Spring 3-27-2013

Emotions in the Twitterverse and Implications for User Interface Design

Anatoliy Gruzd
Dalhousie University, gruzd@dal.ca

Follow this and additional works at: https://aisel.aisnet.org/thci

Emotions in the Twitterverse and Implications for User Interface Design

Anatoliy Gruzd
Dalhousie University
gruzd@dal.ca

Abstract

This study explores the implications of how user interface elements affect the types of messages that are produced as well as the likelihood that, and extent to which, those messages are spread within an online social system such as Twitter.com, a popular online service for sharing short messages. The current paper explores these issues by studying the dissemination patterns of emotional-type messages among Twitter users through automated techniques, coupled with observations from a survey of Twitter users about their willingness to produce or forward messages containing different types of emotional tone. The results show that Twitter users post more positive messages (tweets) than negative, and that positive tweets are 3 times more likely to be forwarded than negative tweets. The findings also suggest that the Twitter user interface may be partially responsible for this (i.e., the interface reduces the likelihood that negative messages will be posted or retweeted). To enable a wider range of discourse on Twitter and to reduce the need for Twitter users to self-censor their tweets, the paper concludes with a potential design solution that will give Twitter users more control over who will receive their tweets, and outlines a future study to evaluate such an interface.

Keywords: Online social networks, user interface, networking systems, personal applications, multi-method

AIS Transactions on Human-Computer Interaction

INTRODUCTION

In the world of business, Starbucks and United Airlines have little in common. However, they share at least one thing; they both are enthusiastically embracing the use of social networking sites such as Facebook and Twitter to stay in touch with their customers, promote new products and services, and create online communities around their brands. This is part of a growing trend among many North American companies to experiment with new ways to reach their customers via the social web (e.g., McLaughlin and Davenport, 2010; Poynter, 2008; Pratt, 2009). As more people join 'social' sites, companies are quickly incorporating these sites into their marketing and customer relationship management plans.

Social networking technology has made it easier for companies to connect to their current and potential customers, and at the same time social networking sites give consumers a stronger say over how companies and other organizations should operate and behave in our society. Online users can find like-minded individuals on the web and share their positive or negative experiences about an organization or an event. Armed with just a connected smart phone and a social networking account, any consumer can potentially cause millions of dollars' worth of damage to a large corporation in a matter of minutes or even seconds. As Greenberg (2010) puts it, in social media "the customer is now in command of the business ecosystem" (p. 411). This is primarily due to the ability of electronic word of mouth (eWoM) to spread online messages instantaneously and to reach a lot more people with ease (Mason, 2008; Trusov et al., 2009). eWoM can be a very effective marketing technique for companies, but once it leaves the company, they have little to no control over if and how the campaign will spread (Jansen et al., 2009). Thus it is especially important for companies and other organizations to maintain their online presence and monitor what people are saying about them online, counter negative language, and provide support and feedback to their customers (Fluss and Rogers, 2010).

Although the Internet abounds with advice from social media 'gurus' and 'experts' on eWoM, it is still unclear from a research perspective why some messages get more attention and are forwarded, and some are not. This paper focuses on one possible aspect that may be responsible for this – the emotional polarity of a message. One hypothesis is that positive online messages tend to be forwarded by more people. Research in this area is inspired by Fowler and Christakis (2008) who found that happiness is contagious in a social network of locally bounded individuals. In addition, the researchers determined that the place occupied by the participants within their social network played a significant part in their level of happiness: "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).

However, at this point 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 this question, this paper investigates posting behavior of users of Twitter.com, a popular online service for sharing short messages called "tweets." Twitter was chosen because of the instantaneous broadcast ability of the messages and the built-in interconnectivity of its membership base. Of particular interest in this research is determining whether the interface design of online social networking systems can influence the way emotions spread among networked members and if so, how. The broad contribution of this research is a better understanding of how HCI can be enhanced through the use of design science research; specifically, based on theories and observations of how the emotional polarity of a message may influence the likelihood that a message will be posted or forwarded (or 'retweeted', using the Twitter slang). This paper attempts to address the following research questions:

- Do Twitter users tend to post primarily positive or negative messages?
- Are positive messages more likely to be forwarded than negative messages?
- Does a user's position within the network correlate with that user's tendency to post positive or negative messages?
- Does the Twitter user interface influence whether certain types of messages are posted or retweeted?

One of the aims of this research is to identify and document patterns in the spread of emotions on the Internet (specifically on Twitter) and to compare these patterns against those found by Fowler and Christakis (2008). This is achieved by using two different methods: automated sentiment analysis and an online survey. First, by applying sentiment analysis to a large sample of Twitter messages, we answer general questions regarding possible links between the emotional polarity of a message and its likelihood of being tweeted or retweeted on Twitter. Second, by conducting an online survey of Twitter users, we attempt to validate and explain some of the results from the automated sentiment analysis and to uncover whether there are any links between the Twitter user interface and the types of tweets that are being posted or forwarded on Twitter. We then conclude with a proposed design change to the Twitter interface that, if implemented, will allow Twitter users more control over who will get their tweets and remove their need to self-censor their tweets.

BACKGROUND

Spread of Emotions in Online Communities

In looking at why people join online communities, research shows that one of the strongest motivators is the positive emotions that these communities foster. A number of prior studies (discussed below) have shown that online communities can increase the levels of happiness in participants. For example, Sum et al. (2007) found that Internet use increases the well-being of older adults by decreasing their isolation. In their study, Sum et al. equated happiness with well-being and feeling connected. This concept of happiness is also reflected in a study by Han et al. (2008), in which women with breast cancer experienced an improved sense of well-being by participating in an online support group. However, contrary to the studies described above, Chou and Lim (2010) found little to no correlation between the use of online social networking sites and individual well-being in a different context. Their results were based on a survey of 66 college students, most of whom were Facebook users. The researchers did acknowledge that online social networking sites "provide conveniences and advantages to contact one's friends" and stated that their future research goal is to "find out whether the well-being of one's online friends is related [to]one's wellbeing" (p.202). This last point is also the subject of a number of recent studies discussed below.

Christakis and Fowler (2008), for example, studied 1,700 Facebook profiles and determined that people with smiling pictures on Facebook are more likely to have online friends who are also smiling in their profile pictures. Zafarani et al. (2010) investigated whether emotions propagate among bloggers in the LiveJournal social network. Based on a sample of 16,444 users, the researchers observed that some emotion propagation does take place in this community of bloggers. Thelwall (2010) conducted sentiment analysis of over 2 million MySpace public comments and found that Myspace users are likely to be friends with people with similar emotional levels. Finally, in a more recent study, Bollen et al. (2011) came to similar conclusions. The researchers analyzed a large sample of 129 million tweets produced by 102,009 users during a 6 month period. Using sentiment analysis, the researchers calculated an individual happiness level, or what they called the Subjective Well-Being (SWB) measure, for each user in the sample. They found that (1) the Twitter population in their sample is moderately happy and (2) users with low SWB tend to connect to other users with low SWB while users with high SWB connect to other users with high SWB. The researchers concluded that "Twitter users are preferentially connected to those with whom they share the same level of general happiness or SWB" (p.238).

These previous studies are highly relevant for our current line of research. Although these studies used different methodologies and relied on different datasets, they independently observed that an online user with a certain emotional state tends to be friends with other users who share similar emotional states. However, these studies stopped short of explaining the main factors that might cause such results. For example, it is still unclear whether this is because people tend to seek out other people with a similar emotional disposition or whether this is in fact because emotions spread online like a biological contagion allowing happy people to pass on their happiness. If the former is true, two people in different emotional states who meet online will more likely stop interacting with each other due to the incompatibility, rather than converge to the same emotional state over time. If the latter is true, then people can start in different emotional states and converge to the same emotional state over time. Although this question still remains largely unanswered, the current paper will shed some additional light on this issue. In particular, the paper seeks to validate the results observed by the previous studies, but focuses on a different context in which Internet users are discussing a global event that is expected to be capable of producing both negative and positive reactions. Furthermore, the paper explores other possible factors that might be in play, such as how the user interface design of the social system might help or hinder the spread of positive or negative messages online.

User Interface Design and Emotions

Also relevant to this paper is research in the field of Human-Computer Interaction (HCI). Researchers in HCI have dealt with users' emotions in one of two distinct ways. The first direction of research concerns how users' emotional states can influence their perceptions of a system. Examples of related work in this direction include studies and theoretical models that seek to explain how users' emotions can influence their experience with a system and the likelihood of its adoption (Loiacono and Djamasbi, 2010; Venkatesh, 2000). The other direction of research explores how a user interface can influence or account for users' emotional states to achieve an end goal. This line of research includes work by Marreiros et al. (2010a,b) and Thagard and Kroon (2006) which has shown that the emotional awareness of users of a group decision making system can help the group to achieve agreements more rapidly. There are also a number of recent commercial projects that have relied heavily on deciphering and understanding users' emotional states of mind. For instance, there is a growing interest in tone management and tone filtering applications to manage social media campaigns. One such application called ToneCheck (http://tonecheck.com) has been released by Lymbix. ToneCheck is an e-mail plug-in that automatically identifies sentences with words or phrases that may convey unintended emotion and then allows the user an opportunity to revise the message. Other applications in this area include a BlackBerry app called MoodMe, which allows BlackBerry users to define color and

pattern that they will see on their Blackberry when receiving a text message containing a certain mood, and personal therapeutic phone applications such as iPhone Moody Me, which helps users record their emotional states and take pictures of what makes them happy so that they can relive those moments by viewing pictures captured earlier.

Although relevant to the general theme of the paper, the previous work in this area of HCI has not considered the implications of how user interface design features affect the types of messages that are produced as well as the likelihood that, and extent to which, those messages are spread within an online social system such as Twitter. The current paper explores these issues in more detail by studying the dissemination patterns of emotional-type messages among Twitter users through automated techniques coupled with information from a survey of Twitter users about their willingness to produce or forward messages containing different types of emotional tone.

Measuring Emotions

One of the key methodological questions to complete this line of research is how to measure emotions and their spread among users of an online social networking site like Twitter. Previous studies on happiness and subjective well-being have employed various methods to detect and measure emotions of online members. One of the most common approaches is to use a survey instrument. For instance, Sum et al. (2007) relied on a survey to measure variables such as intensity of Internet use, life satisfaction, and community scale. By measuring these and other related variables, the researchers were able to identify if Internet use was increasing or decreasing the well-being of the participants. Han et al. (2008) used a similar method in the study of online support groups for women with breast cancer. By evaluating the participants' levels of well-being before and after their participation in an online support group, the researcher determined that the expression of emotions within the group led to an increase in feelings of well-being.

The second approach commonly used to measure happiness and other sentiments online is to use human coders to evaluate a specific piece of text. For example, Berger and Milkman (2010), in their study of the most-emailed New York Times articles, utilized human coders who were asked to rate each sample story according to its type (i.e., practical, inspiring, or surprising) on a 5-point scale. By looking at the average scores and the frequency with which the stories were emailed, the researchers determined which types of stories were most likely to be emailed and found that "affect-laden" and positive articles were the most e-mailed stories.

The third approach, commonly referred to as opinion or sentiment analysis, involves the use of a computer program to analyze text in order to determine its polarity. This automated method tends to be more scalable and more objective as compared to methods involving survey or human coders. As a result, it is usually used when large datasets are involved. For example, Harris and Kamvar (2010) developed a web system called http://wefeelfine.org/ that searches newly-posted blog posts for presence of phrases which start with "I feel" or "I am feeling." When the system finds such phrases, it tries to identify the emotion being expressed in the blog post (e.g., sad, happy, depressed, etc.). Another web project that used a similar approach is the Facebook Gross National Happiness application (Kramer, 2010). This application collects short status updates posted by Facebook users, and then counts how many positive versus negative words are mentioned on a particular day. According to the data collected from this application, the happiest days in the United States corresponded to major holidays such as Thanksgiving and Christmas. To improve the quality of sentiment analysis, a number of studies have also relied on emoticons to classify messagesasbeing positive, negative or neutral (e.g., Read, 2005; Vogel and Janssen, 2009).

Because Twitter has millions of members and because of the dynamic nature of their conversations, automated sentiment analysis was chosen as the primary method for the research in this paper. For an extensive review of different sentiment analysis techniques, see Pang and Lee (2008) or Paltoglou et al. (2010), and Thelwall et al. (2011) specifically for the sentiment analysis of informal messages on the Internet. To validate and interpret the results of the automated sentiment analysis, an online survey was chosen as a complementary research method for this project. The next section will describe the methods used in this paper in more detail.

In sum, this paper attempts to validate some of the observations made in the previous literature regarding how sentiments may spread in an online social network with two main differences. First, instead of looking at messages based on a random sample of users, we focus on (1) the propagation of sentiments around a single global event to reduce the local events' biases and (2) the use of an online survey to measure users' perceptions. We then use the results from the first part of the study to explore issues around user interface design and its potential influence on the spread of certain types of messages, namely those that might convey and contain strong emotional tones.

METHOD

Automated Sentiment Analysis

Twitter was the logical choice for this research as most of the tweets are publicly available and are easily retrieved using the Twitter API. Twitter was also appealing because of how quickly messages spread from user to user. The fast-paced nature of this medium allows the real-time spread of sentiments to be observed through messages being forwarded (or "retweeted"). It is impossible to collect completely unbiased tweets that are not influenced by some local or global events on Twitter. To limit some possible biases, we collected messages about a single popular event. Furthermore, collecting messages on the same topic allows for the comparison of contradicting opinions being expressed about the same event to determine which view is forwarded more often. With these considerations in mind, a data sample was chosen to include only those messages that mentioned the 2010 Winter Olympics. This subset of tweets was chosen because this event was very popular and garnered a lot of attention in the popular press and online media, including Twitter; there was no shortage of tweets that mentioned the Olympics. Because of the competitive nature of the event, we expected that a large proportion of the messages would have strong emotional content and 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. 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 the Twitter search API every hour to retrieve the 100 most recent tweets that mentioned the word "Olympics." During this period, 46,097 tweets were collected.

Automated sentiment analysis techniques were used to identify the polarity of each Twitter message (neutral, negative, positive, or both). Because the goal of this research is neither to develop a new sentiment analysis system nor to improve an existing one, an "off-the-shelf" system called SentiStrength (v2.1) was used. The system was developed by a team of researchers from the University of Wolverhampton in the UK, and is available at http://sentistrength.wlv.ac.uk/. Although there are 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 http://myspace.com/, a formerly popular social networking site, the accuracy of predicting positive and negative emotions was somewhat similar to that of other systems (72.8% and 60.6% for negative and positive emotions, respectively, based on a strength scale of 1–5).As compared to other available methods, SentiStrength showed the highest correlation with human coders (Thelwall et al.,2011).

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,

> *I love, LOVE watching the Olympics...Opening ceremonies now...a sad start though...the killed of an Olympic brother.*

To differentiate between messages that include strong sentiments versus those that are subtle, only messages with a polarity greater than 2 or less than -2 were used for analysis. 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 SentiStrength values of 3 and -1:

> *RT Awesome! @CBCAlerts: Corey Perry scored twice as Team Canada thumped Russia 7-3 in men's hockey quarter-final at Vancouver Olympic[s]*

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:

> *Sad that the Olympics are over. There's just something about international competition.*

Finally, all messages that received polarity values of 1 and -1 were considered to be neutral.

Online Survey

To validate and explain some of the results from the automated sentiment analysis, the second part of the study consisted of an online survey open to Twitter users from the United States. The survey included a total of 14 questions, designed to gauge Twitter users' demographics, how aware they are of their follower-base and whether that awareness influences the types of message that they post or retweet (e.g., Does having work colleagues in an individual's Twitter network cause people to self-censor?)

To administer the survey, we used an online system called Amazon Mechanical Turk (https://www.mturk.com). Mechanical Turk (MT) was created in 2005 as an online crowdsourcing platform to recruit Internet users who are willing to complete what Amazon euphemistically refers to as a "Human Intelligence Task" (HIT) for a nominal fee. Examples of common HITs include completing surveys, proof-editing texts, and transcribing audio files. After accessing the website, MT workers are presented with a list of tasks from job-requesters, such as researchers, sorted by topic, including details of monetary remunerations (usually in the range from $0.10 to $0.50) and the time commitment required to complete a particular task. It has been estimated that approximately 80% of MT workers are from the United States (Paolacci et al., 2010) This factor makes Mechanical Turk an ideal platform for this study, as it has also been determined that most Twitter users currently reside in the United States (Takhteyev et al., 2012). Recruiting survey participants through Mechanical Turk was fast, efficient and economical. Each survey participant on MT is assigned a unique identification code. This code ensures the authenticity of each participant, while at the same time guarantees complete anonymity for them. The authenticity and integrity of each participant on Mechanical Turk was further enhanced by the inclusion of an approval rating for each worker based on how well they performed on previous tasks on MT as rated by other job-requesters. To ensure high quality responses, only MT workers with an approval rating of 95% or higher were allowed to take part in the survey. Other inclusion criteria were residency in the United States and having an active Twitter account. Each participant who completed the survey received $0.30.

In total, 100 participants completed the survey over a three-day period (July 16-18, 2011). The average completion time for the survey was three minutes. The surveyed group was evenly distributed between the sexes (48women, 51men, and one undeclared). The mean age of the respondents was 28 (SD = 9 years), with the majority (64) under 30. This age distribution is reflective of the general Twitter population (Smith, 2011). About half of the survey participants (49) reported using Twitter daily or weekly,43 reported being monthly users, and 8 reported using Twitter more than once daily.

RESULTS

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. (The remaining messages were deemed to be unreliable for sentiment detection due to sentiment values smaller than 3.)There were three times as many positive messages as negative. The results were very similar even after the removal of 2,361 duplicates (messages from the original sample of 46,097that were retweeted). After making this adjustment, 6,673 positive and 2,240 negative unique tweets remained, which yielded the 3 to 1 ratio. This suggests that most people tweeting about the Olympics were generally excited about the event. 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) and at the end, when people were generally sad that the Olympics were about to end.

The dominance of positive and neutral over negative types of messages on Twitter was also confirmed by the survey respondents. Out of 100 participants, 50 stated that they tweet about events that make them happy or excited, while only 18 respondents stated that they post messages that make them feel frustrated, angry or sad. Also when asked about the general tone of most of their tweets/retweets, the majority (71) said that it is "positive" and only 3 people reported that it is "negative." This can be explained by the fact that many Twitter users are conscious of their audience. In fact, about half of the respondents reported considering who will be able to view their messages before they post them, stating that the knowledge of their audience influences the tone and content of their messages. In the comment section for this question, many indicated that they wanted to "sound professional" and were sensitive to how their messages will be perceived by their family members, friends, co-workers, or clients. For instance, one person said "I won't tweet anything insensitive/inappropriate that my coworkers or parents might read"; another explained that "I often refrain from posting tweets with profanity, as I have several professional colleagues that follow me on Twitter"; and another reported being "consistently concerned about how my social media presence could possibly affect how I am viewed professionally." In addition to trying to keep their chatter professional, some respondents were generally concerned with maintaining a "happy" online self-image. For example, one person said "I try not to tweet negative things because I don't want people to think I am unhappy." Other people emphasized the public nature of Twitter and their worries about revealing private information, saying things like "The Internet is so

open you have to be careful what you [share] because anyone can read it." Finally, some explained that they attempt to avoid offending followers with their political or religious beliefs.

Positive versus Negative Retweets

Out of 2,031 messages that were retweeted, 251 messages were deemed to be positive, 98 negative, 992 neutral, and only 7 tweets had both sentiments. On average, positive messages were retweeted 6.6 times which is almost 3 times higher than either negative or neutral messages (see Table 1). Due to the number of outliers identified by IBM SPSS statistics software, we decided to analyze the median (which is less influenced by extreme values than the mean). The median number of retweets for positive tweets was 2; for negative and neutral messages the median was 1. So, even after accounting for outliers, positive tweets were retweeted about twice as often as negative or neutral tweets.

Table 1: Descriptive Statistics for Retweets

Polarity	#Tweets	Maximum # retweets	Mean	Std. Error of Mean	Median	Std. Deviation
Negative	98	19	2.22	0.33	1	3.22
Positive	251	101	6.60	0.67	2	10.54
Neutral	992	113	2.63	0.19	1	5.87

The survey respondents were split almost evenly when asked if they ever retweeted a negative post. Those who said 'yes' (52 respondents) said they retweeted a negative tweet because they found it interesting. Some said that they would retweet a negative message only if "it is about something like a fire, or the death of a loved one." As for the other half who did not retweet a negative tweet (48 respondents), as noted in the previous section, most indicated that they want to stay positive; for example, "I try to stay away from being negative as I believe it brings others down, and that's not fun."

Polarity and User's Network Position

Do 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, the Twitter API was used to find out how many followers each user in the dataset had and how many people they themselves followed ('sources'). In total, posting information was collected on 34,502 users who each posted at least one tweet in the sample dataset. From this initial set of users, 12,626 "celebrity"-type users were excluded. "Celebrity"-type users were defined as users with more than 5,000 followers, and/or 'organizational'- and 'spammer'-type users who posted more than 5,000 messages. This was done to exclude extreme cases which could potentially skew the results.

For each user in the dataset, we determined his/her tendency to post positive or negative tweets. This was calculated based on the strongest sentiment of the available tweets in the dataset. For example, if a user posted three tweets with the corresponding polarity values of -2, 1, and 5, then the strongest sentiment would be 5 (positive). At the end of this filtering process, 5,151 users remained whose sentiment was either positive or negative. Next, we performed a 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." Results revealed that on average, positive users had 35 more followers and 36 more sources than negative users. But, while these results are statistically significant ($p<0.05$), this difference in reality is not large enough to definitively conclude that positive users are more central. Also, having more followers on Twitter does not necessarily mean that a person is more popular. For instance, consider a hypothetical user with 500 followers. This may seem like a lot, but to achieve this the user had to start following 1,500 other users; only 1/3 of them reciprocated by following the user back out of interest, courtesy, or simply automatically. Therefore, a better way to find popular users may be to use a TFS (Twitter Follower-Source) ratio, also known as TFF (Twitter Follower-Friend), which is the ratio of the number of followers to the number of sources. For example, there is a user called TheDodgerhater, who often posts non-flattering comments in general; in particular those messages typically target the Los Angeles Dodgers, a US baseball team. Here is an example of a post from this user about the Olympics from the sample dataset:

> *Every day I find something horrendous about #NBC and their #Olympics coverage. The worst so far was putting #USA v #Canada on MSNBC.*

Despite the tendency to post negative comments, this user has 971 Twitter followers. However, following 1,992 other Twitter users likely enabled TheDodgerhater to develop such a large following, as suggested by this user's low TFS ratio: 971/1,992 = 0.48. Based on this observation, it can be hypothesized that, similar to spammers, "negative" users may have a lower TFS than neutral or positive users. When the means of TFS ratios for positive ($M = 2.54$, $SD = 41.54$) versus negative ($M = 3.07$, $SD = 28.09$) users were compared, the result showed no statistically significant

difference, *p* = 0.68 (two-tailed).This suggests that at least in this sample, the positive or negative tone of messages did not determine whether a user had a low (below 1) TFS ratio on Twitter. Furthermore, based upon the exploration of 500 random users in the current dataset and their tweets (see Figure 1), it is clear that some users have been able to attract a large number of followers and achieve a TFS ratio higher than 1 despite the tendency of their messages to be negative overall (e.g., an account like "omgihatethat"). These are the users represented as red color nodes and who appear near the center of the network in Figure 1. The size of each node represents the number of followers. This is somewhat different from what Fowler and Christakis (2008) observed in their study of face-to-face interactions. This can be attributed 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. This is in line with the previous research on online social networks and emotions discussed in the Background section.
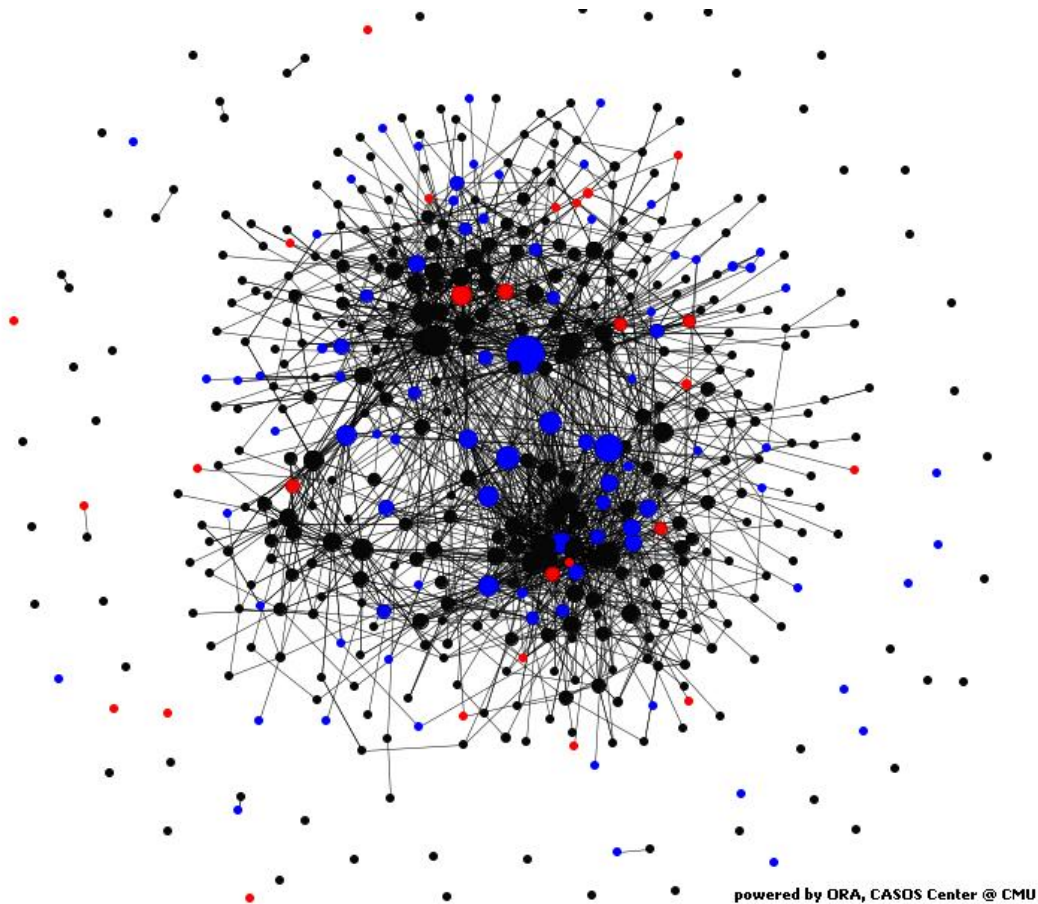
Note: blue nodes = positive users; red = negative; black= neutral; the size of the nodes represents the number of followers in this sample.

Figure 1: 500 Random Twitter Users who Tweeted about the 2010 Winter Olympics

The survey responses also confirmed some of the observations from the quantitative part of the study. Slightly more than half of the respondents reported that they do not think the tone of their posts or retweets influences the number of followers they have (56 respondents) nor the types of individuals who follow them (57 respondents). The remaining 44 and 43 respondents, respectively, either reported thinking that the tone of their posts or retweets affects the number or types of their followers, or that they did not know (see Figure 2). Those who endorsed a relationship between tone and number of followers said things like "people don't want to read negative tweets" and "positive tone increases followers"; one respondent even noticed that "I usually lose followers after posting really negative things." However, at least one respondent noted that "certain people look to either positive, confident or aggressive tones as inviting or powerful. Saying something radical can be influential like Fox News or Bill Maher do on television to gain viewers." This observation is also in line with the sample dataset which shows even "negative" Twitter users can be popular. The homophily effect has also been noted by some respondents, especially by those who endorsed a relationship between tone and types of followers, who left comments such as "people with similar emotions will follow me," "positive people will generally attract other positive people" and "random followers disappear quickly but like-minded followers stick around."

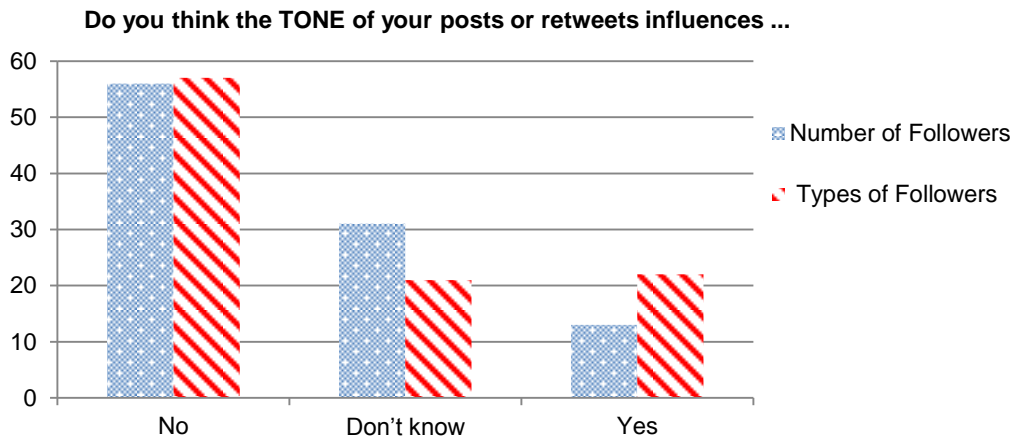**Do you think the TONE of your posts or retweets influences ...**

**Figure 2: Survey Responses: Exploring a Relationship between the Tone of the Messages
and (1) the Number of Followers and (2) the Types of Followers**

There are likely many other factors that influence who follows whom and who retweets whose messages on Twitter, including the content the user posts, social status, how active the user is, and even the user interface design of the social networking platform. Although each of these and other possible factors is worthy of further investigation, the factor that is most relevant to the readers of this journal (and that has not been well studied in the previous literature in this context)is the user interface design. Therefore, the next section will rely on the empirical results from this part of the study to determine if the Twitter user interface influences whether or not certain types of messages with clear and strong emotional tone are posted/retweeted.

DESIGN IMPLICATIONS

From a Human-Computer Interaction perspective, it is known that users' mood may affect how people interact with systems and that a system can be designed to perform mood management (Loiacono and Djamasbi, 2010). But can interface design elements of social networking systems influence a user's tendency to post negative or positive messages? They might. For example, in the case of Twitter, there are at least two system-wide interface elements that may be preventing some users from posting or retweeting messages that are deemed to be negative. One interface element relates to both posting original messages and retweeting somebody else's message, and the other element relates primarily to the retweeting function. Both are discussed below.

First, as noted in the previous section, due to the public nature of conversations on Twitter some users might be reluctant to openly criticize others or disseminate negative messages in a public forum. Furthermore, due to the asymmetric nature of Twitter connections (where one user can follow another, but the other user does not have to follow them back) the sense of community among Twitter users is not as strong as with mutual social ties in systems like Facebook. These observations, coupled with the fact that Twitter users often belong to many different social circles (see, for example, Gruzd et al., 2011), greatly increase users' uncertainty as to who might actually read their Twitter posts and how these might be perceived. This in turn may deter some Twitter users from posting more authentic or personal messages on Twitter, as reported by at least half of the survey participants. For example, if a user feels down one day due to some work-related problems and wants to share this with others, but suspects that among his hundreds of followers there may be a few co-workers, or even clients, then he will probably be more self-conscious about posting such thoughts on Twitter. One possible design consideration to address this concern is to allow users more control over who will receive their messages on Twitter. This will ensure that broadcasting will be more targeted and should reduce the often heard complaint of "I don't want to know what you had for breakfast!" Since Twitter already allows users to organize followers into thematic lists, Twitter can expand this feature to allow users to pick and choose which group(s) will get which messages. Figure 3 demonstrates how this feature might work. Once a user enters a message, but before clicking on the "Tweet" button, the user could select what group(s) should receive the current message by clicking on the corresponding check boxes. The default option would be to send a message to all of one's followers.
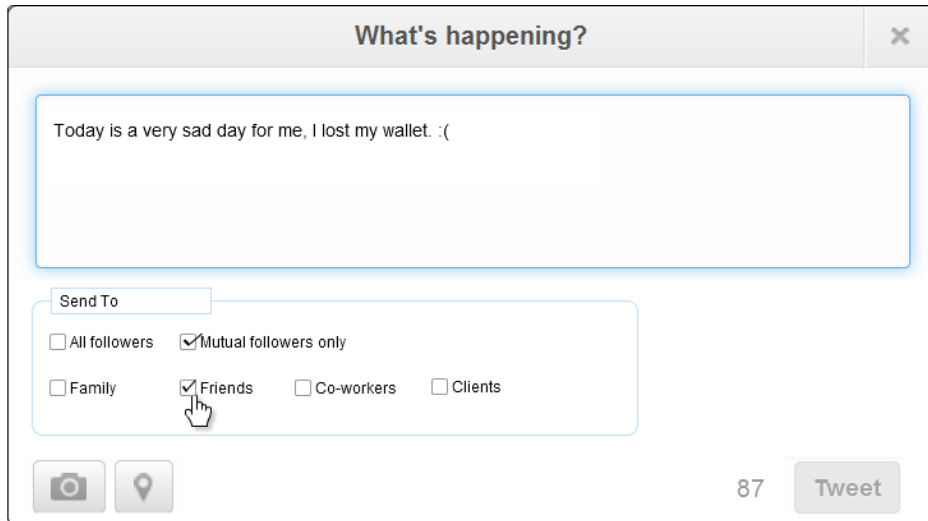
**Figure 3: A Revised Tweeting Pane with an Option to
Specify what Group(s) will Receive the Message**

One of the main inspirations for this design solution came from the survey participants. Specifically, when asked about their preferences as to whom they would direct a tweet if they could, their answers varied significantly depending on the type of message in a hypothetical scenario. Each participant was given four different scenarios based on different message sentiments: a message is (1) positive, (2) negative, (3) sad, or (4) personal in nature; and a multiple choice of five broad groups of potential recipients: (1) all followers, (2) colleagues, (3) family, (4) friends, or (5) strangers. A person could select more than one group that would receive the message. The answers to this question are summarized in Figure 4. According to the survey, Twitter users were most likely to choose to send positive messages to 'all followers' and 'colleagues', personal and sad messages to 'family' and 'friends', and 'negative' messages to 'strangers'. Interestingly, but not unexpectedly, 'colleagues' and 'strangers' were less likely to be selected as recipients of messages containing any emotional sentiments—whether positive or negative. This is because 'colleagues' and 'strangers' were only selected as recipients for all four types of messages 24 and 28 times, respectively; compared to the other three groups of potential recipients (all followers, family, or friends) which were targeted over 120 times each. This result suggests that if given the option to direct messages with strong emotional content to a specific group or groups of followers, a Twitter user would exercise that option; otherwise, most of the respondents would have chosen 'all followers' as an answer in the given four scenarios. In particular, the results suggest that Twitter users would be more comfortable sending personal or sad messages to family and friends, while colleagues and strangers would likely receive more neutral messages without any strong positive or negative sentiments. Such behavior has also been observed on Facebook and is dubbed the "Lowest Common Denominator Culture" by Hogan (2010), who explains this phenomenon as follows: "one might not be posting for one's parents (or children or students) on Facebook, but again, one is posting in light of the fact that these individuals may have access; these individuals define the lowest common denominator of what is normatively acceptable" (p. 383).
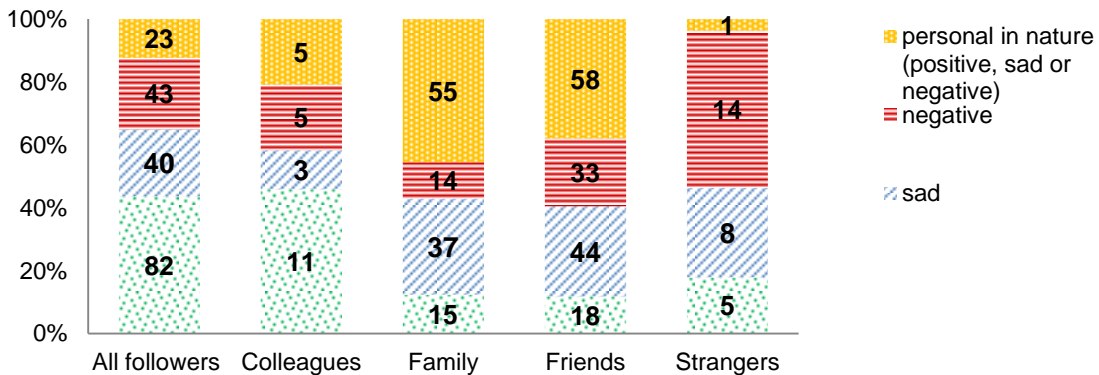
**Figure 4: Survey Responses to the Following Question: "If you could direct your tweet to
a specific group(s) within your network, who would they be?"**

Second, a retweeting convention of using "RT" to denote that a message is a retweet and referencing the original author using "@mention" found in the 2010 version of Twitter was somewhat confusing. This format made it

unnecessarily complicated for one's followers to determine the authorship of a message, especially when a message had been retweeted by multiple users. As noted by boyd et al. (2010), because there is no set of 'rules' when it comes to retweeting, inconsistent syntax creates additional issues surrounding authorship, attribution and conversational fidelity. This in turn may prevent some users from readily retweeting bad news or negative messages as they may fear that a retweet may be perceived as their own, making them the unwitting bearer of bad news. This observation is also supported by the comments left by the survey participants. For example, the following quote from one of the respondents sums it up very well:

> *I realize that if I re-tweet a message, my followers on Twitter will assume that I support the beliefs stated in the re-tweet. So I'm careful about the tone and message content so that I won't be negatively judged by my followers. Otherwise, they might not take me seriously when I need their support.*

Since 2010, when the Twitter dataset for this research was collected, Twitter has launched a "retweet" button. It serves to remove confusion about authorship by generating an attribution to the original sender. For example, Figure 5 shows a message posted by @OReillyMedia and retweeted by @timoreilly. Asshown, "timoreilly" is shown in much smaller font size next to the originator of the message – "OReillyMedia".
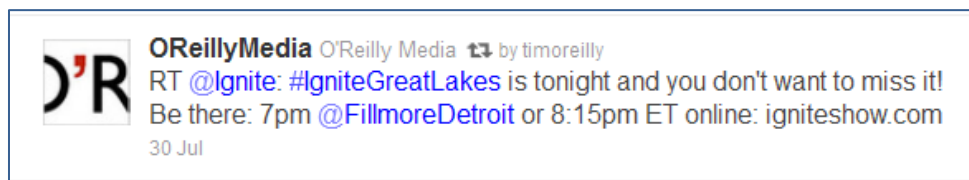
Figure 5: A Sample Tweet Originated by @OReillyMedia and Retweeted by @timoreilly

Unfortunately, this recent change to the interface is unlikely to influence users' retweeting behavior significantly. The user might still have some reservation in retweeting a negative tweet since retweets will appear in a user's timeline (archive of personal tweets). Furthermore, some users still rely on the original syntax of including "RT @"when retweeting others' messages. As a result, there still may be some uncertainty around the authorship of a tweet. For instance, after examining the content of the tweet in Figure 5, which was sent after the addition of the new retweet button, it is clear that the actual originator of this message was @Ignite because of the presence of "RT @Ignite." This means that @OReillyMedia used the original approach to retweet @Ignite's message, and @timoreilly simply pushed the new "retweet" button. These observations were also confirmed by the survey, which was conducted after Twitter changed its interface and added the retweet button. As discussed in the previous section, the survey showed that many were still concerned about coming across as a negative person if they were to retweet a negative post. As one of the respondents pointed out "The people I follow and who follow me are mainly business so I feel I should keep it all in good fun and professional. I try not to retweet anything that would make them think negatively of me."

To remedy these concerns about retweeting negative messages, an interface feature similar to the one proposed in Figure 4, which would allow users to send a retweet to a certain group or groups of people, could also be incorporated in the new retweet button. Although it is not the main focus of the study in this paper, such a design change could also make it easier to share more relevant messages (topic-wise) with different groups of followers. For example, when a person is attending a conference, the number of tweets that the user posts is likely to sky-rocket if the user posts a minute-by-minute commentary on somebody's presentation. This could be useful for conference attendees or people who are interested in this conference, but it may be perceived as noise by others. By utilizing the feature described above, the poster can direct messages only to people who would likely be interested in such commentary. However, to be effective and efficient, the number of pre-defined user-created groups that a Twitter user can create must be kept to a minimum.

The two observations about the Twitter interface made above broadly support Loiacono and Djamasbi's (2010) evidence that there is a connection between systems' design, users' mood, and how users interact with systems. This section also highlights a need for developers and usability experts to revisit and systematically account for mood detection methods and the increasingly social aspects of modern information and communication technologies when designing and evaluating systems.

CONCLUSION AND FUTURE WORK

This study and a few related studies mentioned in the Background section are only beginning to scratch the surface of uncovering the specific mechanisms of how emotions can spread in online communities and how systems design may influence the spread of certain types of messages. The primary aims of this research were: (1) to design and test methodology for studying how positive and negative emotions spread within social networks found in online social

networking sites such as Twitter and (2) to explore whether Twitter's user interface influences the types of messages (with a particular focus on messages containing strong emotional tone) that can spread in an online social system. Starting with a sample of over 46,000 Twitter messages about the 2010 Winter Olympics, the study determined that there were proportionally more positive messages than negative (Research Question 1), and that positive messages were 3 times more likely to be forwarded than negative messages (Research Question 2). We then confirmed this finding via a survey of 100 Twitter users designed to validate and explain the findings from the automated sentiment analysis.

Next, we tested another hypothesis, adopted from Fowler and Christakis' (2008) study about face-to-face interactions, that "positive" people are more connected (have more connections) in a social network. Translating this hypothesis into an online social system such as Twitter would mean that people who tend to post post positive messages should have more followers than those who tend to post negative messages. But, based on the results of the study, this is not the case(Research Question 3). The results suggest that the tendency of some users to post positive or negative messages is less likely to relate to the number of their followers, but more likely to relate to the types of their followers. This is likely due to the fact that most Twitter users have such a wide variety of people who follow them, from professional contacts to close friends to family members and even strangers whom they have never met. As a result, about half of the survey participants reported that they are actively censoring their posts to ensure that their tweets are appropriate for all of their followers. This suggests that, at least for this half of the survey participants, the Twitter interface influences whether certain types of messages are posted or retweeted (Research Question 4).To address this concern, we have proposed a potential design solution to the Twitter interface that will enable users to direct their messages to a pre-defined group or multiple groups of their followers such as 'friends' or 'colleagues'. The expectation is that incorporating such functionality into Twitter will help combat the "lowest common denominator culture" and allow people to express themselves online more freely. The result of the accompanying survey suggested that if given the option to direct messages with strong emotional content to a specific group or groups of followers, Twitter users would exercise that option.

The main limitation of this study is that it relies on a dataset containing Twitter messages about a very specific event – the 2010 Winter Olympics. For this event, the majority of messages were decidedly positive, but for other types of events the balance of positive versus negative messages may be different. For example, Diakopoulos and Shamma (2010) 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). Yet, Jansen et al.'s (2009) study of 150,000 tweets related to brand image building and company-customer relationships on Twitter found that of the tweets that did express sentiment, over 52% expressed positive sentiment, while approximately 33% expressed negative sentiment. Future work in this direction will include the analysis of messages on other topics and in different contexts to determine a set of parameters that can be used to predict the predominance of one sentiment over another in Twitter messages.

The second limitation of this work is that it only included messages with emotional content. An expansion of this work will include the analysis and classification of "neutral" types of messages with more of a focus on the types of information that they carry. The third limitation is that the survey part of this study relied on a non-random sample of 100 US-based Twitter users. While this sample was deemed sufficient to identify some common beliefs and considerations that Twitter users have when posting messages with strong positive or negative sentiments, and was useful in confirming some of the observations and interpretations made based on the sentiment analysis, a random and representative sample of users is needed to achieve stronger generalizability of the results as well as to discover some other possible issues with the current Twitter interface.

Finally, the study did not directly test the proposed interface design; instead it relied on the results of the sentiment analysis and a subsequent survey to identify potential issues with the current Twitter interface, explore the design space, and propose a possible solution. Future work will include a direct evaluation of the proposed solution; allowing users to send messages to certain groups of followers (e.g., friends, family, or colleagues). Initially, as part of a follow up study, the original plan was to build and evaluate a working client for managing a Twitter profile and posting messages with the built-in functionality to direct messages to a group of followers. However, soon after this study was concluded, Google released their new social networking site called Google+ that incorporates many features from both Twitter and Facebook. More importantly, Google+ has a similar functionality to the one proposed in this paper that allows its users to organize followers into groups, or what Google calls 'circles,' and direct messages to a specified circle. And since Google+ also enables Twitter-like, asymmetric relationships, where one user can follow another but the second user does not have to follow the first one back, Google+ is an ideal testing environment to continue this line of research.

More work needs to be done to develop design models and practices that will account for various nuances of social structures in online social networks and take into account how people interact in complex online social environments. The current paper looks into one such nuance: emotions, and how their spread in large socio-technical systems might be influenced by certain design features. As more and more people join social media websites, companies and other

organizations are beginning to see these new communication channels as an additional avenue to reach and engage their stakeholders. What makes these channels so attractive is their ability to propagate messages through so-called electronic Word of Mouth (eWoM). However, as confirmed by this research, not all messages have the same chance to go viral and be heard. This study also demonstrated that how you say something (tone) on social media is just as important as what you say (content).This research suggests that a positive social media campaign would have a better chance of reaching more people than a negative one, at least within the Twitterverse. The broader implication of this study is that the user interfaces of online social environments (and their frequent changes) may influence what people decide to share online, and may impact the social dynamics of an online group as well as their attitudes towards the underlying information system itself. From a more practical side, with a deeper understanding of online social structures and dynamics, it will be possible to develop new innovative applications. Some examples could include an application designed to help moderators of online communities detect "unhappy" clusters of people within their group, thus allowing those moderators to react accordingly to maintain a vibrant and welcoming online community. Alternatively, an email tone filtering tool could be used to help employees within organizations to stay on message. Tools like these would help companies and other organizations to manage their online presence, prepare messages that are more likely to reach the widest audience possible, and monitor what people are saying about their brands so that they can react swiftly and appropriately when negative sentiments are detected.

ACKNOWLEDGEMENTS

REFERENCES

Berger J. and K. L. Milkman (2010) "Social Transmission and Viral Culture," Wharton School Working Paper.

Bollen, J., B. Gonçalves, G. Ruan and H. Mao (2011). "Happiness Is Assortative in Online Social Networks," *Artificial Life* (17) 3, pp. 237-251.

boyd, D., S. Golder and G. Lotan (2010) "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," in *Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS),* pp. 1-10. HI, USA, Janurary 5-8, 2010,

Chou, A.Y. and B.B.L. Lim (2010) "A Framework for Measuring Happiness in Online Social Networks," *Issues in Information Systems*, (XI) 1, pp. 198-203.

Christakis, N. and J. Fowler (2008) "Social Networks and Happiness," Edge: The Third Culture. Retrieved on January 21, 2011 from http://www.edge.org/3rd_culture/christakis_fowler08/christakis_fowler08_index.html

Diakopoulos, N. A. and D. A. Shamma (2010) "Characterizing Debate Performance via Aggregated Twitter Sentiment," in *Proceedings of Conference on Human Factors in Computing Systems*, ACM, pp.1195-1198. Atlanta Georgia, April 10-15, 2010.

Fluss, D. and M. Rogers (2010) "The Growth of Social Media in Customer Service," *CRM Magazine* (14) 9, pp. 8.

Fowler, J. H. and N. A. Christakis (2008) "Dynamic Spread of Happiness in a Large Social Network: longitudinal analysis over 20 years in the Framingham heart study," *British Medical Journal*, 337: A2338.

Greenberg, P. (2010) "The Impact of CRM 2.0 on Customer Insight," *Journal of Business & Industrial Marketing* (25) 6, pp. 410-419.

Gruzd, A., B. Wellman and Y. Takhteyev (2011) "Imagining Twitter as an Imagined Community," *American Behavioral Scientist* (55) 10, pp. 1294-1318.

Han, J. Y., B. R. Shaw, R. P. Hawkins, S. Pingree, F. McTavish and D. H. Gustafson (2008) "Expressing Positive Emotions Within Online Support Groups by Women with Breast Cancer," *Journal of Health Psychology* (13) 8, pp. 1102-1007.

Hogan, B. (2010) "The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online," *Bulletin of Science, Technology & Society* (30) 6, pp. 377-386.

Jansen, B. J., M. Zhang, K. Sobel and A. Chowdury (2009) "Twitter Power: Tweets as Electronic Word of Mouth," *Journal of the American Society for Information Science and Technology* (60) 11, pp. 2169-2188.

Kamvar, S. D. and J. Harris (2011) "We Feel Fine and Searching the Emotional Web," in *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, pp.117-126. Hong Kong, February 9-12, 2011.

Kramer, A. D. I. (2010) "An Unobtrusive Model of 'Gross National Happiness'," In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10).* ACM, pp. 2870290. New York, NY, USA

Loiacono, E. and S. Djamasbi (2010) "Moods and Their Relevance to Systems Usage Models within Organizations: An Extended Framework," *AIS Transactions on Human-Computer Interaction* (2) 2, pp. 55-72.

Marreiros, G., R. Santos and C. Ramos (2010a) "GLSS – Group Learning in Shared Spaces Considering Aspects Like Emotion and Personality," In J. C. Augusto, J. M. Corchado, P. Novais and C. Analide (Eds.), *Ambient Intelligence and Future Trends-International Symposium on Ambient Intelligence (ISAmI 2010)* (Vol. 72) Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 37-44.

Marreiros, G., R. Santos, C. Ramos and J. Neves (2010b) "Context Aware Emotional Model for Group Decision Making," *IEEE Intelligent Systems* (PP) 99, pp.1.

Mason, R.B. (2008) "Word of Mouth as a Promotional Tool for Turbulent Markets," *Journal of Marketing Communications* (14) 3, pp. 207-224.

McLaughlin, C. and L. Davenport (2010) "Brand Community Success Factors: A Study of Two Facebook Brand Community Features," in the *Proceedings of the American Academy of Advertising Conference*, pp. 78-79. Minneapolis, MN, March 18-21, 2011.

Paltoglou, G., S. Gobron, M. Skowron, M. Thelwall and D. Thalmann (2010) "Sentiment Analysis of Informal Textual Communication in Cyberspace," in *Proceedings of ENGAGE 2010, Springer LNCS State-of-the-Art Survey*, pp.13-25. Zermatt, Switzerland, September 13–15, 2010.

Pang, B. and L. Lee (2008) "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval* (1) 1-2, pp. 1-135.

Paolacci, G., J. Chandler and P. G. Ipeirotis (2010) "Running Experiments on Amazon Mechanical Turk," *Judgment and Decision Making* (5) 5, pp. 411-419.

Poynter, R. (2008) "Facebook: the Future of Networking with Customers," *International Journal of Market Research*, 50(1), pp. 11-12.

Pratt, M. (2009) "Cashing in on Tweets," *Computerworld* (43) 29, pp. 22-26.

Read, J. (2005) "Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification," in *Proceedings of the ACL Student Research Workshop (ACLstudent '05).* Association for Computational Linguistics, pp. 43-48. Ann Arbor, MI, USA, June 25-30, 2005.

Smith, A. (2011, June 1) Twitter Update 2011, *Pew Research Center report.* Retrieved June 10, 2010 from http://www.pewinternet.org/~/media//Files/Reports/2011/Twitter%20Update%202011.pdf

Sum, S., M. Mathews, I. Hughes, and A. Campbell (2007) 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, June 28-29.

Takhteyev, Y., A. Gruzd and B. Wellman (2012) "Geography of Twitter Networks. Social Networks," *Special issue on Space and Networks* 34(1), pp. 73-81.

Thagard, P. and F. W. Kroon (2006) "Emotional consensus in group decision making," *Mind & Society* 5(1), pp. 85-104.

Thelwall, M. (2010) "Emotion homophily in social network site messages," *First Monday* 15(4-5).

Thelwall, M., K. Buckley and G. Paltoglou (2011) "Sentiment in Twitter events," *Journal of the American Society for Information Science and* Technology 62(2), pp. 406-418.

Trusov, M., R. E. Bucklin and K. Pauwels (2009) "Effects of Word-of-Mouth versus Traditional Marketing: Findings from an Internet Social Networking Site," *Journal of Marketing* 73(5), pp. 90-102.

Venkatesh, V. (2000) "Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model," *Information Systems Research* (11) 4.

Vogel, C. and J.F. Janssen (2009) "Emoticonsciousness," In A. Esposito, A. Hussain, M. Marinaro and R. Martone (Eds.) Multimodal Signals: Cognitive and Algorithmic Issues. Lecture Notes in Artificial Intelligence, vol. 5398. Springer-Verlag, Berlin, Heidelberg, pp. 271-287.

Zafarani, R., W.D. Cole and H. Liu (2010) "Sentiment Propagation in Social Networks: A Case Study in LiveJournal," In S.K. Chai, J. J. Salerno and P. L. Mabry (Eds.) *Advances in Social Computing* (Vol. 6007). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 413-420

ABOUT THE AUTHOR

Dr. Anatoliy Gruzd is an Associate Professor in the School of Information Management at the Faculty of Management and the Faculty of Computer Science, Dalhousie University (Halifax, NS, Canada). He is also the Director of the Social Media Lab. Dr. Gruzd's research interests include Online Social Networks, Social Networks Analysis, Computer-Mediated Communication and Online Communities. His current research explores how online social media and other Web 2.0 technologies are changing the ways people communicate and disseminate knowledge and information. His research and commentaries have been reported across Canada and internationally in various mass media outlets such as Foreign Affairs, Los Angeles Times, Nature.com, The Atlantic, The Globe and Mail, The National Post, The Canadian Press, CBC TV, CBC Radio, CTV and Global TV.

Transactions on Human-Computer Interaction

ISSN: 1944-3900