well as all combinations of positive emotions,
2. *negative emotion layers*: including the individual negative emotion layers *anger, fear, sadness, disgust*, as well as all combinations of negative emotions,
3. *interlayer*: a combination of positive and negative emotions, and
4. layers including *surprise*: the individual surprise layer and the aggregated emotion layer including *surprise*.

The results in Table 9 show that *polarizing* and *negative* events exhibit a high message-receiving degree (in-degree $deg^{in}(v)$) consistently over all layers in our multiplex network. However, during *positive* events bots tend to predominantly send messages (higher out-degree $deg^{out}(v)$) when communicating expected emotions (i.e. positive emotions). Moreover, when analyzing messages conveying negative emotions during positive events, we found that bots tend to receive such negative messages rather than send them (e.g. as a negative reaction of a human user to a bot-generated tweet).

Network-specific motifs. In addition to re-occurring motifs, our emotion-annotated communication networks also exhibit a number of motifs that appear in a single network only (i.e. motifs that are specific to polarizing, or positive, or negative events only). Such motifs comparatively often 1) contain self-loops (48.15%), or, to a smaller extent, 2) take the form of a closed triad (2.88%).

Self-loop motifs. Motifs containing a self-loop predominantly appear in the Twitter communication about *negative events* (cf. Table1). In particular, we identified 62 motifs containing self-loops that are unique to negative events (53% of all emotion-exchange motifs appearing during negative events). We found that such motifs especially appear when bots and human accounts exchange a *combination* of positive and negative emotions (58% of all emotion-exchange motifs appearing during negative events, excluding the motifs on the surprise layer), rather than for messages where a single emotion (such as anger) is communicated (16.7% of all emotion-exchange motifs appearing during negative events). It is worth mentioning that the nodes producing self-loops by addressing themselves (via @screenname) have predominantly been human accounts rather than bots (see Tables 10 and 11). As pointed by Molyneux and Mouro [52], Twitter users apply self-mentioning (via @screenname) in their tweets to bypass the 140-character restriction and provide more content on of their previous tweets.

As shown in Table 10, it seems that bots tend to attract messages rather then send them during *polarizing* and *negative events*. Interestingly, while engaging in message-exchanges during negative events, bots receive, on average, more positive messages ($\mu = 3.19$, $sd = 2.71$) than negative ($\mu = 1.97$, $sd = 2.32$). However, with respect to messages that have been sent by bots, we found that they send approximately the same amount of positive ($\mu = 0.57$, $sd = 1.43$) and negative ($\mu = 0.55$, $sd = 1.23$) messages during negative events. Our analysis has further shown that bots tend to predominantly attract (i.e. receive) messages when engaging in a direct messaging behavior where mixed emotions are exchanged between bot and human accounts (in this case bots on average receive $mu = 7.94$, $sd = 5.23$ messages). In contrast to the emotion-exchange behavior during positive events (see below), bots tend to predominantly and consistently attract messages on all emotion layers of our multiplex network when communicating during negative events.

However, when analyzing the emotion-exchange motifs that appear during *positive events*, we found different patterns. In particular, during positive events we did not find any self-loop motif on a

Table 9
Average in-degree ($deg^{in}(v)$) and out-degree ($deg^{out}(v)$) for human and bot accounts while communicating positive emotions (joy, trust, anticipation, as well as combinations of positive emotions), negative emotions (fear, disgust, anger, sadness, as well as combinations of negative emotions), a combination of positive and negative emotions (interlayer), and surprise (individual surprise layer, aggregated layer including surprise) in the common set of motifs.

	Polarizing		Positive		Negative	
	$deg^{in}(v)$	$deg^{out}(v)$	$deg^{in}(v)$	$deg^{out}(v)$	$deg^{in}(v)$	$deg^{out}(v)$
Positive emotion layers						
Human	0.34 (0.80)	**1.98 (0.69)**	**2.1 (1.27)**	1.27 (1.01)	1.25 (1.23)	**1.93 (1.69)**
Bot	**3.31 (1.74)**	0.49 (0.93)	0.87 (1.35)	**2.01 (1.60)**	**2.40 (2.2)**	1.40 (1.95)
Negative emotion layers						
Human	0.22 (0.71)	**2.57 (1.09)**	1.38 (0.82)	**1.97 (1.26)**	1.52 (1.36)	1.49 (1.40)
Bot	**4.15 (2.13)**	0.56 (0.86)	**2.00 (2.03)**	0.82 (1.42)	**1.59 (1.80)**	**1.66 (2.18)**
Interlayer (combination of positive and negative emotion layers)						
Human	1.62 (1.58)	**2.26 (1.55)**	**2.27 (1.54)**	1.23 (1.59)	1.36 (1.02)	**1.53 (1.41)**
Bot	**2.12 (2.40)**	1.41 (2.12)	1.26 (2.51)	**2.87 (3.36)**	**1.56 (1.75)**	1.34 (1.85)
Inclusion of surprise						
Human	1.08 (0.57)	**1.61 (1.13)**	**2.78 (1.69)**	1.56 (1.38)	1.91 (1.72)	**2.19 (1.94)**
Bot	**1.59 (0.84)**	0.91 (1.35)	1.48 (2.54)	**3.91 (3.36)**	**2.73 (1.82)**	2.11 (2.62)

Table 10
Average in-degree ($deg^{in}(v)$) and out-degree ($deg^{out}(v)$) for human and bot accounts while communicating positive emotions (joy, trust, anticipation, combination of positive emotions), negative emotions (fear, disgust, anger, sadness, combination of negative emotions), a combination of positive and negative emotions (interlayer), and the inclusion of surprise (surprise, aggregated) in motifs containing self-loops.

	Polarizing		Positive		Negative	
	$deg^{in}(v)$	$deg^{out}(v)$	$deg^{in}(v)$	$deg^{out}(v)$	$deg^{in}(v)$	$deg^{out}(v)$
Positive emotion layers						
Human	0.5 (0.58)	1.5 (1.73)	**4.01 (4.12)**	2.24 (1.85)	2.2 (3.13)	**4.03 (4.42)**
Bot	**5 (0)**	**3 (0)**	0.02 (0.13)	**3.36 (1.45)**	**3.19 (2.71)**	0.57 (1.43)
Negative emotion layers						
Human	**2 (2.05)**	**1.5 (0.51)**	–	–	0.98 (0.96)	**1.79 (1.51)**
Bot	0 (0)	1 (0)	–	–	**1.97 (2.32)**	0.55 (1.23)
Interlayer (combination of positive and negative emotions)						
Human	2 (2.3)	**4 (2.45)**	**4.5 (4.01)**	2.5 (1.77)	3.06 (1.54)	**5.78 (4.58)**
Bot	**4 (1.41)**	0 (0)	0 (0)	**4 (0)**	**7.94 (5.23)**	1.41 (3.23)
Inclusion of surprise						
Human	**2.6 (2.54)**	**2.1 (1.06)**	**4.34 (4.34)**	**3.17 (2.68)**	0.98 (1.21)	**2.45 (2.43)**
Bot	0.13 (0.52)	1.13 (0.52)	0.02 (0.13)	2.31 (1.27)	**3 (2.71)**	0.58 (1.06)

Table 11
Bot ratio in emotion-exchange motifs containing self-loops. Nodes with a bot ratio > 33% are depicted in a darker color. The labelling is adjusted for the presence of self-loops.

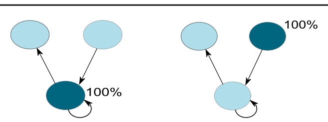
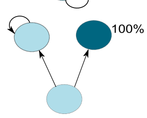
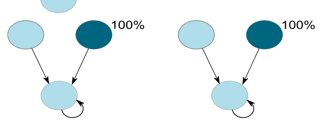
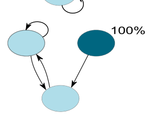
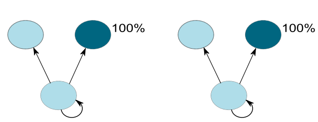
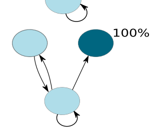
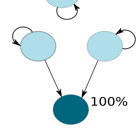
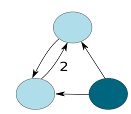
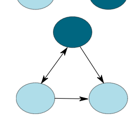
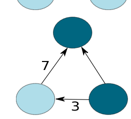
Label	Polarizing / Positive / Negative
$031CL^R$	
$031DL^L$	
$121DL^L$	
$031UL^L$	
$031UL^R$	
$121DL^L$	
$031DL^R$	
$121UL^R$	
$041UL^L$	

Table 12
Position of bot accounts in closed triad motifs. Nodes with a bot ratio > 33% are depicted in a darker color.

Layer	Motif	Data-set	Layer
021D		Positive	Aggregated, positive
120U		Negative	Interlayer
120D		Negative	Interlayer

negative emotion layer, i.e. no self-loop motifs emerge when bots and humans exchange negative messages during positive events. However, we found that bots send comparatively more positive messages during positive events ($\mu = 3.36$, $sd = 1.45$) as compared to human accounts ($mu = 2.24$, $sd = 1.85$). Furthermore, we found that during positive events human accounts receive significantly more messages on all emotion layers (see Tables 10 and 11).

Closed triad motifs. The closed triads[9] formed in human-bot interactions are depicted in Table 12. Our findings show that during *negative events* humans and bots predominantly form closed triads on the *interlayer* where they exchange a combination of positive and negative emotions. During *positive events*, closed triads occur when humans and bots exchange positive emotional messages as well as on the aggregated layer (i.e. when humans and bots exchange a combination of messages including positive and negative emotions, as well as surprise). In this context, it is worth mentioning that during *polarizing events*, where it is expected that OSN users engage in discussions, we found no involvement of bots in a closed triad. Although closed triad motifs exist in our polarizing data-set, they occur only when humans exchange messages with other humans, indicating that in polarizing events closed triad motifs are characteristic for a human-like communication.

5. Discussion

Our findings show that humans, in contrast to bots, conform to the base emotion of an event, confirming corresponding "offline" social studies. For example, Heath [32] demonstrated that people predominantly disseminate positive messages during positive events and negative messages during negative events in a word-of-mouth manner. However, by comparing our findings to the ones in

[9] Note that these motifs are not generalized and show the actual number of message exchanged between two nodes as edge weights.

the related work, we cannot confirm the results reported by Dickerson et al. [19], according to which humans disagree more with the base sentiment of an event compared to bot accounts. One possible explanation for this result is that we analyzed tweets extracted from 24 different events while Dickerson et al. [19] studied a single event only (the 2014 Indian election).

Everett et al. [24] further noted that bots may send messages that diverge from the base emotion of the respective event in order to deceive human users. In our study, there was a noticeable deviation from the base sentiment during *positive* events. More specifically, we found that bots tend to be more negative during positive events ($d_{p-n} = 0.666$ for human accounts, and $d_{p-n} = 0.282$ for bot accounts) and more positive during negative events ($d_{n-p} = 0.192$ for human accounts, and $d_{n-p} = 0.005$ for bot accounts), see also the t-test results in Table 4. As an exception to this finding, we discovered that if bots engage in a direct message exchange with human accounts during *negative events*, they tend to exhibit a more human-like behavior and conform to the base mood of the event (i.e. in negative events, bots predominantly send negative emotional messages, such as news about a death of a celebrity or a terror attack in which they include the screenname of a Twitter user. An example of a bot-generated tweet reads: *"Follow @human-screenname live streaming from #Aleppo #StandWithAleppo*").

Identifying emotionally polarizing bots during *polarizing events* can serve as an indicator for an attempt of opinion swaying. In our study, we found that bot-generated retweets have an impact on the perceived emotionality in the Twitter discourse. In particular, our results indicate that during *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*. The bot-generated messages below have been sent in the run-up to the 2016 US presidential elections:

- "*#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*".

In the run-up to the 2016 Austrian presidential elections, bots disseminated messages such as [see also 44]:

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

For the election events we analyzed, messages that support one particular candidate make up the vast majority (99.37%) of positive bot-generated tweets (87.8% of those messages are retweets). Therefore, we can conclude that bots clearly follow a strategic agenda during elections (and probably during polarizing events in general).

Our findings further indicate that bots tend to spread (retweet) more negative emotions during *positive events* (86.41% retweets) as compared to humans. Such bot-generated tweets found in our data-set either: A) express a negative opinion about a prospectively positive 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: B) surprise the readers by injecting topic-wise unrelated negative content. For example, human-generated Thanksgiving tweets in our data-set are predominantly positive (congratulating and festive messages). In contrast, bots generated topic-wise unrelated tweets that convey negative emotions and were injected into the Thanksgiving discourse by using the *#thanksgiving* hashtag, e.g.:

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

As noted in related studies [see, e.g., 61], we also observed a relatively low involvement of bots in direct messaging behavior (via @screenname) on Twitter. However, though low in involvement, those bots who directly communicate with human users send emotional messages with a higher frequency compared to human accounts. This finding might lead to the assumption that the same pattern will be observable in specific triadic patterns that are formed as bots communicate with human accounts. Interestingly, however, this was not the case. Instead, our analysis showed that the corresponding emotion-exchange motifs more often emerge as bots *receive* messages (via @screenname mentioning) rather then send them. This finding was especially evident for the *polarizing* and *negative events* we analyzed (see Table 1). This also suggests that during polarizing and negative events communication patterns which emerge when bots predominantly send emotional messages (rather than receive them) are sporadic and, in general, not representative for our real-world networks (i.e. those patterns are not statistically significant and over-represented and therefore most often do not qualify as emotion-exchange motifs).

In contrast, the *positive events* in our data-set showed a comparatively higher involvement of bots in message-sending behavior. This might be due to the nature of the corresponding OSN discourse since negative events (such as a terror attack) or polarizing events (such as political elections) can lead to a discussion of a higher emotional arousal than a positive event (such as a happy birthday greeting). Therefore, it is likely that messages related to polarizing or negative events create more heated reactions of a higher emotional intensity from human users (message senders). An example from our data-set reads (the screennames have been anonymized):

- sent by a bot (news): "*Follow @human-A-screenname live streaming from #Aleppo #StandWithAleppo*" (at 2016-11-29, 01:25:02)
- replied to by a human (personal reaction): "*Kids being bombed with chlorine-filled barrels #aleppo #syria @bot-A-screenname*" (at 2016-11-30, 10:15:06)

Thus, we found that bots use emotional content as means to attract more attention for their own tweets in terms of content popularity and visibility (expressed via likes and retweets), and also inspire, attract, and potentially steer emotional reactions from human users in a direct messaging behavior.

In a previous study, we investigated communication patterns (motifs) that emerge during riot events [see, 42]. The results of that investigation indicated that social media discussions of humans and bots do not result in statistically significant patterns if we look at message exchanges only and disregard the emotional tone of the respective message exchange. However, when considering emotions the differences between human and bot accounts become more evident. By systematically analyzing communication patterns mined from 24 real-world events, we found that in a direct interaction with humans (i.e. direct and/or bilateral communication instead of broadcast messages) bots generally send expected emotions and in turn receive a comparatively higher volume of messages conveying shifted emotions. This finding intuitively diverges from the general trend that bots author comparatively more tweets (broadcast messages) which convey shifted emotions than human accounts and reveals that while engaging in a one-to-one message sending behavior with human accounts (i.e. in direct and/or bilateral communication), bots conform to the base mood of an event and behave more human-like.

6. Conclusion

In this paper, we systematically analyzed 4.4 million tweets related to 24 real-world events that have been generated by 1.3 million user accounts, 35.2 thousand of which were identified as bots. Each of the 24 events has been categorized as either positive (e.g., public holiday, birthday of a celebrity), negative (e.g., terror attacks, acts of war), or polarizing (e.g., elections, death of a controversial political figure).

We examined the differences between bots and human accounts with respect to the specific emotions they disseminate. Our findings indicate that humans generally conform to the base emotion of the respective event. Bots, on the other hand, contributed to the higher intensity of shifted emotions (e.g., positive emotions during negative events) to receive more attention (in terms of likes or retweets) to their content. In particular, our results indicated that emotional bot-generated tweets attract more likes and retweets if the tweet conveys shifted emotions. Interestingly, emotionally neutral tweets authored by bots did not receive much attention in terms of likes and retweets. We further discovered that bots are not only most active during polarizing events ($N(tweets)_{pol} = 3\%$, $N(tweets)_{pos} = 0.6\%$, $N(tweets)_{neg} = 2\%$) but they also tend to emotionally polarize during controversial events (such as elections). Such a polarization was evident especially in the election-related data-sets. For example, we found that bots inject shifted emotions into topic-wise unrelated Twitter discussions (e.g., negative messages related to the 2016 US presidential election that include a #*thanksgiving* hashtag).

Thus, the emotions conveyed in tweets may serve as a valuable indicator for distinguishing bot or a human activity. However, we identified one exception to this finding – while communicating via direct @-mentioning (instead of sending broadcast messages), bots tend to act more human-like and conform to the base mood of the respective event. In this context, we identified a number of characteristic communication patterns that arise when humans and bot accounts directly exchange messages (via @screenname). Our initial assumption was that bots would predominantly take the role of message senders (e.g. acting as spammers). However, our analysis revealed that in the *emotion-exchange motifs* we identified, bots more often act as message receivers rather then message senders. This pattern was especially evident during polarizing and negative events, where bots receive emotional reactions from human users.

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. Extending our work on the emotion-exchange motifs, we further plan to examine the role of specific user accounts and the temporal aspects of the formation of emotion-exchange motifs.

Conflict of interest declaration

We declare that we do not have any conflict of interest concerning this paper and that there has not been any financial support by third-parties that could have influenced its outcome.

Acknowledgments

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

Appendix A

Table A.13
Search terms used in the data extraction procedure.

Event	Search terms
1) Death of Fidel Castro	#Castro, #FidelCastro, #CastroDeath, Fidel Castro
2) 2016 Austrian presidential elections	#vdb2016, #vdb, #VanderBellen, #bpw2016, #mehrdennje, #NorbertHofer, #NorbertHofer2016, #Hofer, #Austrianelection, #Austria #election
3) 2016 US presidential elections	#VoteTrumpPence, #maga, #MakeAmericaGreatAgain, #VoteHillary, #HillaryClintonForPresident, #ImWithHer, #StrongerTogether, #USElections2016, #USElection, #ElectionDay, Donald Trump, Hillary Clinton
4) The Walking Dead season 7 premiere	#twd, #thewalkingdead, #Lucille #Negan, #WalkingDead
5) Rosberg winning Formula 1	Nico Rosberg win, #F1Finale, #AbuDhabiGP, Rosberg champion, Rosberg triumph, #NicoRosberg
6) Murray winning ATP	Murray #ATPFinals, Andy Murray, andy_murray
7) Rosberg retirement message	#NicoRosberg retire, #NicoRosberg farewell, #NicoRosberg farewell, #Rosberg retire
8)"Beauty and the Beast" trailer release	#BeautyAndTheBeast
9) "Fantastic beasts" trailer release	#FantasticBeasts, #FBAWTFTmovie, #FantasticBeastsandWhereToFindThem
10) ComiCon Vienna	#viecc
11) Miley Cyrus birthday	#HappyBirthdayMiley, #MileyCyrus birthday, @MileyCyrus birthday
12) New Pentatonix album released	#PTXMerryGentlemen
13) Ellen Degeneres medal of freedom	#MedalofFreedom Ellen, Ellen Medal of freedom
14) Thanksgiving	#Thanksgiving
15) Erdogan's threats to EU	#Erdogan warn EU, #Erdogan threat EU, #Erdogan blackmail EU
16) US anti-Trump protests	#NotMyPresident, #AntiTrump, #TrumpRiot, #TrumpProtest, #NeverTrump
17) Death of Leonard Cohen	#LeonardCohen, Leonard Cohen
18) Death of Colonel Abrams	#ColonelAbrams, Colonel Abrams
19) Aleppo bombings	#StandWithAleppo, #Aleppo, Aleppo bombing
20) Seattle shooting	#SeattleShooting, #seattle #shooting
21) Lufthansa strike	#Lufthansa strike, #Lufthansa cancel
22) Ransomware in Seattle	#SanFrancisco railway, #SanFrancisco hack, #SanFrancisco ransom
23) Yellowstone incident	#yellowstone #hotpot
24) Earthquake in central Italy	#PrayForItaly, #Earthquake #Italy, #ItalyEarthquake, #Terremotocentroitalia, #terremoto #Italy, Italy earthquake

References

[1] M. Abdul-Mageed, L. Ungar, Emonet: Fine-grained emotion detection with gated recurrent neural networks, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, 2017, pp. 718–728, doi:10.18653/v1/P17-1067.

[2] N. Abokhodair, D. Yoo, D.W. McDonald, Dissecting a social botnet: growth, content and influence in twitter, in: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW, ACM, New York, NY, USA, 2015, pp. 839–851, doi:10.1145/2675133.2675208.

[3] U. Alon, Network motifs: theory and experimental approaches, Nat. Rev. Genetics (8) (2007) 450–461, doi:10.1038/nrg2102.

[4] V. Barash, N. Milic-Frayling, M.A. Smith, A method for discovering dynamic network motifs by encoding topic propagation, in: Proceedings of the IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 01, IEEE Computer Society, Washington, DC, USA, 2013, pp. 451–458.

[5] V. Batagelj, A. Mrvar, A subquadratic triad census algorithm for large sparse networks with small maximum degree, Soc. Netw. 23 (3) (2001) 237–243, doi:10.1016/S0378-8733(01)00035-1.

[6] M.E. Beber, C. Fretter, S. Jain, N. Sonnenschein, M. Müller-Hannemann, M.-T. Hütt, Artefacts in statistical analyses of network motifs: general framework and application to metabolic networks, J. R. Soc. Interf. 9 (77) (2012) 3426–3435, doi:10.1098/rsif.2012.0490.

[7] E.A. Bender, E. Canfield, The asymptotic number of labeled graphs with given degree sequences, J. Combinatorial Theory Series A 24 (3) (1978) 296–307.

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

[9] N. Chavoshi, H. Hamooni, A. Mueen, Debot: Twitter bot detection via warped correlation, in: Proceedings of the IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 817–822, doi:10.1109/ICDM.2016.0096.

[10] N. Chavoshi, H. Hamooni, A. Mueen, Identifying correlated bots in Twitter, in: E. Spiro, Y.-Y. Ahn (Eds.), Proceedings of the 8th International Conference on Social Informatics, SocInfo Part II, Springer International Publishing, Cham, Bellevue, WA, USA, 2016, pp. 14–21, doi:10.1007/978-3-319-47874-6_2.

[11] N. Chavoshi, H. Hamooni, A. Mueen, On-demand bot detection and archival system, in: Proceedings of the 26th International Conference on World Wide Web Companion, WWW Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2017, pp. 183–187, doi:10.1145/3041021.3054733.

[12] N. Chavoshi, H. Hamooni, A. Mueen, Temporal patterns in bot activities, in: Proceedings of the 26th International Conference on World Wide Web Companion, WWW Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2017, pp. 1601–1606, doi:10.1145/3041021.3051114.

[13] Z. Chu, S. Gianvecchio, H. Wang, S. Jajodia, Who is tweeting on Twitter: Human, bot, or cyborg? in: Proceedings of the 26th Annual Computer Security Applications Conference, ACSAC, ACM, New York, NY, USA, 2010, pp. 21–30, doi:10.1145/1920261.1920265.

[14] M. Coletto, K. Garimella, A. Gionis, C. Lucchese, Automatic controversy detection in social media: a content-independent motif-based approach, Online Soc. Netw. Media 3-4 (2017) 22–31, doi:10.1016/j.osnem.2017.10.001.

[15] E. Colleoni, A. Rozza, A. Arvidsson, Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data, J. Commun. 64 (2) (2014) 317–332, doi:10.1111/jcom.12084.

[16] L.P. Cordella, P. Foggia, C. Sansone, M. Vento, A (sub)graph isomorphism algorithm for matching large graphs, IEEE Trans. Pattern Anal. Mach. Intell. 26 (10) (2004) 1367–1372, doi:10.1109/TPAMI.2004.75.

[17] C.A. Davis, O. Varol, E. Ferrara, A. Flammini, F. Menczer, Botornot: A system to evaluate social bots, in: Proceedings of the 25th International Conference Companion on World Wide Web, WWW Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2016, pp. 273–274, doi:10.1145/2872518.2889302.

[18] J.A. Davis, S. Leinhardt, The Structure of Positive Interpersonal Relations in Small Groups, Houghton Mifflin, Boston, 1972.

[19] J.P. Dickerson, V. Kagan, V.S. Subrahmanian, Using sentiment to detect bots on Twitter: are humans more opinionated than bots? in: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2014, pp. 620–627, doi:10.1109/ASONAM.2014.6921650.

[20] M.E. Dickison, M. Magnani, L. Rossi, Multilayer Social Networks, Cambridge University Press, 2016, doi:10.1017/CBO9781139941907.

[21] A. Duma, A. Topirceanu, A network motif based approach for classifying online social networks, in: Proceedings of the IEEE 9th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI), 2014, pp. 311–315, doi:10.1109/SACI.2014.6840083.

[22] C. Edwards, A. Edwards, P.R. Spence, A.K. Shelton, Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter, Comput. Human Behav. 33 (2014) 372–376, doi:10.1016/j.chb.2013.08.013.

[23] K. Erciyes, Complex Networks: An Algorithmic Perspective, CRC Press, Inc., 2014.

[24] R.M. Everett, J.R.C. Nurse, A. Erola, The anatomy of online deception: What makes automated text convincing? in: Proceedings of the 31st Annual ACM Symposium on Applied Computing, SAC, ACM, New York, NY, USA, 2016, pp. 1115–1120, doi:10.1145/2851613.2851813.

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

[26] E. Ferrara, O. Varol, C. Davis, F. Menczer, A. Flammini, The rise of social bots, Commun. ACM 59 (7) (2016) 96–104, doi:10.1145/2818717.

[27] B.L. Fredrickson, The role of positive emotions in positive psychology: the broaden-and-build theory of positive emotions., Am. Psychol. 56 (2001) 218–226.

[28] N.A. Ghani, S. Hamid, I.A.T. Hashem, E. Ahmed, Social media big data analytics: a survey, Comput. Human Behav. (2018), doi:10.1016/j.chb.2018.08.039.

[29] Z. Gilani, R. Farahbakhsh, J. Crowcroft, Do bots impact Twitter activity? in: Proceedings of the 26th International Conference on World Wide Web Companion, WWW Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2017, pp. 781–782, doi:10.1145/3041021.3054255.

[30] S.J. Guo, The 2013 Boston marathon bombing: publics emotions, coping, and organizational engagement, Public Relat. Rev. 43 (4) (2017) 755–767, doi:10.1016/j.pubrev.2017.07.003.

[31] S. Gurukar, S. Ranu, B. Ravindran, Commit: a scalable approach to mining communication motifs from dynamic networks, in: Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, SIGMOD '15, ACM, New York, NY, USA, 2015, pp. 475–489, doi:10.1145/2723372.2737791.

[32] C. Heath, Do people prefer to pass along good or bad news? Valence and relevance of news as predictors of transmission propensity, Organ. Behav. Hum. Decis. Proces. 68 (2) (1996) 79–94.

[33] P. Howard, A. Duffy, D. Freelon, M. Hussain, W. Mari, M. Mazaid, Opening closed regimes, what was the role of social media during the Arab Spring? Project Inf. Technol. Polit. Islam (2011) 1–30.

[34] C. Hutto, S. Yardi, E. Gilbert, A longitudinal study of follow predictors on twitter, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13, ACM, New York, NY, USA, 2013, pp. 821–830, doi:10.1145/2470654.2470771.

[35] H.K. Kim, J. Niederdeppe, The role of emotional response during an H1N1 influenza pandemic on a college campus, J. Public Relat. Res. 25 (1) (2013) 30–50, doi:10.1080/1062726X.2013.739100.

[36] H.S. Kim, S. Lee, J.N. Cappella, L. Vera, S. Emery, Content characteristics driving the diffusion of antismoking messages: Implications for cancer prevention in the emerging public communication environment, J. Nat. Cancer Inst. Monogr. 47 (2013) 182–187.

[37] N. Kollanyi, P. Howard, S. Woolley, Bots and automation over Twitter during the U.S. election, Data Memo. Oxford UK Project Comput. Propag. 2016 (4) (2016) 1–5.

[38] A.D.I. Kramer, J.E. Guillory, J.T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, Proc. Nat. Acad. Sci. 111 (24) (2014) 8788–8790, doi:10.1073/pnas.1320040111.

[39] E. Kušen, M. Strembeck, An analysis of emotion-exchange motifs in multiplex networks during emergency events, Appl. Netw. Sci. 4 (1) (2019) 8:1–8:33, doi:10.1007/s41109-019-0115-6.

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

[41] E. Kušen, M. Strembeck, Investigation of emotion exchange motifs in bot/human interactions during riot events, in: Proceedings of the fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 2018, pp. 19–26, doi:10.1109/SNAMS.2018.8554860.

[42] E. Kušen, M. Strembeck, On message exchange motifs emerging during human/bot interactions in multilayer networks: The case of two riot events, in: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2018, pp. 1274–1279, doi:10.1109/ASONAM.2018.8508275.

[43] E. Kušen, M. Strembeck, On the public perception of police forces in riot events–the role of emotions in three major social networks during the 2017 G20 Riots, in: Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS), 2018.

[44] E. Kušen, M. Strembeck, Politics, sentiments, and misinformation: An analysis of the twitter discussion on the 2016 austrian presidential elections, Online Soc. Netw. Media 5 (2018) 37–50, doi:10.1016/j.osnem.2017.12.002.

[45] E. Kušen, M. Strembeck, 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, pp. 13–22. Madeira, Portugal.

[46] E. Kušen, M. Strembeck, G. Cascavilla, M. Conti, 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), ACM, 2017, pp. 321–324.

[47] E. Kušen, M. Strembeck, M. Conti, Emotional Valence Shifts and User Behavior on Twitter, Facebook, and YouTube, Springer International Publishing, pp. 63–83. doi:10.1007/978-3-030-02592-2_4.

[48] J. Lee, W. Xu, The more attacks, the more retweets: Trumps and clintons agenda setting on twitter, Public Relat. Rev. 44 (2) (2018) 201–213, doi:10.1016/j.pubrev.2017.10.002.

[49] C. Llewellyn, L. Cram, A. Favero, R.L. Hill, Russian troll hunting in a brexit twitter archive, in: Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries, JCDL, ACM, New York, NY, USA, 2018, pp. 361–362, doi:10.1145/3197026.3203876.

[50] C. Mascaro, D. Agosto, S.P. Goggins, One-sided conversations: The 2012 presidential election on Twitter, in: Proceedings of the 17th International Digital Government Research Conference on Digital Government Research, ACM, New York, NY, USA, 2016, pp. 112–121, doi:10.1145/2912160.2912185.
[51] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: Simple building blocks of complex networks, Science 298 (5594) (2002) 824–827.
[52] L. Molyneux, R.R. Mouro, Political journalists' normalization of Twitter, Journal. Stud. 0 (0) (2017) 1–19, doi:10.1080/1461670X.2017.1370978.
[53] M.E.J. Newman, S.H. Strogatz, D.J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E 64 (2) (2001) 026118.
[54] P. Nulty, Y. Theocharis, S.A. Popa, O. Parnet, K. Benoit, Social media and political communication in the 2014 elections to the european parliament, Electoral Stud. 44 (2016) 429–444, doi:10.1016/j.electstud.2016.04.014.
[55] O. Oh, K. Kwon, H.R. Rao, An exploration of social media in extreme events: Rumor theory and Twitter during the Haiti earthquake 2010, in: Proceedings of the 31st International Conference on Information Systems, 2010.
[56] J. Pan, Y. Liu, X. Liu, H. Hu, Discriminating bot accounts based solely on temporal features of microblog behavior, Phys. A Stat. Mech. Appl. 450 (2016) 193–204, doi:10.1016/j.physa.2015.12.148.
[57] A. Paranjape, A.R. Benson, J. Leskovec, Motifs in temporal networks, in: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, WSDM '17, ACM, New York, NY, USA, 2017, pp. 601–610, doi:10.1145/3018661.3018731.
[58] R. Plutchik, The nature of emotions, Am. Sci. 89 (4) (2001).
[59] R. Procter, F. Vis, A. Voss, Reading the riots on Twitter: methodological innovation for the analysis of big data, Int. J. Soc. Res. Methodol. 16 (3) (2013) 197–214, doi:10.1080/13645579.2013.774172.
[60] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, F. Menczer, Detecting and tracking political abuse in social media, in: Proceedings 5th International AAAI Conference on Weblogs and Social Media (ICWSM), 2011, pp. 297–304.
[61] F. Riquelme, P. Gonzalez-Cantergiani, Measuring user influence on Twitter: a survey, Inf. Process. Manag. 52 (5) (2016) 949–975.
[62] R. Rotabi, K. Kamath, J. Kleinberg, A. Sharma, Detecting strong ties using network motifs, in: Proceedings of the 26th International Conference on World Wide Web Companion, WWW '17 Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2017, pp. 983–992, doi:10.1145/3041021.3055139.
[63] S. Savage, A. Monroy-Hernandez, T. Höllerer, Botivist: calling volunteers to action using online bots, in: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, CSCW, ACM, New York, NY, USA, 2016, pp. 813–822, doi:10.1145/2818048.2819985.
[64] U. Schimmack, Affect measurement in experience sampling research, J. Happiness Stud. 4 (1) (2003) 79–106, doi:10.1023/A:1023661322862.
[65] W.E. Schlauch, K.A. Zweig, G. Theory, N. Analysis, Influence of the null-model on motif detection, in: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2015, pp. 514–519.
[66] C. St Louis, G. Zorlu, Can Twitter predict disease outbreaks? BMJ 344 (2012), doi:10.1136/bmj.e2353.
[67] H. Stefanie, T.D. Bowman, K. Holmberg, A. Tsou, C.R. Sugimoto, L. Vincent, Tweets as impact indicators: examining the implications of automated bot accounts on Twitter, J. Assoc. Inf. Sci. Technol. 67 (1) (2016) 232–238.
[68] J. Surma, Social exchange in online social networks. the reciprocity phenomenon on facebook, Comput. Commun. 73 (2016) 342–346, doi:10.1016/j.comcom.2015.06.017. Online Social Networks.
[69] J. Suttles, N. Ide, Distant supervision for emotion classification with discrete binary values, in: A. Gelbukh (Ed.), Computational Linguistics and Intelligent Text Processing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 121–136.
[70] A. Topirceanu, A. Duma, M. Udrescu, Uncovering the fingerprint of online social networks using a network motif based approach, Comput. Commun. 73 (2016) 167–175.
[71] S. Tsugawa, H. Ohsaki, Negative messages spread rapidly and widely on social media, in: Proceedings of the ACM on Conference on Online Social Networks, 2015, pp. 151–160, doi:10.1145/2817946.2817962.
[72] S. Valenzuela, N. Park, K.F. Kee, Is there social capital in a social network site? Facebook use and college students' life satisfaction, trust, and participation1, J. Comput. Med. Commun. 14 (4) (2009) 875–901, doi:10.1111/j.1083-6101.2009.01474.x.
[73] G. Warner-Soderholm, A. Bertsch, E. Sawe, D. Lee, T. Wolfe, J. Meyer, J. Engel, U.N. Fatilua, Who trusts social media? Comput. Hum. Behav. 81 (2018) 303–315, doi:10.1016/j.chb.2017.12.026.
[74] S. Wernicke, Efficient detection of network motifs, IEEE ACM Trans. Comput. Biol. Bioinf. 3 (4) (2006) 347–359, doi:10.1109/TCBB.2006.51.
[75] J. Xie, C. Zhang, M. Wu, Y. Huang, Influence inflation in online social networks, in: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, IEEE Press, 2014, pp. 435–442.
[76] X. Yan, P. Jiang, Effect of the dynamics of human behavior on the competitive spreading of information, Comput. Human Behav. 89 (2018) 1–7, doi:10.1016/j.chb.2018.07.014.
[77] U. Yaqub, S.A. Chun, V. Atluri, J. Vaidya, Analysis of political discourse on twitter in the context of the 2016 US presidential elections, Gov. Inf. Q. 34 (4) (2017) 613–626.
[78] Q. Zhao, Y. Tian, Q. He, N. Oliver, R. Jin, W.-C. Lee, Communication motifs: a tool to characterize social communications, in: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, ACM, New York, NY, USA, 2010, pp. 1645–1648.