Negative messages spread rapidly and widely on social media

Negative Messages Spread Rapidly and Widely on Social Media

Sho Tsugawa
Faculty of Engineering, Information and Systems
University of Tsukuba
Ibaraki 305–8573, Japan
s-tugawa@cs.tsukuba.ac.jp

Hiroyuki Ohsaki
School of Science and Technology
Kwansei Gakuin University
Hyogo 669-1337, Japan
ohsaki@kwansei.ac.jp

ABSTRACT

We investigate the relation between the sentiment of a message on social media and its virality, defined as the volume and the speed of message diffusion. We analyze 4.1 million messages (tweets) obtained from Twitter. Although factors affecting message diffusion on social media have been studied previously, we focus on message sentiment, and reveal how the polarity of message sentiment affects its virality. The virality of a message is measured by the number of message repostings (retweets) and the time elapsed from the original posting of a message to its Nth reposting (N-retweet time). Through extensive analyses, we find that negative messages are likely to be reposted more rapidly and frequently than positive and neutral messages. Specifically, the reposting volume of negative messages is 1.2–1.6-fold that of positive and neutral messages, and negative messages spread at 1.25 times the speed of positive and neutral messages when the diffusion volume is large.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Science

General Terms

Human Factors

Keywords

Social media, Twitter, Information diffusion, Retweet, Sentiment

1. INTRODUCTION

On social media, such as Twitter and Facebook, users post many messages including their opinions and feelings. One of the most successful social media, Twitter, allows users to post *tweets*, which are short messages with a limit of 140 characters. As of early 2014, 240 million users were posting over 500 million tweets on Twitter each day [33].

Some of the messages posted on social media are disseminated to many other users by word-of-mouth, which affects trends and public opinions in society. Social media users can disseminate

messages to their friends via functionalities, such as *retweeting* in Twitter and *share* in Facebook. This word-of-mouth message diffusion on social media is an important mechanism that influences public opinion and can affect brand awareness and product market share [3]. Therefore, information diffusion in social media has attracted the attention of many researchers [4, 14, 17, 18, 22, 26–28].

As we will discuss in Section 2, factors affecting word-of-mouth message diffusion in social media have been analyzed extensively [17, 22, 28]. Researchers often focus on Twitter as one of the largest social media, and investigate the relation between features extracted from a tweet and its virality. For instance, it has been shown that tweets with features such as URLs, hashtags, and emotional words are more likely to be retweeted than those without these features [22]. It has also been shown that the tweet topic and the number of followers of the tweet publisher are major factors affecting tweet diffusion [17, 28].

We focus on *sentiment* as a factor affecting message diffusion, and examine the effects of positive and negative sentiment in each tweet on its virality on Twitter. Behaviors of social media users are not necessarily objective and legitimate, and psychological and emotional factors are expected to affect the users' behaviors.

The relation between message sentiment and the virality of the message, defined as the volume and the speed of the message diffusion, has been studied [14, 26, 27]. However, different results have been reported for the volume of message diffusion. For instance, Gruzd *et al.* showed that positive tweets are retweeted more than negative tweets [14], whereas Stieglitz *et al.* showed the opposite [27]. Moreover, most studies focus on only the volume of diffusion and do not focus on the diffusion speed. Although Stieglitz *et al.* [27] performed pioneering work analyzing the relation between tweet sentiment and diffusion speed, their analyses used the time interval between the original tweet and only the first retweet as a measure of diffusion speed.

This paper aims to reveal how the sentiment of a tweet affects its virality in terms of both diffusion volume and speed on Twitter by using a large-scale dataset containing 4.1 million tweets. Our main contributions are as follows.

- We investigate 4.1 million non-domain-specific tweets to understand general effects of the sentiment of a tweet on its virality in social media. Previous studies used domain-specific tweets, such as tweets related to the Olympics [14] and political elections [26, 27], and show different results. We used a dataset of mixed domain tweets, and examined the general relation between sentiment and virality in general situations.

- We reveal that negative messages are typically more viral in terms of diffusion volume than positive and neutral messages. Psychology studies suggest that negative things have

a strong effect on people than positive things [6, 24, 31]. We provide empirical evidences of the existence of such bias on social media.

- We also reveal that negative messages spread faster than positive and neutral messages when the diffusion volume is large. We used the time interval between the original tweet and the Nth retweet (N-retweet time) to measure its diffusion speed. By collecting a large number of tweets, we obtained a dataset including tweets with a large retweet count. To the best of our knowledge, this is the first study to investigate the relation between the sentiment and diffusion speed of tweets with large diffusion volume.

The remainder of the paper is organized as follows. Section 2 introduces works related to analyses of message diffusion on social media. In Section 3, we introduce the theoretical background and research questions. Section 4 explains the methodology and dataset used for the analyses. Section 5 shows the results, and Section 6 discusses the implications of the results and the limitations of the work. Finally, Section 7 contains our conclusions.

2. RELATED WORK

Factors affecting retweetability of tweets (i.e., probability of retweet) have been analyzed in previous work [15, 22, 28]. Suh *et al.* analyzed 74 million tweets, and showed that the presence of hashtags and URLs significantly affects retweetability, whereas the number of past tweets does not [28]. Naveed *et al.* analyzed 60 million tweets, and showed that the presence of emotional words, hashtags, and URLs are major factors affecting retweetability [22].

Hansen *et al.* investigated the relation between emotions contained in a tweet and its retweetability [15]. Analysis of approximately 560,000 tweets showed that for tweets about news, negative tweets have higher retweetability than positive tweets, whereas the opposite is true for non-news tweets. These studies have focused on retweetability; however, in this work, we focused on the volume and speed of retweets.

Factors affecting the volume of retweets have been analyzed [14, 17, 27]. Hong *et al.* showed that tweet topics determined by topic modeling, which is a widely used natural language processing technique [9], and the number of followers of the tweet publisher are useful features for predicting the volume of retweets [17].

The relation between tweet sentiment and the volume of tweet diffusion has been examined [14, 26, 27]. Gruzd *et al.* analyzed 46,000 tweets related to the Winter Olympics in 2010, and found that positive tweets have a larger number of retweets than negative tweets [14]. In contrast, Stieglitz *et al.* analyzed approximately 170,000 tweets related to political elections in Germany [26, 27], and revealed that negative and positive tweets have a larger volume of retweets than neutral tweets [26, 27]. Moreover, in one dataset they showed that negative tweets had a larger volume of retweets than positive tweets, whereas in the other there was no significant difference in retweet volume between positive and negative tweets [27]. These studies used domain-specific tweets, where the tweets were related to specific social events, and reached different conclusions. Our study uses larger-scale non-domain-specific tweets, and investigates the relation between the sentiment of a tweet and its diffusion volume, eliminating the effects of the tweet domain.

Analyses of the relation between message sentiment and diffusion speed is limited. Stieglitz *et al.* investigated the relation between tweet sentiment and retweet speed [27]. They used the time interval between the original tweet and the first retweet (1-retweet

time) as a measure of retweet speed, and showed that there was no significant difference between retweet speed of positive and negative tweets. Extending the methodology of their work, we used the time interval between the original tweet and the Nth retweet as a measure of diffusion speed, and investigate the effects of message sentiments on its diffusion speed.

Prediction of the volume of retweets is a related and active research topic [11, 20]. Cheng *et al.* predicted the volume of retweets with machine learning techniques [11]. Although these studies have constructed prediction models using several features, we examine the effects of the features (message sentiment in this study) on the retweet volume. Our results can be used to predict retweet volume and provide several suggestions for improving marketing and designing new functionality in social media, which is discussed in Section 6.

3. THEORY AND RESEARCH QUESTIONS

Psychology studies suggest that negative things have a stronger effect on people than positive things, which is called *negativity bias*, and this bias exists in many situations [6, 24, 31]. Moreover, positive and negative emotions affect virality [7, 8]. Psychological arousal increases virality, and news articles evoking positive and negative emotions often go viral [8]. Therefore, it is expected that negative tweets are retweeted more than positive and neutral tweets, and that positive tweets are retweeted more than neutral tweets.

However, empirical observations of the relation between tweet sentiment and retweet volume are limited; therefore, it is still unclear whether negativity bias exists in social media. As discussed in Section 2, tweets with different domains show different relations [14, 27]. Therefore, we tackle the following question using large-scale non-domain-specific tweets.

RQ 1 *How is tweet sentiment related to the retweet volume?*

As negativity bias theory suggests, negative emotion in a tweet may increase the reaction speed to the tweet. However, as discussed in Section 2, analyses of the relation between tweet sentiment and diffusion speed are also limited. Our second research question is as follows.

RQ 2 *How is tweet sentiment related to the retweet speed?*

In what follows, we tackle these two research questions by analyzing large-scale tweet data.

4. METHODOLOGY

In this section, we explain the dataset and methodology that we used to answer our research questions.

4.1 Overview

We collected tweets on Twitter, and investigated the relation between the sentiment of each tweet and its virality. To focus on users with the same culture and to eliminate the effects of different time-zones, we used tweets from Japanese twitter users. Following the method in Ref. [27], we categorized the tweet sentiment as *positive*, *negative*, and *neutral*.

The tweet sentiment was determined by using two methods: objective classification using a dictionary of positive and negative words [29, 30]; and subjective classification by several people. For objective classification, we determined the sentiment of each tweet by counting the number of affective words used in the tweet. Since such objective classification could cause classification errors, we also used subjective classification of a subset of collected tweets.

Table 1: Distribution of the number of retweets in the dataset

Section	Number of retweets	Number of tweets
1	2–10	3,748,449
2	11–25	318,640
3	26–50	111,527
4	51–75	37,174
5	76–100	18,616
6	101–250	33,847
7	251–500	10,359
8	501–750	2,903
9	751–1000	1,227
10	1001 or more	2,295

Table 2: Statistics for the collected tweet dataset, D_A

	Mean	Median	Std. dev.
Number of retweets	9.70	3	70.80
Number of URLs	0.39	0	0.53
Number of hashtags	0.27	0	0.70
Number of followers	6237.30	515	36220.61

The two classification methods were used to check the robustness of the results. Details of these methods are explained in Section 4.3.

For each original tweet, we calculated the number of retweets and the time interval between the original tweet and the Nth retweet (N-retweet time). We investigated the relation between these measures and tweet sentiment.

4.2 Dataset

Using the Twitter application programming interface (API), we collected Japanese retweets posted during July 25-31 2013[1]. Retweets where the original tweet was posted before 25 July 2013 were discarded. For each original tweet, we counted the number of retweets and extracted original tweets that were retweeted multiple times, namely tweets with a retweet number of more than one. This was intended to focus on tweets with a certain amount of retweet volume. We obtained 4,285,037 original tweets, referred to hereafter as tweets. There were no special social events such as the Olympic and political elections during the period of data collection. The distribution of the number of retweets in the dataset is shown in Table 1. Table 1 shows that our dataset included tweets with a large diffusion volume. Because the distribution of the number of retweets is heavy-tailed [21] and a large retweet diffusion is a rare event, previous studies [14,27] use tweets with relatively small diffusion. In contrast, by collecting a large number of tweets, our dataset includes a sufficient number of tweets with a large diffusion volume, which allows us to analyze N-retweet time for a large retweet count, N.

From the 4,285,037 tweets, we chose 8,000 tweets for determining sentiment by manual evaluations. For obtaining 8,000 tweets, we used stratified sampling rather than random sampling to extract tweets with different diffusion volumes. We classified all tweets into 10 sections shown in Table 1, and we randomly chose 800 tweets for each section. We denote the dataset of all tweets as D_A, and the 8,000 sampled tweets as D_S. Statistics about collected tweet data, D_A, are shown in Table 2.

---

[1] We used the Search API in Twitter REST API v1.1, and collected Japanese tweets using the query q=RT, lang=ja.

Table 3: Examples of positive and negative words. The English translation of the Japanese words listed in the dictionary are shown.

Positive	Negative
Happy, laugh, pretty, favorite	Sad, dislike, sick, fear
good, comfortable, smile	bad, horrible, tired
celebrate, beautiful, love	unlucky, anxiety, sorry

4.3 Methods for Inferring Tweet Sentiment

We inferred the sentiment of each tweet in dataset D_A by using a dictionary of affective words. The dictionary is compiled by manual evaluation of a dictionary of positive and negative words extracted according to a technique in Refs. [29, 30]. The dictionary contains 2,871 positive words and 3,534 negative words. Examples of words are show in Tab. 3. We used MeCab [1] for morphological stemming of the Japanese tweet text, and obtained words used in each tweet. For each tweet, we counted the number of positive and negative words listed in the dictionary. We classified each tweet by the following rules: a tweet that had at least one positive word and no negative words was positive; a tweet that had at least one negative word and no positive words was negative; a tweet that had no positive and negative words was neutral; and other tweets were discarded. Following these rules, we obtained 863,830 positive tweets, 343,910 negative tweets, 2,929,324 neutral tweets, and 147,973 tweets were discarded. Previous research has [15, 22] used similar dictionary-based approaches to analyze the relation between tweet sentiment and virality. Therefore, this approach is reasonable for classifying large-scale tweet data.

Moreover, we inferred the sentiment of each tweet in dataset D_S by manual evaluation. We recruited 11 annotators from the undergraduate and graduate students in our laboratory. Annotators were instructed to read the tweets independently, and tag each tweet as *positive*, *negative*, *neutral*, or *uncertain*. For each tweet, three annotators independently gave a sentiment tag for the tweet. Following the method used in the sentiment analysis task in the SemEval workshop [23], we adopted *majority vote* for determining the sentiment label of each tweet. We discarded tweets that were given three different tags by the three annotators, and tweets that were given two or more uncertain tags. If two of the three annotators gave the tweet the same tag, the tweet was classified as having the sentiment corresponding to the tag. Using this method, we obtained 1,432 positive tweets, 976 negative tweets, and 4,737 neutral tweets (total of 7,145 tweets), and these tweets were used in the analyses. We discarded 855 tweets, of which 140 tweets were uncertain.

We examined the agreement between the objective classification using the dictionary of affective words, and subjective classification (Table 4). The overall agreement between objective and subjective classifications was approximately 60%. Evaluating the sentiment of short messages automatically is challenging [5], and the overall agreement is often low. However, the proportion of tweets classified as the opposite sentiment was only 2%, which suggests that objective classification can be used for our analysis, particularly for comparing negative and positive tweets.

4.4 Measures of Diffusion Volume and Speed

We obtained the number of retweets for each tweet and N-retweet time as measures of diffusion volume and speed, respectively. Each retweet has a timestamp and the ID of the original tweet. For each original tweet, T, we counted the number of retweets of tweet T. We obtained the N-retweet time of tweet T by calculating the in-

Table 4: Tweet sentiment obtained by subjective and objective classifications

	Positive (subj.)	Negative (subj.)	Neutral (subj.)	Uncertain (subj.)	Discard (subj.)	
Positive (obj.)	559	95	870	5	134	1,662
Negative (obj.)	69	286	384	6	93	838
Neutral (obj.)	751	513	3,321	123	440	5,149
Discard (obj.)	57	82	158	5	49	351
	1,432	976	4,737	140	715	8,000

Table 5: Variables used in regression analysis

Variable	Description
RTnum	Number of retweets
NRTtime	Time interval between original tweet and Nth retweet
pos	Categorical variable that shows the tweet is positive
neg	Categorical variable that shows the tweet is positive
follower	Number of followers
URL	Categorical variable for whether the tweet includes a URL
hash	Categorical variable for whether the tweet includes a hashtag

terval between the time tweet T was posted and the time the Nth retweet was posted.

4.5 Methods for Statistical Analysis

Initially, we examined the mean and distribution of the measures of message virality for the message sentiments. We classified all tweets as positive, negative, or neutral. For each category, we obtained the mean and distribution of the number of retweets and N-retweet time. When analyzing dataset D_S, we estimated the mean of the number of retweets in the population because dataset D_S was obtained from dataset D_A by biased sampling. The method of estimating the mean number of retweets of positive tweets was as follows. Let μ_i^p be the sample mean of the number of retweets of positive tweets in section i (Table 1) and in dataset D_S, and let f_i^p be the number of positive tweets in section i and in dataset D_A divided by the number of positive tweets in dataset D_A. The mean number of retweets of positive tweets was estimated as $\sum_i f_i^p \mu_i^p$.

Next, we performed regression analysis to investigate the effects of message sentiment on its virality considering other factors related to retweet behavior. We used the variables shown in Table 5. Following the method in Ref. [27], we used the presence of URLs, hashtags, and the number of followers as control variables because these factors affect message diffusion [22,27,28]. Using these control variables, we examined the effects of message sentiment on its virality eliminating the effects of other factors. We did not include a variable for the activity of twitter users because this does not affect message diffusion [27]. We did not use dataset D_S for regression analysis because it was obtained from biased sampling.

Following the method in Ref. [27], we used a binomial regression model for regression of *RTnum* because the variance of the number of retweets is large (Tables 1 and 2). In the negative binomial regression model, the relation between dependent and independent variables is modeled as

$$
\begin{aligned}
\log(RTnum) =& \beta_0 + \beta_1\,URL + \beta_2 hash \\
& + \beta_3 \log(follower) + \beta_4 pos + \beta_5 neg,
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
RTnum =& e^{\beta_0} \times e^{\beta_1\,URL} \times e^{\beta_2 hash} \\
& \times follower^{\beta_3} \times e^{\beta_4 pos} \times e^{\beta_5\,neg},
\end{aligned} \tag{2}
$$

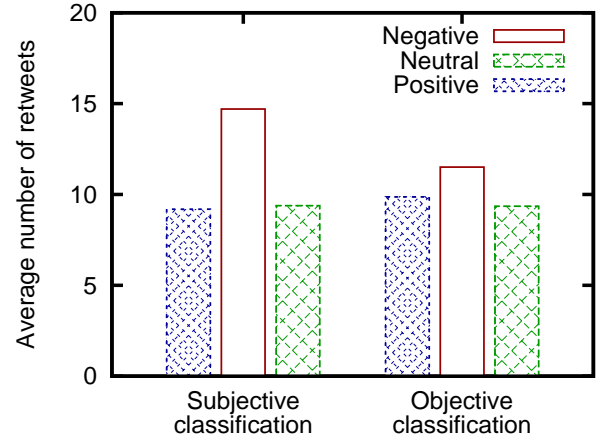Figure 1: Relation between tweet sentiment and the mean number of retweets. Left-hand bars show the estimated mean values obtained from dataset D_S, and the right-hand bars show the simple mean values obtained from dataset D_A. The population is the tweets whose number of retweets is more than one. Retweet volume of negative tweets is larger than those of positive and neutral tweets.

where β_n is the regression coefficient. Note that *follower* is log transformed because the distribution of the number of followers is heavy-tailed. For the regression of *NRTtime*, we used a simple linear regression model.

5. RESULTS

5.1 Analysis of Descriptive Statistics

To address **RQ1**, we examined the mean number of retweets for each category based on tweet sentiment (Fig. 1). Bars on the left-hand side of the figure show the results obtained from dataset D_S, and bars on the right-hand side show the results obtained from dataset D_A. The results of dataset D_S show estimated mean values that are explained in Section 4.5.

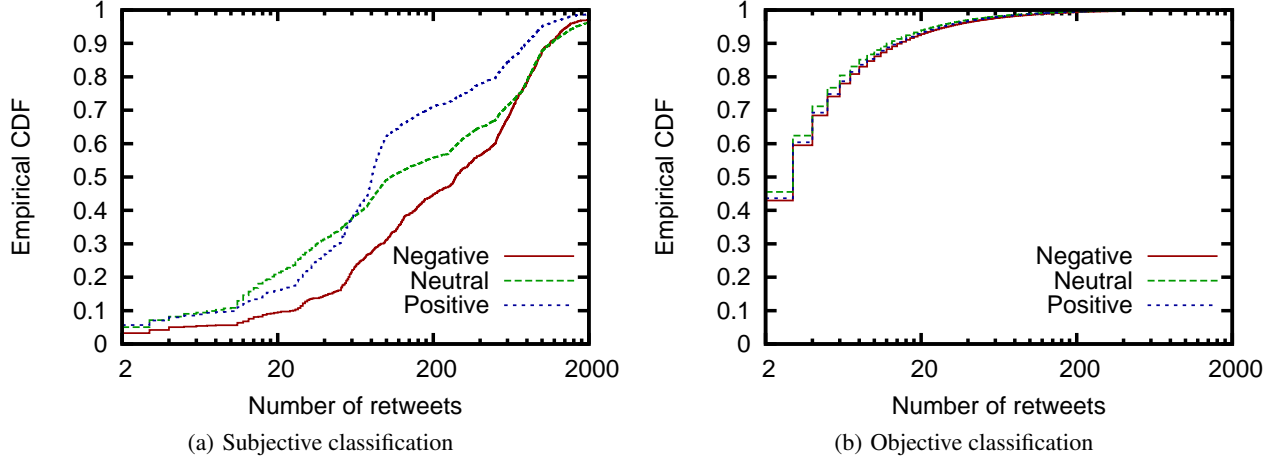(a) Subjective classification   (b) Objective classification

Figure 2: Cumulative distribution of the number of retweets for each category. Note that (a) shows the cumulative distributions of the retweet volume of the sampled tweets, not the total population. Negative tweets tend to have a larger retweet volume than positive and neutral tweets.
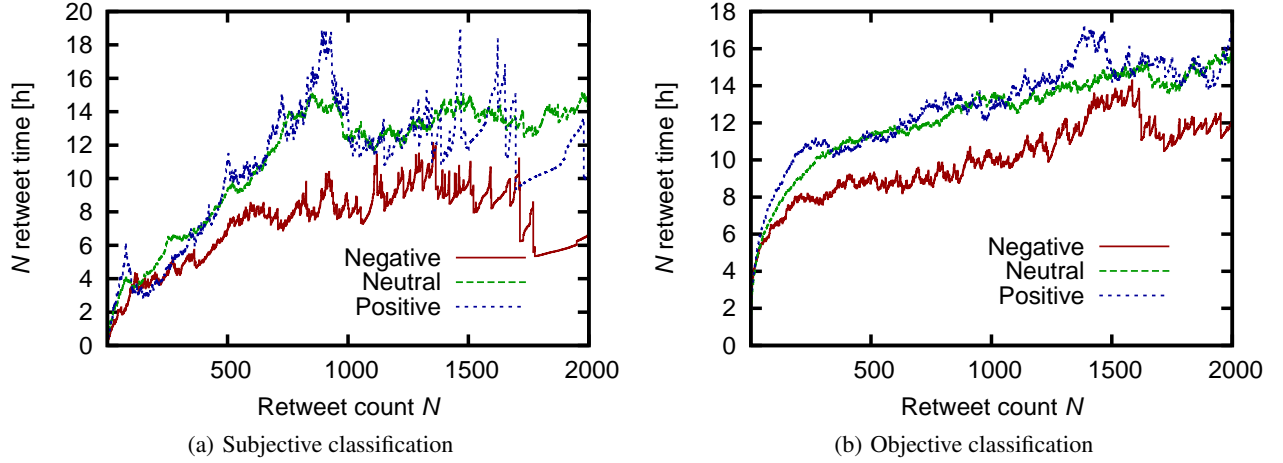

(a) Subjective classification   (b) Objective classification

Figure 3: Average N-retweet time for each category. Average N-retweet time of negative tweets is shorter than those of positive and neutral tweets.

Figure 1 shows that the retweet volume of negative tweets is approximately 1.2–1.6-fold that of neutral tweets, and the retweet volumes of positive and neutral tweets are similar to each other. We performed the pairwise test on the results of dataset D_A using the Steel-Dwass [12, 25] method, and found that there were significant differences in the number of retweets between any two categories based on sentiment ($p < 0.05$). These results suggest that the retweet volume of negative tweets is larger than that of neutral and positive tweets and the retweet volume of positive tweets is similar to neutral tweets. The differences of the mean values obtained with datasets D_S and D_A may be caused by the difference between objective and subjective classifications (Table 4).

Next, we investigated the distributions of the number of retweets for each category (Fig. 2). Figure 2 confirms that negative tweets tend to have a larger retweet volume than positive and neutral tweets. We can also find that positive tweets tend to have slightly larger retweet volume than neutral tweets (Fig. 2 (b)).

Next, we tackled retweet speed to answer **RQ2** by using average N retweet time. Figure 3 shows average N-retweet times for each

category. Average N-retweet time was obtained by calculating the average N-retweet time for tweets that were retweeted at least N times. Because the number of samples with a large retweet count, N, is limited, the average values fluctuate if N is large.

Figure 3 shows that average N-retweet time of negative tweets is shorter than those of positive and neutral tweets. In particular, when $N > 100$, the average N-retweet time of negative tweets is approximately 20% shorter than those of positive and neutral tweets. Note that the fraction of tweets retweeted more than 100 times is only 1% in the collected dataset. Namely, tweets with a retweet count of $N > 100$ have high virality in terms of diffusion volume. These results suggest that negative tweets spread faster than positive and neutral tweets, particularly for tweets with large diffusion volume. The diffusion time of negative tweets was approximately 20% shorter than that of positive and neutral tweets, namely the diffusion speed of negative tweets was about 1.25-fold that of positive and neutral tweets. In contrast, the N-retweet time of positive tweets was slightly longer than that of neutral tweets.
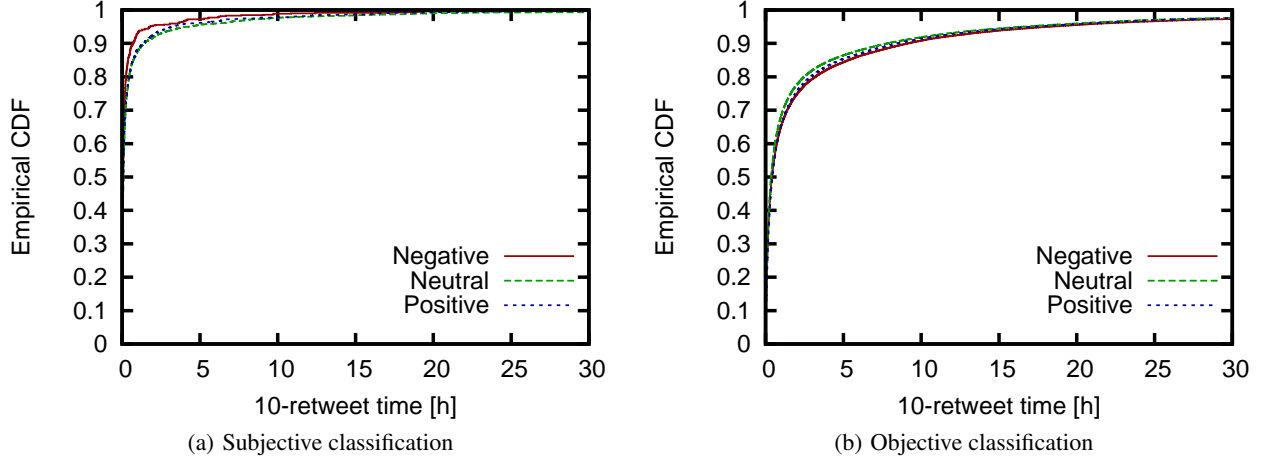
(a) Subjective classification   (b) Objective classification

Figure 4: Cumulative distribution of 10-retweet time for each category. 10-**retweet time for negative tweets and tweets with other sentiment is similar.**

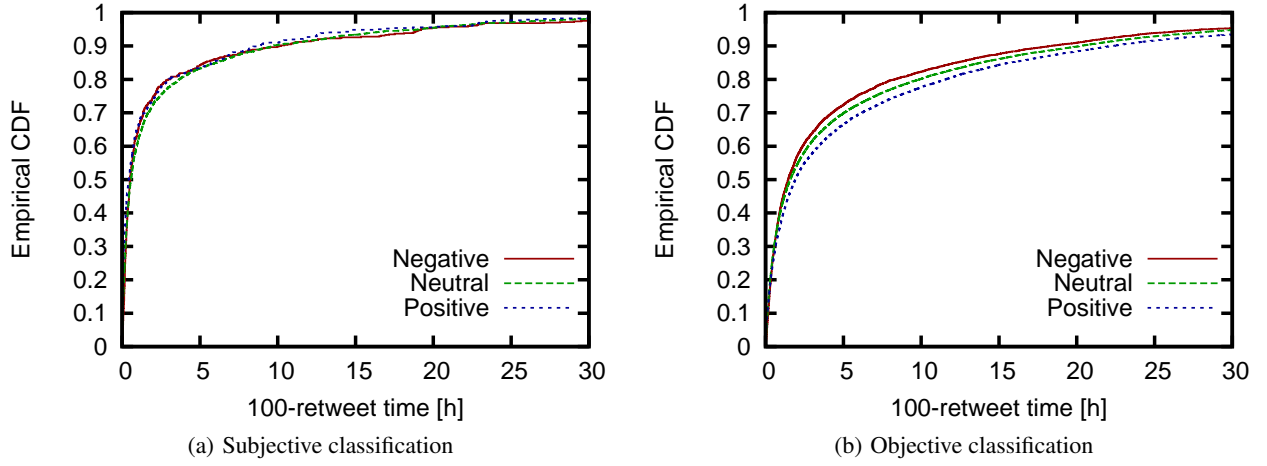(a) Subjective classification   (b) Objective classification

Figure 5: Cumulative distribution of 100-retweet time for each category. 100-**retweet time for negative tweets and tweets with other sentiment is similar.**

We investigated the distribution of N-retweet time of tweets for each category. Figures 4, 5, and 6 show the cumulative distributions of N-retweet time for each category when $N = 10, 100$, and 1000, respectively.

These results confirm that negative tweets spread faster than neutral and positive tweets do if the retweet count, N, is large. Figure 6 shows that the diffusion speed of negative tweets is faster than tweets with other sentiment when $N = 1000$. In contrast, Figs. 4 and 5 show that N-retweet time for negative tweets and tweets with other sentiment is similar. The difference in N-retweet time between positive and neutral tweets was only observed in Fig. 5(b). The pairwise test with the Steel-Dwass method [12, 25] shows that there is a significant difference in 10-, 100-, and 1000-retweet time among tweet sentiment categories in dataset D_A ($p < 0.05$).

These analyses show similar results from datasets D_S and dataset D_A, which suggests that the results are robust. For **RQ1**, our results suggest that in terms of retweet volume, negative tweets were the most viral and the virality of positive tweets was similar to neutral tweets. For **RQ2**, negative tweets spread faster than neutral and

positive tweets, particularly when the retweet count was large, and positive and neutral tweets spread at similar speeds.

5.2 Regression Analysis

The results in the previous section show that the message sentiment and virality are closely related to each other. In this section, we perform regression analysis to investigate the relation between message sentiment and virality, eliminating the effects of other factors affecting message diffusion. We performed negative binomial regression analysis for investigating the effects of message sentiment on diffusion volume. The dependent variable was *RTnum*, and the independent variables were *pos*, *neg*, *follower*, *URL*, and *hash*. Table 6 shows the regression analysis results. The regression coefficient, β, and the values of e^{β} for each variable are shown in the table to demonstrate the effects of each independent variable on the dependent variable.

The result of the regression analysis shows that whether the sentiment of a tweet is negative or positive increases its number of retweets in the model. The strength of the effect of each variable can be estimated from the regression coefficient, e^{β} (Eq.(2)).

(a) Subjective classification
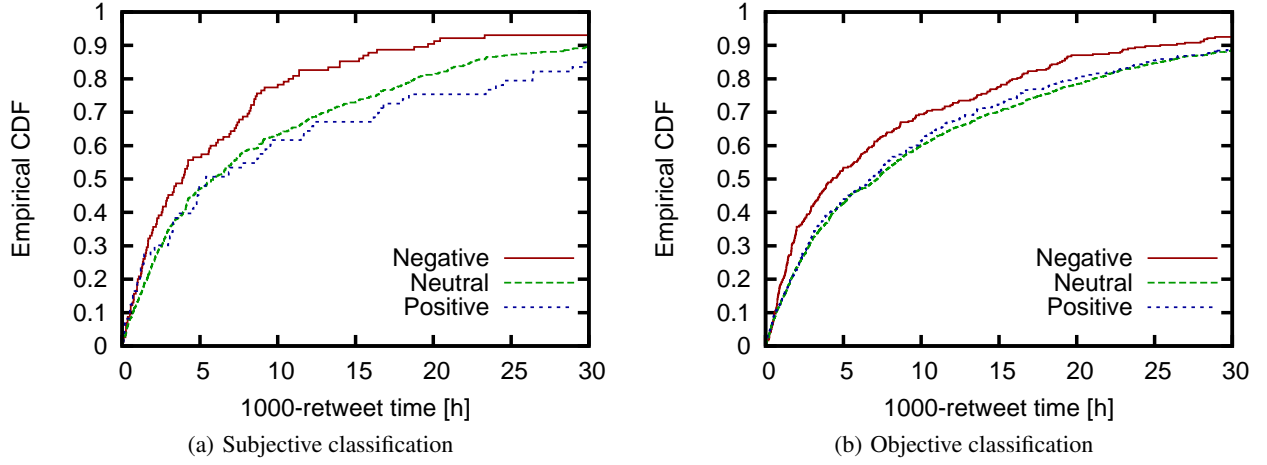

(b) Objective classification

Figure 6: Cumulative distribution of 1000-retweet time for each category. The diffusion speed of negative tweets is faster than tweets with other sentiment when $N = 1000$.

The regression coefficient of *neg* suggests that negative tweets are retweeted 36.5% more often than neutral tweets, which is consistent with the results in the previous subsection. This indicates that negative sentiment is a major driving factor of tweet diffusion, because the regression coefficient of *neg* is comparable with *hash*, which is a major driving factor for retweets [22, 28]. In addition, positive sentiment in a tweet increases retweet volume, although the effect is weaker than other factors. In summary, this result shows that negative sentiment is a strong driving factor for retweet diffusion and that positive sentiment is not a strong driving factor for retweet diffusion, although it slightly affects diffusion volume.

Note that pseudo R^2 of our model is low. Message diffusion on social media is often difficult to explain, and there are many other driving factors. In this analysis, we can conclude that the effects of negative and positive sentiment are statistically significant and the effect of negative sentiment is similar to that of hashtags. We do not claim that we can model the retweet volume only using these variables. We should also note that the value of pseudo R^2 of our model is lower than that obtained in [27]. This is because our dataset does not include tweets that are not retweeted. URLs or hashtags in tweets are strong factors affecting whether the tweets are retweeted or not [22,28]. Therefore, we can generally construct more accurate model explaining *RTNum* from these independent variables if the dataset includes tweets with no retweet than if the dataset only includes tweets with more than one retweet.

Finally, we examined the relation between message sentiment and its diffusion speed by regression analysis. We used 100-*RTtime*, and 1000-*RTtime* as dependent variables. In addition to the independent variables used in the diffusion volume regression analysis, we used *RTnum* as an independent variable. This is because tweets with a large diffusion volume are considered to spread fast. In the following analyses, a linear regression model was used. Tables 7 and 8 show the regression results for the dependent variables of 100-*RTtime*, and 1000-*RTtime*, respectively.

Table 8 indicates that the presence of negative sentiment in a message decreases the 1000-retweet time ($p < 0.1$). This result is consistent with the observation in the previous subsection that negative tweets spread fast when the number of retweets is large. Table 7 shows that the presence of negative sentiment in a message does not significantly affect 100-retweet time. This result suggests that negative sentiment does not have a significant effect on diffu-

Table 6: Negative binomial regression results for *RTnum*. ***: significant at the 1% level, **: significant at the 5% level, *: significant at the 10% level.

Dependent variable: *RTnum*		
Independent variables	Coeff. β	e^{β}
pos***	0.131	1.139
neg***	0.311	1.365
log(*follower*)***	0.203	
URL***	0.546	1.726
hash***	0.291	1.338
constant***	0.467	
Pseudo R^2		0.030
Num. of observations		4,137,064

sion speed when the diffusion volume is small. Looking at other control variables, as intuitively expected, we can find that *follower*, *URL*, and *RTnum* significantly affect diffusion speed.

These results do not show that positive sentiment increases diffusion speed. Positive sentiment in a tweet does not significantly affect 1000-retweet time and positively and significantly affect 100-retweet time.

Our findings are summarized in Table 9. We can conclude that negative tweets spread more widely than positive and neutral tweets, and it is suggested that negative tweets spread faster than tweets with other sentiments, particularly for tweets with a large diffusion volume,. The effect of positive sentiment is weaker than that of negative sentiment, although positive tweets are retweeted slightly more than neutral tweets. Moreover, the diffusion speed of positive tweets is similar to that of neutral tweets, although for tweets with a small diffusion volume, positive tweets sometimes spread slower than neutral tweets.

Table 7: Regression results for 100-*RTtime*[h]. ***: significant at the 1% level, **: significant at the 5% level, *:significant at the 10% level.

Dependent variable: 100-*RTtime* [h]	
Independent variables	Coeff. β
pos***	1.149
neg	0.052
log(*follower*)***	-0.632
URL***	1.889
hash***	1.992
RTnum***	-0.003
constant***	11.855
R^2	0.040
Num. of observations	48,814

Table 8: Regression results for 1000-*RTtime*[h]. ***: significant at the 1% level, **: significant at the 5% level, *: significant at the 10% level.

Dependent variable: 1000-*RTtime* [h]	
Independent variables	Coeff. β
pos	0.941
neg*	-1.922
log(*follower*)**	-0.331
URL***	5.055
hash	0.339
RTnum***	-0.002
constant***	17.365
R^2	0.080
Num. of observations	2,194

6. DISCUSSION

6.1 Findings and Implications

Our study shows that negative tweets are more viral than positive tweets in terms of retweet volume. This is a strong evidence of existence of negativity bias [6, 24, 31] on social media. As discussed in Section 2, prior work by Stieglitz *et al.* [27] only partly supported negativity bias, and Gruzd *et al.* [14] showed opposite results. These studies targeted domain-specific tweets, and as discussed in Ref. [27], the tweet domain alters how tweet sentiment affects the virality. However, our study investigates the effects of tweet sentiment after eliminating the effects of tweet domains. Consequently, our study shows that negative tweets are generally more viral than positive tweets, which indicates negativity bias on social media.

The results for retweet speed also partly support negativity bias. We investigated the relation between tweet sentiment and N-retweet time. For a large retweet count, N, negative tweets spread faster than positive and neutral tweets. Stieglitz *et al.* [27] only used 1-retweet time, and found that there was no significant difference in retweet speed between positive and negative tweets. Our study shows that negative tweets spread faster than positive tweets when the diffusion volume is large. To the best of our knowledge, ours is the first study to show the effects of sentiment on diffusion speed of tweets with a large diffusion volume.

Our results also show that the effects of positive sentiment in a tweet on its virality are weak. This contradicts the results in Refs. [14, 22, 26, 27] suggesting that positive and negative sentiment in a message increase its virality. One possible cause of this difference between our study and previous studies might be the nationality. Ours is the first study to use Japanese tweets to investigate the relation between tweet sentiment and virality. The language and cultural difference may affect the results because usage patterns of Twitter users differ across languages [16]. However, more analyses are necessary to reveal the cause of this.

Our results have several implications. First, it is important for companies to address negative opinions about their products on social media. Even if there are the same number of users with positive as those with negative opinions, negative opinions may spread faster and further, and thus reach a larger number of people than the positive opinions. Second, it is important to track the sentiment of individual tweets to prevent unintentional tweet diffusion. Recently, negative rumors and misinformation spread on social media, known as *flaming*, have posed serious problems, and blocking rumor spread is of interest to researchers [10, 32]. Our results suggest that individual users should take care to avoid unnecessary negative terms to prevent the unintentional information spread. A mechanism to detect and alert users to tweet sentiment may be an effective approach.

6.2 Limitations

While we used a large-scale dataset including 4.1 million tweets, it was still a sample of messages on social media. We studied Twitter as a social media platform, and only analyzed Japanese tweets. We chose Twitter because of its availability of large-scale data; however, to generalize the results, it is necessary to analyze data from other platforms. Most previous studies used English tweets [14, 15, 22], some used German tweets [26, 27], whereas we used Japanese tweets. Our study shows that for Japanese tweets, tweet sentiment is a major driving factor for retweets. However, the research methodologies of this study are different from previous studies, particularly regarding tweet topics, and Twitter usage patterns are different across languages [16]. Therefore, the differences among different languages should be investigated. For examining the generalizability of our results, we are also interested in several tasks such as expanding the data collection period, and investigating messages during several social events (e.g., national festival holidays).

We used a simple approach for objective classification of large-scale tweets based on their sentiment [15, 22]. Although we obtained similar results from the datasets constructed by objective and subjective classifications, using a more sophisticated method to determine tweet sentiment should produce better results. Because tweets are short it is difficult to determine tweet sentiment and there several studies about determining tweet sentiment accurately [2, 5, 13, 19]. In future work, we intend to apply these techniques to our dataset, and validate the results in this paper.

7. CONCLUSION

We investigated the relation between the sentiment of a tweet and its virality in terms of diffusion volume and speed by analyzing 4.1 million tweets on Twitter. We used the number of retweets and N-retweet time as measures of tweet virality. We found that negative tweets spread more widely than positive and neutral tweets, and that negative tweets spread faster than positive and neutral tweets when the diffusion volume was large. We showed that the diffu-

Table 9: Summary of findings

RQ		Conclusion	Supporting results
1: Retweet volume	Negative vs. neutral	Negative is larger	Figs. 1, 2, and Tab. 6
	Negative vs. positive	Negative is larger	Figs. 1, 2, and Tab. 6
	Positive vs. neutral	Positive is slightly larger	Tab. 6
2: Retweet speed	Negative vs. neutral	Negative is faster for large diffusion volume	Figs. 3, 6, and Tab. 8[*]
	Negative vs. positive	Negative is faster for large diffusion volume	Figs. 3, 6, and Tab. 8[*]
	Positive vs. neutral	Neutral is slightly faster for small diffusion volume	Figs. 3(b), 5(b), and Tab. 7

[*] Tab. 8 is not so strong evidence, but supports this conclusion.

sion volume of negative tweets was 1.2–1.6-fold that of positive and neutral tweets, and that the diffusion speed of negative tweets was 1.25-fold that of positive and neutral tweets when the diffusion volume was large.

Acknowledgements

8. REFERENCES

[1] Mecab: Yet Another Part-of-Speech and Morphological Analyzer. http://mecab.sourceforge.net.

[2] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau. Sentiment analysis of Twitter data. In *Proceedings of the Workshop on Languages in Social Media (LSM'11)*, pages 30–38, June 2011.

[3] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts. Everyone's an influencer: Quantifying influence on Twitter. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM'11)*, pages 65–74, Feb. 2011.

[4] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. The role of social networks in information diffusion. In *Proceedings of the 21st International Conference on World Wide Web (WWW'12)*, pages 519–528, Apr. 2012.

[5] L. Barbosa and J. Feng. Robust sentiment detection on Twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING'10)*, pages 36–44, Aug. 2010.

[6] R. F. Baumeister and E. Bratslavsky. Bad is stronger than good. *Review of General Psychology*, 5(4):323–370, Dec. 2001.

[7] J. Berger. Arousal increases social transmission of information. *Psychological Science*, 22(7):891–893, July 2011.

[8] J. Berger and K. L. Milkman. What makes online content viral? *Journal of Marketing Research*, 49(2):192–205, Apr. 2012.

[9] D. Blei, A. Ng, and M. Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993l–1022, Jan. 2003.

[10] C. Budak, D. Agrawal, and A. El Abbadi. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th International Conference on World Wide Web (WWW'11)*, pages 665–674, Mar. 2011.

[11] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In *Proceedings of the 23rd International Conference on World Wide Web (WWW'14)*, pages 925–936, Apr. 2014.

[12] M. Dwass. Some *k*-sample rank-order tests. In *Contributions to Probability and Statistics*, pages 198–202. Stanford University Press, 1960.

[13] P. Gonçalves, M. Araújo, F. Benevenuto, and M. Cha. Comparing and combining sentiment analysis methods. In *Proceedings of the first ACM Conference on Online Social Networks (COSN'13)*, pages 27–38, Oct. 2013.

[14] A. Gruzd, S. Doiron, and P. Mai. Is happiness contagious online? A case of Twitter and the 2010 Winter Olympics. In *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS'11)*, pages 1–9, Jan. 2011.

[15] L. Hansen, A. Arvidsson, F. Nielsen, E. Colleoni, and M. Etter. Good friends, bad news - Affect and virality in Twitter. *Future Information Technology*, 185:34–43, Dec. 2011.

[16] L. Hong, G. Convertino, and E. H. Chi. Language matters in Twitter: A large scale study. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM'11)*, pages 518–521, July 2011.

[17] L. Hong, O. Dan, and B. Davison. Predicting popular messages in Twitter. In *Proceedings of the 20th International Conference on World Wide Web (WWW'11)*, pages 57–58, Apr. 2011.

[18] D. Kempe, J. M. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'03)*, pages 137–146, Aug. 2003.

[19] E. Kontopoulos, C. Berberidis, T. Dergiades, and N. Bassiliades. Ontology-based sentiment analysis of Twitter posts. *Expert Systems with Applications*, 40(10):4065–4074, Aug. 2013.

[20] A. Kupavskii, L. Ostroumova, A. Umnov, S. Usachev, P. Serdyukov, G. Gusev, and A. Kustarev. Prediction of retweet cascade size over time. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM'12)*, pages 2335–2338, Oct. 2012.

[21] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th*

International Conference on World Wide Web (WWW'10), pages 591–600, Apr. 2010.

[22] N. Naveed, T. Gottron, J. Kunegis, and A. Alhadi. Bad news travel fast: A content-based analysis of interestingness on Twitter. In *Proceedings of the ACM Web Science Conference 2011 (WebSci'11)*, pages 1–7, June 2011.

[23] S. Rosenthal, P. Nakov, S. Kiritchenko, S. M. Mohammad, A. Ritter, and V. Stoyanov. SemEval-2015 task 10: Sentiment analysis in Twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval'15)*, pages 451–463, June 2015.

[24] P. Rozin and E. B. Royzman. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4):296–320, Nov. 2001.

[25] R. G. D. Steel. A rank sum test for comparing all pairs of treatments. *Technometrics*, 2(2):197–207, May 1960.

[26] S. Stieglitz and L. Dang-Xuan. Political communication and influence through microblogging—an empirical analysis of sentiment in Twitter messages and retweet behavior. In *Proceedings of the 45th Hawaii International Conference on System Science (HICSS'12)*, pages 3500–3509, Jan. 2012.

[27] S. Stieglitz and L. Dang-Xuan. Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4):217–247, 2013.

[28] B. Suh, L. Hong, P. Pirolli, and E. Chi. Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In *Proceedings of the 2nd IEEE International Conference on Social Computing (SocialCom'10)*, pages 177–184, Aug. 2010.

[29] H. Takamura, T. Inui, and M. Okumura. Extracting semantic orientations of words using spin model. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL'05)*, pages 133–140, June 2005.

[30] H. Takamura, T. Inui, and M. Okumura. Extracting semantic orientations using spin model. *IPSJ Journal*, 47(2):627–637, Feb. 2006. (in Japanese).

[31] S. E. Taylor. Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110(1):67–85, July 1991.

[32] S. Wen, J. Jiang, Y. Xiang, S. Yu, W. Zhou, and W. Jia. To shut them up or to clarify: Restraining the spread of rumors in online social networks. *IEEE Transactions on Parallel & Distributed Systems*, 25(12):3306–3316, Dec. 2014.

[33] S. Yang, A. Kolcz, A. Schlaikjer, and P. Gupta. Large-scale high-precision topic modeling on Twitter. In *Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'14)*, pages 1907–1916, Aug. 2014.