

# Is Happiness Contagious Online?

## A Case of Twitter and the 2010 Winter Olympics

Anatoliy Gruzd  
Dalhousie University  
gruzd@dal.ca

Sophie Doiron  
Dalhousie University  
sp936215@dal.ca

Philip Mai  
Dalhousie University  
philipmai@dal.ca

### Abstract

*Is happiness contagious online? To answer this question, this paper investigates the posting behavior of users on Twitter.com, a popular online service for sharing short messages. Specifically, we use automated sentiment analysis to study a large sample of over 46,000 Twitter messages that reference the 2010 Winter Olympics. We determined that there are more positive messages than negative, and that positive messages are 3 times more likely to be forwarded than negative messages. However, we were not able to confirm with a reliable degree of certainty that the emotional context of messages is directly related to the user's position in the Twitter network. It is likely that there are other factors involved as well. For example, we found that negative users were more prolific posters than positive users, suggesting their more argumentative and passionate nature. This paper concludes with some implications for the Twitter community and a description of our follow-up study.*

### 1. Introduction

Is happiness contagious? Can it spread from person to person? If we are talking about a network of locally bounded individuals with many face-to-face interactions consisting of “strong” ties, the answer is a definite “yes”, according to Fowler and Christakis (2008). However, it is still unclear if the same can be said for online social networks consisting of people who live in far flung places with few or even no face-to-face interactions and where the majority of ties are considered to be “weak”. To shed some light on the latter question, this paper investigates posting behavior of users on Twitter.com, a popular online service for sharing short messages called “tweets”. Twitter makes an ideal case to determine whether happiness is contagious online because of the instantaneous broadcast-ability of the messages and the built-in inter-connectivity of its membership base.

As part of their now famous Framingham Heart study, Fowler and Christakis followed 4,739 people and observed their levels of happiness over the course of 20 years. To determine levels of happiness, they relied on a commonly used scale called the Center for Epidemiological Studies Depression scale (CES-D) developed by the United States National Institute of Mental Health. This scale works by comparing each participant's varying levels of happiness with respect to the different types of social relationships in their lives. From their study, the researchers were able to determine that the place occupied by the participants within their social network played a significant part in their level of happiness; more specifically, “people at the core of their local networks seem more likely to be happy, while those on the periphery seem more likely to be unhappy” (p. 6) [7]. In addition, the researchers were able to determine that the level of happiness of the participants' close relationships had a direct influence on their own overall level of happiness; in fact, the results showed that clusters of happiness can and do form within real-world, face-to-face social networks, and that happiness spreads from person to person within these networks and that the happiness extends up to three degrees of separation. The question at hand is whether or not happiness can also spread within online social networks in a similar way.

In our research, we seek to determine patterns in the spread of happiness on the internet (specifically on Twitter) and to compare these patterns with those found in the Christakis and Fowler study. Our research questions include the following:

- How do we automatically measure the positive or negative polarity of twitter messages?
- Do Twitter users tend to post primarily positive or negative messages?
- Are positive messages more likely to be forwarded than negative messages?

- Can a user's position within the network be an indicator of that user's tendency to post positive or negative messages?

## 2. Background

### 2.1. Happiness in online communities

With the growing number of social networks now being created on the internet, the definition of social network and of community is rapidly changing. There are now countless places on the internet where individuals can go to interact with each other. Such internet spaces include chat rooms, social networking sites, online support groups, blog spots, and other online communities that bring together people who share common interests. As Easter argues, these online spaces fulfill the basic need for a "third place" for people to spend their time outside of their home and work environments [6]. Before the existence of the internet, this concept of a 'third place' would have been anything from a coffee shop to a book store; now, with the internet being widespread and relatively accessible, more and more individuals are turning to the internet as a 'third place' to spend their time. Some individuals seek internet communities as a leisure activity, as is the case in Easter's work; others turn to the internet in the hopes of finding support groups to help them with difficult life situations, or to find new ways of establishing relationships and connecting with loved ones.

Many researchers in recent years have tried to determine the driving forces behind the growing number of online communities; in looking at factors that motivate people to turn to online communities, researchers have found that one of the strongest motivators was the positive emotions that these communities foster. Many results show that participating in online communities can reduce feelings of isolation and increase levels of happiness of the participants. For example, in the study conducted by Sum et. al., the results showed that internet use increases the well being of older adults by reducing the isolation and boredom that results from a lack of meaningful friendships [17]. It was found that internet use by older adults is effective in keeping them connected to loved ones when limited mobility becomes a factor in their lives. The authors explain that "travel through cyberspace does not require physical movement, so even elders with disabilities can contact with social networks from their home" (p.1). This

study also found that the internet can "reinforce people's connection with their surrounding social world [and] possibly reduce the onset of depression" (p. 3).

As family and community structures are collapsing in western societies, the internet offers alternative ways of connecting with one another in a world where individuals are becoming increasingly independent and socially isolated [6]. In this context, happiness is equated with feeling connected and maintaining meaningful relationships, both of which are enabled and/or enhanced by the internet. This concept of happiness is also reflected in the study by Han et. al., in which women with breast cancer experienced improvements in their sense of well being by participating in an online support group [10]. These online support groups have been proven to be effective in "reducing patients' depression, stress, and cancer-related trauma" (p. 1003); by allowing people with similar medical conditions to come together and support each other over the internet regardless of geographical, social, or economic distance, these online support groups enable and enhance the healing process through connectivity and shared experience.

These studies clearly demonstrate that happiness and well-being can be generated by the participation in online communities and social networks; however, what these studies do not identify is whether or not happiness can be spread via online social networks in the same way as it does in face-to-face networks as shown by Fowler and Christakis' study. Several of the studies on internet use and happiness have alluded to the fact that the participation in an online social network generates happiness in ways that are similar to that of face-to-face interactions; however, there is little research into whether or not happiness can spread like a biological contagion within online social networks. But before we can delve more deeply into answering this question,, we must first determine the best approach to take on such an undertaking.

### 2.2. Methods used to measure happiness on the internet

Previous studies on happiness in relation to the internet have employed a wide variety of methods to gauge happiness. One of the most common approaches used by researchers to measure happiness on the internet is to use a survey instrument. The study by Sum et. al. on the participation of older adults on the

internet utilized a survey to measure the following: intensity and history of internet use, internet breadth scale, changes in life satisfaction, and community scale [17]. By measuring these components, the researchers were able to determine whether or not the internet was increasing the well-being of the participants. The study of online support groups for women with breast cancer employed a similar method, using a series of tests both before and after the participants were given access to the Comprehensive Health Enhancement Support System (CHESS). By evaluating the participants' levels of happiness before and after their use of the CHESS, the authors were able to determine that the expression of emotions within the online support group led to an increase in well-being for the participants. Though the use of surveys was effective for these studies, surveys have many flaws and are not always sufficiently accurate. They are also very time consuming and unwieldy for a large-scale sample. Both of the studies discussed above were conducted within a relatively small sample of participants (222 participants for the first study, and 231 for the second one), and so the use of a survey was appropriate. Additionally, surveys only serve to measure participants' perceived levels of happiness or well-being, and self-perception is often inconsistent or inaccurate.

The second approach commonly used by researchers to measure happiness and other sentiments online is to use human coders to rate a specific piece of text using a detailed coding system. Berger and Milkman, in their study of the most-emailed New York Times articles, utilized human coders to "quantify the extent to which each article contained practical information, inspired awe, or evoked surprise" (p.11) [1]. Coders were given detailed descriptions of each article type (i.e practical, awe inspiring, or surprising), and were asked to rate each sample article according to these definitions using a 1-5 point scale. Scores were averaged across all the coders, and then standardized. By looking at the scores and the rate at which the articles were emailed, along with information about where the articles were situated on the New York Times website, researchers were able to determine which types of articles are most likely to be emailed. They found that affect-laden and positive articles are more likely to be the most e-mailed stories on any given day. (p. 2).

The third approach used by researchers is an automated text analysis called opinion or sentiment

analysis that uses a computer algorithm to process and analyze text in order to determine its polarity. For example, Crockett relied on this method to crawl the blogosphere in search of statements that represent different emotions [3]. In another study conducted by Dodds & Danforth [5], the researchers relied on a scale called the Affective Norms for English Words (ANEW). This scale allows researchers to come up with a very specific emotional score for any piece of text, thus measuring the happiness (or unhappiness) of song lyrics, blog posts, public addresses, etc. For an extensive review of different sentiment analysis techniques see [15], and [14, 18] specifically for the sentiment analysis of informal messages on the internet. In short, the automated method is more scalable to larger datasets and more objective than the survey method or human coders.

In sum, because Twitter has millions of members and because of the dynamic nature of the Twitter membership, we deemed the first and second approaches to be impractical for our purposes. Instead, we decided to rely primarily on sentiment analysis. The next section will describe our method in more detail.

### 3. Method

Twitter was a logical choice for this research as most of the tweets are publicly available and are easily retrieved using Twitter API. Twitter was also appealing because of how quickly messages spread from user to user. The fast-paced nature of this medium allowed us to observe in real time the spread of happiness through messages being forwarded (or "retweeted").

Since it is impossible to collect completely unbiased tweets that are not influenced by some local or global events on Twitter, we decided to collect messages on a single popular event. This allowed us to control our results for this particular event. Furthermore, by collecting messages on the same topic, we can compare contradicting opinions being expressed about the same event and determine which one has a higher chance of being passed on.

With this consideration, we chose to narrow down our sample of messages to include only messages that mentioned the 2010 Winter Olympics. This subset of tweets was chosen because this event was a well covered, popular event that garnered a lot of attention in the popular press and online media, including Twitter. There was no shortage of tweets that mentioned the Olympics, and because of the

competitive nature of the event, we could expect that a large proportion of the messages would have strong emotional content. We anticipated that many of the messages would have either very positive or very negative polarities, given the strong emotional investment that most fans tend to make in support of their national teams and athletes. For example, there were many messages congratulating winning teams or expressing disappointment in teams or athletes whose performances were poor.

The messages were collected from February 12, 2010 (the first day of the Olympics) to March 4, 2010 (a few days after the closing ceremony) by sending an automated request to Twitter search API every hour to retrieve the 100 most recent tweets that mentioned the word “Olympics”. For this period, we collected 46,097 tweets.

As mentioned in the previous section, we chose to use automated sentiment analysis techniques to identify polarity of Twitter messages (neutral, negative, positive or both). And since the goal of this research is neither to develop a new sentiment analysis system nor to improve an existing one, we chose to use an “off-the-shelf” system called SentiStrength (v2.1), developed by a team of researchers from the University of Wolverhampton in the UK, available at <http://sentistrength.wlv.ac.uk>. Although there is a number of other open and commercial text mining and natural language processing tools that can perform sentiment analysis such as Lexalytics and Opinion Observer, SentiStrength was specifically designed to analyze informal short online messages. Based on the formal evaluation of this system (conducted by the developers) on a large sample of status updates from Myspace.com, a popular social networking site, the accuracy of predicting positive and negative emotions was somewhat similar to that of other methods/systems (72.8% and 60.6% for negative and positive emotions correspondingly, based on a strength scale of 1–5), and as compared to other available methods, SentiStrength showed significantly the highest correlation with the human coders [18].

SentiStrength works as follows. The system assesses each message separately on positive and negative scales and returns two numbers: a positive polarity value (1 to 5) and a negative polarity value (-1 to -5). The advantage of having two polarity values is that it makes it possible to identify messages that have both positive and negative emotions at the same time.

These are the messages for which absolute values are equal to each other and greater than 1. For instance,

“awesome opening on CTV for the 2010 olympics, so America why is NBC not showing it live?”.

To differentiate between messages that include strong sentiments versus those that are subtle, we only focused on messages with a polarity greater than 2 or less than -2. These messages were then deemed positive if the positive value was higher than the negative absolute value. Here is an example of a positive message with the SentiStrength values of 3 and -1:

“I friggin' love the Olympics. I never thought I could care so much for people or sports that I just learned about six minutes beforehand.”

Alternately, messages were deemed negative if the negative absolute value was higher than the positive value. For example, the following message was classified as negative with the values 1 and -4:

“Horror at 95mph: Luger killed after smashing into concrete pillar at Vancouver Olympics [url]”

Finally, all messages that received polarity values of 1 and -1 were considered to be neutral. The next section will present the results of the sentiment analysis and describe its implications.

## 4. Results

### 4.1. Positive versus negative tweets

Of the total 46,097 tweets, 37% (17,218) were neutral, 15% (7,064) were strongly positive, 5% (2,344) were strongly negative, and only 298 tweets had both sentiments, according to SentiStrength. There were 3 times as many positive messages as negative. The results were very similar even after we removed 2,361 duplicates (messages that were retweeted, from the original sample of 46,097). There were 6,673 positive and 2,240 negative tweets, which yielded the 3 to 1 ratio.

This suggests that people tweeting about the Olympics were generally excited about the event. This is well demonstrated in Figure 1, which shows that there were more messages that mentioned the word “happy” per day (Figure 1.a) than there were messages that mentioned the word “sad” (Figure 1.b). The majority of “sad” messages were posted at the

beginning of the Olympics (mostly about the tragic death of a 21-year-old Georgian luge competitor Nodar Kumaritashvili) and at the end, when people were generally sad that the Olympics were over.

It should be noted that for this event the majority of messages with a strong sentiment were positive, but for other events and circumstances the balance of positive versus negative messages may be different. For example, Diakopoulos and Shamma analyzed over 3,000 tweets posted during the live presidential TV debate in 2008, and they found that the majority of tweets during the debate (41.7%) were negative, while 25.1% were positive, 6.8% were mixed (messages with both positive and negative components), and the remaining 26.4% were tagged as other (contained non-evaluative statements or questions) [4]. However, in the Jansen et. al.'s 2009 study of 150,000 tweets related to brand image building and company-customer relationships on Twitter, the researchers found that of the tweets that did express sentiment, over 52% expressed positive sentiment, while approximately 33% expressed negative sentiment [8]. Our future work in this direction will include the analysis of messages on other topics and contexts to determine a set of parameters that can be used to predict the predominance of one sentiment over another in Twitter messages.

#### 4.2. Positive versus negative REtweets

The results from the previous section may be expected for a global celebratory event such as the 2010 Winter Olympics; however, what is really interesting is that, looking only at the retweeted (forwarded) messages, we found that there are almost 2,5 times as many positive retweets as compared to negative retweets. Out of 2,031 that were retweeted, 251 messages were deemed to be positive, 98 negative, 992 neutral, and only 7 tweets had both sentiments. (The remaining messages were deemed to be unreliable for sentiment detection due to the sentiment values smaller than 3.)

To ensure that this discovery is not due to the fact that there were simply more positive messages than negative messages in our dataset, we also calculated the average number of times when each type of message (positive, negative and neutral) was retweeted (See Table 1). On average, positive messages were retweeted 6.6 times which is almost 3 times higher than either negative or neutral messages. In addition,

due to a number of outliers identified by SPSS Statistics, we also decided to look at the median number, which is less influenced by extreme values than the mean. The median for positive tweets was 2, and for negative and neutral messages it was only 1. So, even after accounting for outliers, positive tweets were on average retweeted twice as often versus only once for negative or neutral tweets.

However, it is an accepted fact that if a user has a lot of followers, his or her tweets will be more likely to be forwarded than those whose authors have fewer followers [16]. We see this in our data as well. For Twitter users whose messages were retweeted at least once, their median number of followers was 520. This is significantly higher than the average number of followers for the typical Twitter user, estimated anywhere between 27 to 85 followers according to the earlier studies [12, 13]. This means that regardless of the emotional polarity of the posts, a user with more than the average number of followers will likely get more retweets as well. However, within this group of users with a higher than average number of followers, we did not find any statistically significant relationship between the number of followers and the number of retweets based on the regression analysis (See the scatterplot in Figure 2). This suggests that based on our sample, the level with which Twitter messages are retweeted for users with more than the average number of followers is also influenced by factors other than just the number of followers. And based on the observation above regarding the polarity of retweets, we believe that the polarity of each message is one such factor. Specifically, it seems that positive messages are more contagious than negative. This does not necessarily prove that positive messages affect people more than negative, but that people are more likely to share good news than bad news during a sporting event.

Our conclusion is different from the one reached by Heath [11], where the author found that bad news stories are more likely to be mentioned and passed on to acquaintances. (Since Heath's study focused on general news, not just sport-related news, our comparison with this study is very tentative.) This difference might be accounted for by the fact that Twitter messages are transmitted within a virtual public environment where posted messages are accessible by millions of potential followers. Furthermore, this difference in message forwarding behavior might be the result of the fact that, as noted by Boyd, Golder and

Lotan, there are no set “rules” when it comes to retweeting, and inconsistent syntax creates issues surrounding authorship, attribution and conversational fidelity [2]. This in turn may give people pause from readily retweeting bad news as he or she may fear being the bearer of bad news. Another possible explanation has to do with the nature and purpose of the Twitter service. The majority of Twitter users joined and use the service with social motives in mind (e.g., “have fun”, “keep in touch with friends”) rather than for information motives (e.g., “share information”, “learn interesting things”) [9], with only about 4% of tweets being news stories (pearanalytics.com). So, Twitter users may be less interested in reading and sharing public news stories (whether positive or negative) with their followers as opposed to using it more as a social medium to share thoughts and news of a more personal nature and simply to keep in touch with like-minded people.

#### 4.3. Polarity and user’s network position

The next question that we tried to answer is whether or not a user's position in the Twitter network correlates with that person's tendency to post positive or negative messages. More specifically, we wanted to know whether or not people who tend to post positive messages also tend to be more central in the social network than those who tend to post negative messages. To answer this question, we used the Twitter API to find out how many followers each user in our dataset had and how many people they themselves followed. (For simplicity, we will refer to the latter number as “sources”.) In total, we collected posting information on 34,502 users who each posted at least one tweet. From this initial set of users, we excluded 12,626 “celebrity”-type users, defined as users with more than 5,000 followers, and/or “organization” and “spammer”-type users who posted more than 5,000 messages. This was done to exclude extreme cases which could potentially skew the results in one direction or the other. Next, for each user we determined his/her tendency to post positive or negative tweets. This was calculated based on the strongest sentiment of the available tweets in our dataset. For example, if a user posted three tweets with the corresponding polarity values of -2, 1, and 5, then the strongest sentiment will be 5 (positive). At the end of this filtering process, we ended up with 5,151 users whose sentiment was either positive or negative. Next we used t-test to compare the means of the number of followers and the means of the number of sources for

two groups of users, “positive” and “negative”. We found that on average, positive users had 35 more followers and 36 more sources than negative users. This seems to suggest that more positive users are also more central. But, admittedly, while these results are statistically significant, the difference is not large enough to give a definitive answer to confirm whether or not positive users are more central.

Regardless of how the above question gets answered with a larger sample, based upon our exploration of the top 100 negative users in our current dataset and their tweets, it is clear that some users have been able to attract a large number of followers despite the fact that their messages tended to be negative overall (e.g., an account like “omgihatethat”). This is somewhat different from what Fowler and Christakis observed in their study of face-to-face interactions. We attribute this to the fact that the internet in general and Twitter in particular is much more conducive to homophily than the physical world; it is infinitely easier to find a large group of like minded people who share your world views on the internet than in real life. But as the first part of this study showed, negative messages initiated by such communities are less likely to spread beyond their niche network on Twitter than positive messages.

There are likely many other factors that influence who follows whom and who retweets whose messages on Twitter, including the content the user posts, his/her social status, how active the user is, etc. For example, in our dataset of 5,151 users, we found that negative users were more prolific than positive users (counting from when they first joined Twitter); on average, posting 253 more messages than positive users, based on t-test (statistically significant). This may be because negative users are more passionate about their strongly held views and are more motivated to express their frustration, as indicated by their higher posting behavior. To determine to what extent the polarity of messages may influence a user’s network position, we will need to control for many of these factors. In our follow-up study, we will increase our sample size of Twitter users and will collect messages over a longer period, possibly 3 months. This larger sample dataset will allow us to conduct other types of statistical testing such as trying to determine whether there is causality between the tendency to post negative or positive messages and the user’s position within the network. More specifically, we will seek to determine whether people at the core of their local networks are

more likely to post positive messages, or if it is the other way around; are people who tend to post positive messages a more likely to be at the core of a network? To supplement our automated analysis, we are also planning to survey users in our sample about their perceived happiness and find out reasons why they choose to follow or not follow particular Twitter users.

## 5. Conclusions

We are only beginning to scratch the surface and uncover some of the specific mechanisms of how emotions can spread in online communities; however, this study is an important first step in this line of research. The primary aim of this initial research was to design and test methodology for studying how positive and negative emotions spread within social networks on Twitter. Starting with a large sample of over 46,000 messages, we determined that in the context of conversations about the 2010 Winter Olympics, there were more positive messages than negative, and that positive messages were 3 times more likely to be forwarded than negative messages. However, we were not able to confirm with a reliable degree of certainty that the emotional context of messages is directly related to the user's position in the Twitter network. It is likely that there are other factors involved as well. For example, we found that negative users were more prolific posters than positive users, suggesting a possible correlation with their more argumentative and passionate nature.

The results shed light on how the emotional tone and content of online messages may influence users' online interactions and the formation of social connections on the internet. The study also opened many doors for further research into questions like, "Do people cluster based on the average polarity of the messages that they post?" and "Is the contagion of messages dependent on the author being an individual or an organization?"

From a practical perspective, the results of this study point to the fact that if you want your messages to be forwarded and to reach more people, you need to make sure that both the tone and content of your online messages are positive overall. The results of this study show that even subtle nuances in the emotional content of a message can have a major impact on the receiver, and on the degree to which a message will be retweeted. Thus, even in the age of Twitter, the old adage about attracting more bees with honey than

vinegar still holds true.

## Acknowledgments

This work was partially supported by the Social Sciences and Humanities Research Council (SSHRC) grant. We would like to thank Sreejata Chatterjee for her assistance with data collection as well as Geoffrey Allen and Kathleen Staves for their help with the manual evaluation of Twitter messages. Also we would like to thank anonymous reviewers for providing very helpful comments.

## 6. References

- [1] J. Berger and K. L. Milkman, "Social transmission and viral culture", Wharton School Working Paper, 2010.
- [2] d. boyd, S. Golder and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter", HICSS-43. IEEE: Kauai, HI, 2010.
- [3] S. Crockett, "The art of feelings", *New Statesman*, 135(4800), 2006, pp. 20-21.
- [4] N. A. Diakopoulos and D. A. Shamma, "Characterizing debate performance via aggregated Twitter sentiment", Conference on Human Factors in Computing Systems, ACM, Atlanta Georgia, 2010.
- [5] P. S. Dodds and C. M. Danforth, "Measuring happiness of large-scale written expression: Songs, blogs, and presidents", *Journal of happiness studies*, 2009.
- [6] J. Easter, "Happiness and the internet", Master Thesis, Design Interactions, Royal College of Art, 2007.
- [7] J. H. Fowler and N. A. Christakis, "Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study", *British Medical Journal*, 2008, 337: A2338.
- [8] B. J. Jansen, M. Zhang, K. Sobel and A. Chowdury, "Twitter power: Tweets as electronic word of mouth", *Journal of the American Society for Information Science and Technology*, 60(11), 2009, pp. 2169-2188.
- [9] P. Johnson and S. Yang, "Uses and gratifications of Twitter: An examination of user motives and satisfaction of Twitter use", Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication, Boston, MA, August 5, 2009.
- [10] J. Y. Han, B. R. Shaw, R. P. Hawkins, S. Pingree, F. McTavish and D. H. Gustafson, "Expressing positive emotions within online support groups by women with breast cancer", *Journal of Health Psychology*, 13(8), 2008, pp. 1102-1107

[11] C. Heath, “Do people prefer to pass along good or bad news? Valence and relevance of news as predictors of transmission propensity”, *Organizational Behavior and Human Decision Processes*, 68(2), 1996, pp. 79-94.

[12] B. A. Huberman, D. M. Romero, and F. Wu., “Social networks that matter: Twitter under the microscope”, *First Monday*, 14(1-5), 2009.

[13] R. J. Moore, “New Data on Twitter’s Users and Engagement”, *RJMetrics report*, January 26, 2010, available at <http://themetricssystem.rjmetrics.com/2010/01/26/new-data-on-twiters-users-and-engagement>

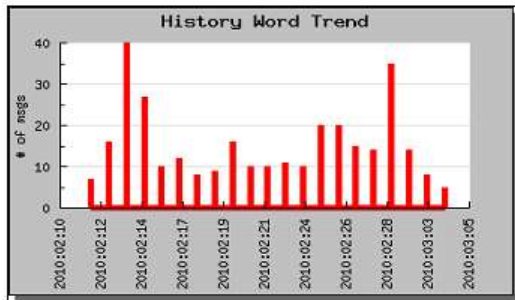
[14] G. Paltoglou, S. Gobron, M. Skowron, M. Thelwall and D. Thalmann, “Sentiment analysis of informal textual communication in cyberspace”, In the proceedings of *ENGAGE 2010*, Springer LNCS State-of-the-Art Survey, Zermatt, Switzerland, Sept.13 – 15, 2010, p.13-25.

[15] B. Pang and L. Lee, “Opinion mining and sentiment analysis”, *Foundations and Trends in Information Retrieval*, 1(1-2), 2008, pp. 1-135.

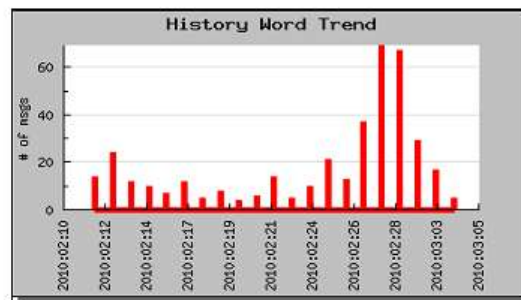
[16] B. Suh, L. Hong, P. L. Pirolli, E. H. Chi, “Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network”, *Second IEEE International Conference on Social Computing (SocialCom)*, August 20 - 22, 2010, Minneapolis, MN.

[17] S. Sum, M. Mathews, I. Hughes, and A. Campbell, “Participation of older adults in cyberspace: social inclusion and wellbeing”, Paper presented at the *4th National Conference on Depression in the Elderly; Successful Aging: Countering Depression in Old Age*, Sydney, Australia, 2007.

[18] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, “Sentiment strength detection in short informal text”, *Journal of the American Society for Information Science and Technology*, In press.



a

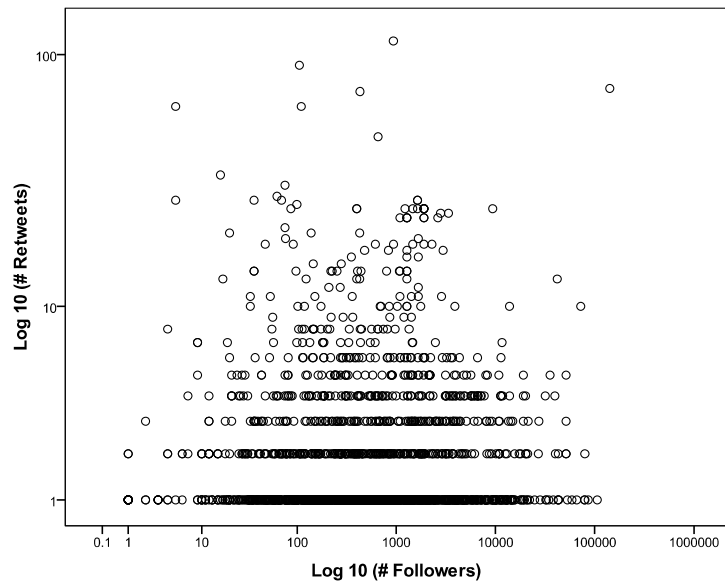


b

**Figure 1. Frequency occurrence distribution of messages per day mentioning the word - “happy” (a) and “sad” (b)**

Polarity	#Tweets	Maximum	Mean	Std. Error of Mean	Median	Std. Deviation
Negative	98	19	2.22	0.33	1	3.22
Neutral	992	113	2.63	0.19	1	5.87
Positive	251	101	6.6	0.67	2	10.54

**Table 1. Descriptive Statistics for Retweets**



**Figure 2. The number of retweets versus the number of followers for users with less than 500,000 followers**