Exploring User Purpose Writing Single Tweets [*]

Arifah Che Alhadi   Thomas Gottron   Steffen Staab
WeST – Institute for Web Science and Technologies
University of Koblenz-Landau
56070 Koblenz, Germany
{alhadi,gottron,staab}@uni-koblenz.de

ABSTRACT

Microblogging is a new way of communication among people which allows them to disseminate messages via web, mobile phone, email or instant messaging. In 2010, microbloggers generated 65 million messages a day on Twitter alone. Our hypothesis is that tweeting is an activity which its users perform in order to stills some needs. In this paper, we describe an approach for analysing user purposes in writing single tweets and organize these purposes to taxonomy. We find that people use microblogging for eight different purpose, e.g. promotion, social interaction and expressing emotions. We aim to classify the tweet into categories of purposes.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval: Information Filtering

General Terms

Theory

Keywords

Social Microblogging, Twitter, Purpose, Classification

1. INTRODUCTION

Twitter has gained attention from web users as a way of social communication. Twitter employs a social networking model called 'following' in which each user can choose who they want to 'follow' in order to receive tweets from without an explicit need to get a permission. The objective of this

research is to identify the purpose of people writing single tweets in Twitter and organize these purposes to the taxonomy. "Purpose" has been defined as "a result, end, mean, aim, or goal of an action intentionally undertaken" [2]. In [1], purpose is defined as "why you do something or why something exists". When we are able to identify the purposes, we are better able to understand the mechanisms underlying microblogging and eventually we may even improve the user interfaces by building an intention-driven user interface for writers and readers.

Due to the restricted 140 characters of tweets, we look into what categories should a Twitter messages be classified. Usually classifications are into content categories, e.g. health, politics, or education. In this sense our work is very different because we aim at a classification of Twitter messages into categories of purposes.

The contribution of this paper is an approach to identify the purposes in tweeting. We analyze the purposes of people in tweeting with the aim of answering the question of what are the purposes underlying these activities by such a mass number of tweets. We believe there are various undiscovered purposes of why people keep posting the messages. We aim at a classification of such needs of purposes. For instance, *"The Call for Papers of #iswc2011 is out! Check it out http://bit.ly/gCquHq http://fb.me/SzLUZYub"*, we classified into promotion or marketing under the subcategory of event; and share resources under subcategory of hyperlink.

Outline The outline of the paper is base on our research methodology as showed in Figure 1. First we provide a brief overview of related research for short text classification and users' intentions using microblogging in Section 2. In Section 3, we explain our approach to discover as many purposes as possible in Twitter usage for both read and write the messages. Followed by data synthesis in Section 4. In Section 5, we disccused the outcome of brainstorming session. Finally in Section 6, we present the conclusions that may be drawn from our work.
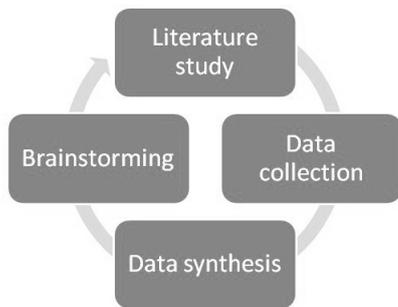
Figure 1: Research Methodology

2. RELATED WORK

Classification of tweets has been well researched recently. Naaman et al. [8] categorized the tweets based on their content to 9 categories. The categories are as follows: information sharing, self promotion, opinions or complaints, statements and random thoughts, me now, questions to followers, presence maintenance, Anecdote (me) and anecdote (others). They focused on the evaluation of how message content varies by user activity, personal networks and usage patterns. The categories are based on user centric such as his/her current status, his/her opinion, or posting question to his/her follower. However, our work focuses on identifying the purpose of users writing those message content for wider audience.

Sankaranarayanan et al. [9] classified tweets into two categories: news and non-news. The idea is to capture tweets that contain information related to the breaking news. Since there is a lot of rich information being transmitted across the globe, this two category are so general. Michelson and Macskassy classify tweets based on their content and generate a topic of interest profile for each user. On the other hand Irani et al. [5], Kwak et al. [7] and Weng et al. [12] classify the tweets based on global trending topics. Sriram et al. group the tweets into the following five categories: news, events, opinions, deals and private messages [10]. None of these classification approaches however categorize the tweet based on the intent of the user posting every single tweets.

Several studies on users intentions using microblogging platform showed that people use microblogging as informal learning [4], business branding, organization communication [13], discussion channel, sharing information/URL [6], etc. However, there is no work on user intentions in individual level of single tweets, but only anecdotal reports in non-scientific websites or magazine

From the different perspective, user intention in annotation resources [11] also can play a major role in supporting a user's social search. Stromhaier [11] propose a novel idea in tag recommendation; purpose tagging which focuses on capturing aspects of intent("what it can be used for"). He stated that keywords or tags issued by user exhibit his/her intent in annotation the resources. It will help bridging the gap between user intent expressed in search queries and the resources users expect to retrieve.

3. DATA COLLECTION

We conducted an online survey to find out the purposes of using Twitter which people use Twitter. Participants were asked for what purposes they used Twitter for both read and write tweets. We provide five purposes as an example to give an idea what we meant by purpose. Participants were allowed to select more than one purpose and give their own purposes under 'others'. As a result, we collected 187 responses and 49 new purposes were added by the participants. The survey was distributed by posting the survey hyperlink to Twitter, Facebook, newsgroup in Koblenz and Mainz, personal contact with researcher from Dallas, France and United Kingdom. We also ask their help to distribute it to their friends, colleagues and students.

4. DATA SYNTHESIS

The fourth step synthesized the data from the literature study and the survey. The data was manually analyzed and classified in order to identify the categories of users' purposes and organized it to build the taxonomy. As example: instant update included in category of social interaction with people.

5. BRAINSTORMING

Finally, we validated the categories of our taxonomy by conducting brainstorming sessions. Our analysis revealed 8 major categories for microblogging as follow:

1. Social interaction with people
2. Promotion or marketing
3. Share resources
4. Give or require feedback
5. Broadcast alert/urgent information
6. Require/raise funding
7. Recruit worker
8. Express emotions

There were also numerous tweets for which we were unable to identify a user purpose in writing those tweets. Thus we will classify it as 'other' purpose. The subcategories and an example tweet for each major category in our taxonomy are given in the Table 1.

Identification of tweet purposes have been carried out manually on a random selection of 1,000 tweets from Choudhury [3] dataset in order to deduce the features for each category of purposes. The dataset consist of 9,998,756 tweets. We use Amazons Mechanical Turk(AMT)[1] as a platform for the human judgement of the given tweet to classify based on user purposes in tweeting. For every tweet, we assigned to five Mturk users to evaluate and categorize it based on the major category. We also allow them to categorize single tweet with multiple categories. Premilinary results are shown in Figure 2. The highest category was social interaction with people (797 tweets) and the lowest is require/raise
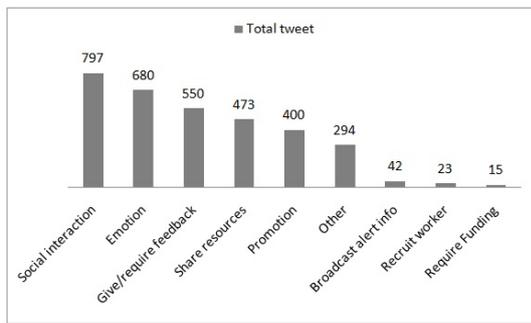
---

[1]https://www.mturk.com

Figure 2: Premilinary Result

funding (15 tweets). However, such figures normaly changes as these results are taken from a sample of 1000 tweets respectively.

6. CONCLUSION

In this research study we have analyzed the purposes of people in tweeting and it can be useful for classifcation of tweets. We hope our approach can support the user experience in searching and browsing information in Twitter. Currently, we are working on automated approaches of detecting features that allow for capturing tweet purposes. Then an automated classification of tweets into purposes will be determined from it.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Cambridge advanced learner's dictionary. website, 2011.
http://dictionary.cambridge.org/dictionary/british/purpose.

[2] Wikipedia the free encyclopedia. website, 2011.
http://en.wikipedia.org/wiki/Purpose.

[3] M. D. Choudhury, Y.-R. Lin, H. Sundaram, K. S. Candan, L. Xie, and A. Kelliher. How does the data sampling strategy impact the discovery of information diffusion in social media? In *Proc. Conf. on Weblogs and Social Media*, pages 34–41, 2010.

[4] G. Grosseck and C. holotescu. Can we use twitter for educational activities. In *PASIS 2010: Proc. of Pasific Asia Conf. on Information System*, pages 173–178, July 2010.

[5] D. Irani, S. Webb, C. Pu, and K. Li. Study of trend-stuffing on Twitter through text classification. In *Proc. Collaboration, Electronic Messaging, Anti-Abuse and Spam Conf.*, 2010.

[6] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: understanding microblogging usage and communities. In *WebKDD/SNA-KDD '07: Proc. of the 9th WebKDD and 1st SNA-KDD Workshop*, pages 56–65, August 2007.

[7] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proc. Int. World Wide Web Conf.*, pages 591–600, April 2010.

[8] M. Naaman, J. Boase, and C.-H. Lai. Is it really about me?: message content in social awareness streams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, CSCW '10, pages 189–192, 2010.

[9] J. Sankaranarayanan, H. Samet, M. D. Teitler, Benjamin E.and Lieberman, and J. Sperling. TwitterStand: News in tweets. In *Proc. Int. Conf. on Advances in Geographic Information Systems*, pages 42–51, November 2009.

[10] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas. Short text classification in Twitter to improve information filtering. In *Proc. Int. Conf. on Research and Development in Information Retrieval*, pages 841–842, 2010.

[11] M. Strohmaier. Purpose tagging: capturing user intent to assist goal-oriented social search. In *SSM '08: Proceedings of the 2008 ACM Workshop on Search in Social Media*, pages 35–42. ACM, 2008.

[12] J. Weng, E.-P. Lim, J. Jiang, and Q. He. TwitterRank: Finding topic-sensitive influential twitterers. In *Proc. Int. Conf. on Web Search and Data Mining*, pages 261–270, February 2010.

[13] F. D. L. Wigand. Twitter in government: Building relationships one tweet at a time. *Information Technology: New Generations, Third International Conference on*, 0:563–567, April 2010.

Table 1: Categories in Taxonomy

Major Categories	Subcategories	Example of tweet
Social interact with people	Colleague, family, friend, celebrity, client, stakeholder, consumer, politician.	"@slimmingjess91 Cool! Is it your 1st or 2nd year? What do you want to do your degree in? Obv science related lol x"
Promotion or marketing	Self promotion, product, team, event, news site, business, blogs, institute, project, music, movie, deals, contest, product, sponsorship, campaign, club.	"The Call for Papers of iswc2011 is out! Check it out http://bit.ly/gCquHq http://fb.me/SzLUZYub"
Share resources	photo, video, hyperlink, mp3	"Corvette Lovers | Flickr - Photo Sharing! http://bit.ly/fpuk17"
Give/require feedback	Questions/answer, advice, recommendation, survey, opinion, support, idea, vote	"Does anyone else get dizzy watching computer animations? I get vertigo or a type of seasickness when I simply view some of them."
Broadcast alert/urgent information	Emergencies, traffic, weather, environment disasters.	"RT @NJ_News:Accident update:3 buses, motorcycle accident in Lincoln Tunnel causes hour-long delays, multiple injuries http://bit.ly/frRazs"
Require/raise funding	Health, campaign, personal, charity, project.	"Cagsil-Help Homeless In America Projects, http://hubpages.com/t/189f48 #America #Homeless #Shelters #Mission #Funding #Aid #Help #Truth"
Recruit worker	Staff, volunteer, consultant	"What are YOU doing this spring? We know you like movies, so why not be a volunteer for TFF 2011? We need you! http://bit.ly/gR6gLV"
Express emotions	Complaint, protest, compliment, positive or negative feeling.	"Congrats to Tom Heath and Chris Bizer - their new #linkeddata book is published! - http://bit.ly/?xWHy"
Other	None of the categories	"beer beer beer beer"