Emotional Divergence Influences Information Spreading in Twitter

René Pfitzner and **Antonios Garas** and **Frank Schweitzer**
Chair of Systems Design, ETH Zurich
Kreuzplatz 5, 8032 Zurich
Switzerland

Abstract

We analyze data about the micro-blogging site Twitter using sentiment extraction techniques. From an information perspective, Twitter users are involved mostly in two processes: information creation and subsequent distribution (*tweeting*), and pure information distribution (*retweeting*), with pronounced preference to the first. However a rather substantial fraction of tweets are retweeted. Here, we address the role of the sentiment expressed in tweets for their potential aftermath. We find that although the overall sentiment (polarity) does not influence the probability of a tweet to be retweeted, a new measure called *emotional divergence* does have an impact. In general, tweets with high emotional diversity have a better chance of being retweeted, hence influencing the distribution of information.

1 Introduction

An important step to understand how today's society work is to understand how people communicate and share information. Especially communication via social networking platforms is of profound interest to the scientific community . With this manuscript we contribute to the recent interest in analyzing social media and micro-blogging websites. Especially, we consider the micro-blogging platform *Twitter*[1]. Twitter has become a valuable source for quantitative socio-researchers in the last few years. One of the first quantitative, data-driven studies of *Twitter* has been presented by Huberman et al. (Huberman, Romero, and Wu 2009), where the social network of users within the platform was linked to their tweeting-behavior. Later studies addressed more questions concerning the structure of the underlying social network and the influence of users (Kwak et al. 2010; Cha et al. 2010), the dynamics of tweets and topics (Zaman et al. 2010; Kwak et al. 2010; Huang and Thornton 2010), the sentiment expression in tweets (Thelwall, Buckley, and Paltoglou 2011; Pak and Paroubek 2010; Bollen, Pepe, and Mao 2011) or the demographics of twitter users (Mislove et al. 2011). Twitter and other social mediae provide scholars with a valuable source of data that could be used for sociological research, as for example in (Goncalves, Perra, and Vespignani 2011) where the socio-cognitive concept of Dunbar's number has been validated. In (Bollen, Mao, and Zeng 2011) a correlation between the swing of collective mood as expressed in Twitter messages and a particular stock market index was identified, which is now used at least in one hedge fund[2]. In (Asur and Huberman 2010) Twitter messages were used to forecast the box-office revenues of movies based on the volume of their mentions on the micro-blogging platform. One important value of Twitter and other public social media platforms is that they allow us to *listen* to (public) conversations and to *infer* people's opinions or moods based on their public statements. This has for example been illustrated by the *Pulse of the Nation* study[3] or by (Golder and Macy 2011). In general, Twitter and other similar social mediae provide us with insights into people's social behavior and their individual reflections of our everyday life on a scale that has never been possible before. To evaluate this huge amount of information poses a hard and important problem for the quantitative sciences, and our manuscript contributes to this latter undertaking. By combining statistical analysis with modern state-of-the-art sentiment classification technologies, we identify the role of emotions in the processes of information spreading over Twitter.

2 Data and sentiment classification

Our study is based on public messages distributed by users of the micro-blogging platform Twitter. Twitter is a social media service which allows users to create a profile and write up to 140 character long messages, which are called tweets and which are (mostly) publicly available. A social network of users is built as one can choose to *follow* all messages of other users. In the remaining, we will refer to the user who follows another user as *follower* and to the user who is being followed as *followee*. The dataset we use for this study was introduced in (Thelwall, Buckley, and Paltoglou 2011), and contains almost 35 Mio. tweets in English language from February to March 2010 coming from almost 3 Mio. different accounts. It was provided from the data company *Spinn3r* as part of their free access program for researchers. To analyze the sentiment expressed in tweets, we use the SentiStrength (Thelwall et al. 2010) classifier, which

[1]http://www.twitter.com

[2]http://www.derwentcapitalmarkets.com/

[3]http://www.ccs.neu.edu/home/amislove/twittermood

was built especially to cope with sentiment detection in short informal text. SentiStrength combines a lexicon-based approach with more sophisticated linguistic rules, e.g. taking care of negotiations, misspellings, the "sloppy" use of words like *love* in every-day language (*I love the Internet*) as well as punctuation or emoticons; for details see (Thelwall et al. 2010). It is thus much better suited for analyzing sentiment of tweets than e.g. using the LIWC (Linguistic Inquiry and Word Count, www.liwc.net) library, which was complied for coping with longer documents (Thelwall et al. 2010). The SentiStrength classifier assigns two values to each tweet: a measure of positive and a measure of negative sentiment, both on absolute integer scales ranging from 1 to 5, with 1 denoting no sentiment and 5 denoting high sentiment value. To be more precise:

- The positive sentiment score p, with $p \in [1, 5]$, is basically equal to the sentiment score of the most positively classified word in the tweet, adjusted by linguistic rules.
- The negative sentiment score n, with $n \in [-5, -1]$, is basically equal to the sentiment score of the most negatively classified word in the tweet, adjusted by linguistic rules.

An overall binary classification, b, of the tweet (i.e. whether it is positive $b = 1$ or negative $b = -1$ or neutral $b = 0$) is given based on:

$$b = \begin{cases} 1 & \text{if } |n| < |p| \\ -1 & \text{if } |n| > |p| \\ 0 & \text{if } |n| = |p| \end{cases} \tag{1}$$

where the binary classification b (sometimes called *sentiment polarity*) gives the overall sentiment classification of a tweet.

3 Tweeting and Retweeting

The creation and distribution of potentially new (publicly available) information on Twitter is called *tweeting*. An interesting subset of tweets are the so called *retweets*. Retweets are basically forwarded tweets: if a user likes a tweet by one of his followees he can choose to distribute it to his followers. Properties of retweets have been studied before e.g. in (Boyd, Golder, and Lotan 2010; Kwak et al. 2010) and (Zaman et al. 2010), where latter proposed a probabilistic model to predict the success of a tweet based on it's statistics and it's senders social network. With this manuscript we go beyond the standard statistical analysis. Instead we focus on the role of *sentiment*, and especially of *emotional divergence* (which will be introduced later), on the possible success of a tweet. We define a tweet to be successful if it is retweeted. The more often a tweet is retweeted, the more successful it is. In the data of all tweets, a retweet can be recognized if it is a regular expression of the kind *RT* $\{$*user name*$\}$*:*$\{$*text*$\}$.

Retweets are not rare

In our data set we find that 9% of all tweets are retweets - a number that is notably more than the 3% reported earlier (Boyd, Golder, and Lotan 2010). Based on our result, we conclude that retweets make up indeed a substantial fraction of all traffic in Twitter. From a research point of view, this is interesting in various respects: first, retweets are some kind of word-of-mouth effect and the beneficiary is the user who originally shared the tweet. Hence, a natural question to ask is: "What do I have to do in order to have my tweets retweeted by many users?" Second, retweets are a subset of all tweets. As such, it is interesting to study whether the underlying user dynamics are different, i.e. is retweeting governed by other processes than tweeting? We will address the second question in the next subsection.

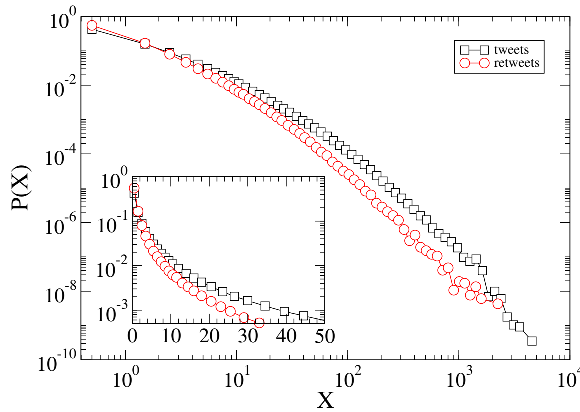

Figure 1: Probability density functions of number of tweets and number of retweets sent per user, given that the user tweeted or retweeted at least once. The data is nicely fit by $P_t(x) \sim \ln \mathcal{N}(0.173, 2.672)$ ($R^2 = 0.999$) and $P_r(x) \sim \ln \mathcal{N}(-0.281, 1.912)$ ($R^2 = 0.999$). Inset: Zoom of the small X area, where we see that the two distributions already start to diverge.

Retweets and Tweets originate from the same dynamics

The underlying microscopic dynamics of a process are reflected in its statistics. To obtain a first insight in the dynamics of tweeting and retweeting, we hence study the statistics of tweets and retweets. In figure 1 we show the probability density function $P(X)$ that a user posts X tweets and retweets, respectively. For both statistics we only consider *active* users, i.e we exclude users with zero tweets/retweets. Our data suggests that most users tweet only very few times, if they tweet at all. Indeed, the average number of tweets (arithmetic mean) is 7.4, the median is only 2 tweets. Both distributions are heavy tailed, implying that there are only few individuals who tweet a lot.We note that these distributions are not power-laws. Indeed, they are both better approximated by similar log-normal distributions. From the theory of stochastic processes, it is well known that log-normal distributions arise when the underlying dynamics are so called *multiplicative stochastic processes*. Hence, for a future modeling approach of the process of tweeting and retweeting, a framework based on multiplicative stochastic processes would be suitable. However, writing down such a model is beyond the scope of this paper.

Emotional Content of Tweets and Retweets

Most quantitative studies dealing with social media and especially Twitter are centered around statistical, structural or technical questions. However, an interesting question that is still open is about the type of content that usually is retweeted. Here we apply *SentiStrength* (Thelwall et al. 2010) to get an insight about the emotional content of retweets. We ask two principal questions:

- Are emotionally *positive, negative or neutral* tweets more likely to be retweeted? This question has previously been asked in (Hansen et al. 2011), based on a much smaller data set though. Also, tweets were classified by the rather generic ANEW classifier (Bradley and Lang 1999).
- Are *emotionally diverse* tweets more or less likely to be retweeted? This question extends the first one. It however shifts focus from the overall sentiment of a tweet to the emotional divergence encountered in the tweet, which we introduce below.

To address the first question, we consider emotional polarity, i.e. the overall emotional value of tweets and retweets. In table 1 we show the results of this analysis, suggesting that there is no significant difference in emotional polarity for tweets and retweets. This is in line with (Hansen et al. 2011). We should stress here that the emotional classification shown in table 1 provides support to what is known as the *Pollyanna Hypothesis* (Boucher and Osgood 1969), which states that there is a bias towards positivity in human language expression, in line with another recent analytical result (Garcia, Garas, and Schweitzer 2011).

	-1	0	$+1$
pure tweets	19.9%	33.8%	46.3%
retweets	19.8%	31.4%	48.8%

Table 1: Percentage of negative, neutral and positive sentiment in pure tweets and retweets in our data set. With $p = 0.1991$ (χ^2-test), we cannot reject the null hypothesis that emotional polarity of tweets and retweets is identically distributed.

Emotional divergence

In order to address the second question, we introduce the notion of *emotional divergence*. We define emotional divergence as the (normalized) absolute difference between the positive and the negative sentiment score delivered by SentiStrength:

$$d = \frac{p - n}{10}. \tag{2}$$

Whereas emotional polarity measures the *overall* emotion expressed in a text, emotional divergence measures the (extreme) *span* of expressed emotions. For example, according to SentiStrength the sentence "*I love hating you*" will be classified as *love=3* and *hate=-4*, hence resulting in binary emotional polarity $b = -1$. However, there might be high contrast in the emotional information of the used words (as is the case in the example) and emotional divergence

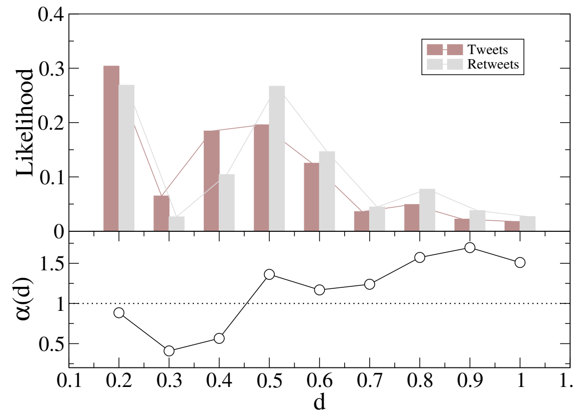

Figure 2: Emotional divergence of tweets and retweets. (Top) Likelihood of tweets and retweets to have emotional divergence (d). (Bottom) Likelihood ratio (α) .

is able to capture this effect. Indeed, in the example sentence emotional divergence is very strong $d = 0.7$, reflecting the emotionally contrasting nature of the statement. Emotional divergence can be considered as a natural extension of emotional polarity, going beyond measuring the overall emotional content. Instead it captures the extent of emotional expression present in text. Given that positive sentiment is measured by integer values $p \in [1, 5]$, and negative sentiment is measured by integer values $n \in [-1, -5]$, emotional divergence d will be an integer value $d \in [0.2, 1]$. As we pointed out earlier, emotional polarity does not play a role for the spreading probability of information in form of retweets. However, we expect emotional divergence to have an effect and hence to add some deeper insight in the role of emotions in information distribution. In figure 2 (top) we show the histogram of emotional divergence for tweets and retweets in our data set. We find that content with high emotional divergence is more likely to be placed among retweets than among pure tweets (i.e. tweets that were not retweeted). This effect is visualized even better by the likelihood ratio

$$\alpha(d) = \frac{P(d|R)}{P(d|T)}, \tag{3}$$

where $P(d|R)$ and $P(d|T)$ are respectively the likelihoods of retweets and tweets to show emotional divergence d, see figure 2 (bottom). There is a clear threshold of $d = 0.4$ dividing the low-divergence regime from the high-divergence regime. A message with emotional divergence $d \leq 0.4$ is much more likely for tweets, whereas a message with higher emotional divergence $d > 0.4$ is much more likely to be found in the set of retweets. The ratio α has a peak at value $d = 0.9$. The likelihood of finding a retweet (over the whole set of retweets) with emotional divergence $d = 0.9$ is $\alpha \sim 1.7$ times higher than finding a simple tweet with same emotional divergence. From the likelihoods shown above, we are able to make predictions about the retweeting probabilities via Bayes theorem:

$$P(R|d) = \frac{P(d|R)P(R)}{P(d)}. \tag{4}$$

With the retweet probability $P(R) = 0.09$ and $P(d) = P(d|R)P(R) + P(d|T)P(T)$, we find that a tweet with $d = 0.9$ has an absolute $P(R|d = 0.9) = 0.14$ chance of being retweeted, whereas a tweet with $d = 0.3$ performs much worse $P(R|d = 0.3) = 0.03$. The existence of a clear and consistent threshold is quite remarkable. Hence, emotional divergence, as a measure of contrary emotional content, is a suitable distinguishing feature for the success of tweets.

4 Discussion

Based on the micro-blogging platform Twitter, we find that emotional divergence adds some additional insight into the question of which tweets are most successful in being retweeted. Also, we should note that the observed usefulness of emotional divergence might be a) special to Twitter and the users engaged in the service and b) a finite size effect, due to the limited length of a tweet. Indeed, in longer texts one would expect emotional divergence to always be high simply due to statistics, i.e. the increased chance of using a very negative and a very positive word. This conjecture remains to be tested. Also, our study treats every Twitter user as equally successful, neglecting the underlying structure of the social network. It would be interesting to include this as well as other dimensions (e.g. is the tweet news/non-news, does it feature a link, etc.) to the problem.

5 Summary

We study the phenomenon of retweeting in the micro-blogging platform Twitter using sentiment classification techniques. We first show that retweets and tweets are most likely due to the same dynamics and that retweets make up a rather substantial fraction of all tweets. We then provide evidence that a new measure called emotional divergence yields insight into the retweet probability of a tweet. We show that highly emotional diverse tweets can have up to almost five times higher chances of being retweeted.

6 Acknowledgements

This work was supported by a European Union grant by the 7th Framework Programme, Theme 3: Science of complex systems for socially intelligent ICT. It is part of the CyberEmotions project (contract 231323).

References

Asur, S., and Huberman, B. 2010. Predicting the future with social media. In *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 492–499. IEEE.

Bollen, J.; Mao, H.; and Zeng, X.-j. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2(1):1–8.

Bollen, J.; Pepe, A.; and Mao, H. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In *Fifth International AAAI Conference on Weblogs and Social Media*.

Boucher, J., and Osgood, C. 1969. The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*.

Boyd, D.; Golder, S.; and Lotan, G. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *Proceedings of the 43rd Hawaii International Conference on Systems Sciences*.

Bradley, M., and Lang, P. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida.

Cha, M.; Haddadi, H.; Benevenuto, F.; and Gummadi, K. P. 2010. Measuring user influence in twitter: The million follower fallacy. *4th International AAAI*.

Garcia, D.; Garas, A.; and Schweitzer, F. 2011. Positive words carry less information than negative words. *Arxiv preprint arXiv:1110.4123v1*.

Golder, S. a., and Macy, M. W. 2011. Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science* 333(6051):1878–1881.

Goncalves, B.; Perra, N.; and Vespignani, A. 2011. Validation of Dunbar's number in Twitter conversations. *Arxiv preprint arXiv:1105.5170*.

Hansen, L.; Arvidsson, A.; Nielsen, F.; Colleoni, E.; and Etter, M. 2011. Good friends, bad news-affect and virality in twitter. *Future Information Technology* 34–43.

Huang, J., and Thornton, K. 2010. Conversational tagging in twitter. *Proceedings of the 21st* 173.

Huberman, B.; Romero, D.; and Wu, F. 2009. Social networks that matter: Twitter under the microscope. *Available at SSRN: http://ssrn.com/abstract=1313405* 14(1):8.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? *Proceedings of the 19th international conference on World wide web - WWW '10* 591.

Mislove, A.; Lehmann, S.; Ahn, Y.; Onnela, J.; and Rosenquist, J. 2011. Understanding the Demographics of Twitter Users. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*.

Pak, A., and Paroubek, P. 2010. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In Chair), N. C. C.; Choukri, K.; Maegaard, B.; Mariani, J.; Odijk, J.; Piperidis, S.; Rosner, M.; and Tapias, D., eds., *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA).

Thelwall, M.; Buckley, K.; Paltoglou, G.; Cai, D.; and Kappas, A. 2010. Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* 61(12):2544–2558.

Thelwall, M.; Buckley, K.; and Paltoglou, G. 2011. Sentiment in Twitter Events. *Journal of the American Society for Information Science* 62(2):406–418.

Zaman, T. R.; Herbrich, R.; van Gael, J.; and Stern, D. 2010. Predicting information spreading in twitter. *Workshop on Computational Social Science and the Wisdom of Crowds (NIPS 2010)*.