

INFO 290
 Prof. Irwin King
 19. 09. 2011
 Kim, Eungchan

Project proposal

The Method of Overcoming a 140-character limit in Twitter: URL links

Introduction

According to Wikipedia, Twitter is defined as “an online social networking and micro-blogging service,” which was founded in 2006. Through Twitter, people can send and read text-based posts, known as tweets, of up to 140 characters. In the sense of text-based service, Twitter is somehow similar to traditional text messaging and e-mail. Twitter is a one-to-many communication tool rather than one-to-one method. Since 2006, the range of Twitter usage has covered from sharing individuals’ opinions or thought to informing the current news and even, advertising products. And Twitter is applicable platform to various devices such as desktop PC, tablet PC, and even smart phones. It is no doubt that Twitter is a usable and light-weight tool to express their opinions.

Twitter users, however, can only use 140 characters so many skeptics might wonder why Twitter restricts to writing space to 140 characters and how much tweet can deliver users’ thought thoroughly. Here is an interesting story why Twitter limits its writing spaces to 140 characters.[1] The 140-character limit is based on not a scientific reason but an empirical reason. In 1985, a German typewriter discovered 160-character maximum would be enough space as the writing communication tool after counting the number of letters, numbers, punctuation marks and spaces on the papers. That empirical research bounded the maximum words of the text messaging to 160-characters. Twitter uses 20-characters for the users’ unique address and the rest of 160-characters which is 140-characters for users’ writing space.

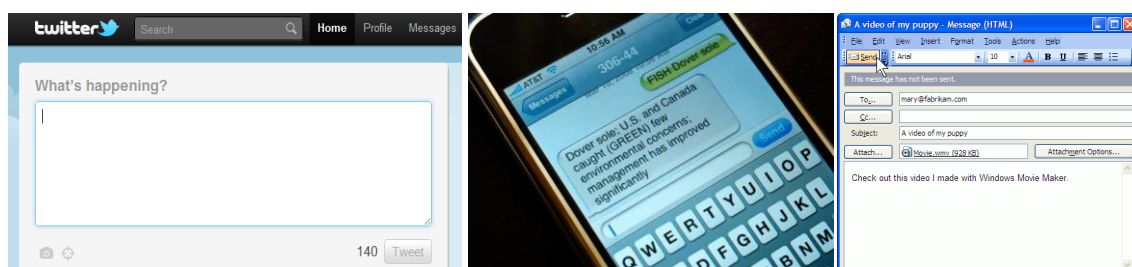


Figure 1. Various types of writing spaces such as Twitter, text messaging and email

Motivation

I have used Twitter since March in 2010. I mostly communicate with my friends via Twitter. But, when I tried to share some interesting articles on the web, it was really hard to fit the content within 140 characters. I should choose one of the following routes around that limitation:

- 1) Using the abbreviation of the term
- 2) Splitting the message into a couple of tweets
- 3) Using the URL links to the article



Figure 2. Three methods to overcome the characters limit: using the abbreviation of the term, splitting the message into a couple of tweets and using the URL links to the article

The first method may be okay to my close “followers,” whose timelines have my tweets stream, but can make most of “followers” confused of the meaning and even mislead “followers.” The second route is really troublesome for both Twitter users and “followers.” Especially, the interface of Twitter enables the newer tweet to be placed at the upper. In other words, new tweet is posted on the top of timeline. For this reason, if users split his/her tweet to a couple of tweets, it should be rearranged by time order, otherwise “followers” have to read tweets from down to up. Finally, it seems to me that the third method is the most useful way to overcome the character limitation. Even though it is somehow risky for “followers” not knowing what the URL links point to, the “followers” are willing to click the URL links due to their previous choice to be my “followers.” In addition, many techniques of using URL links appears such as sharing links to longer articles, discussions, posts, videos — anything that lives behind a URL links. For example, foursquare uses URL links to show their location. In this sense, I want to do my research about the usage of URL links in Twitter sphere.

Research Question

I address the following research questions: When do people use URL links in their tweets? What is the ratio of URL links tweets to non-URL links tweets in a specific user? Why do users use URL links in their tweets? What is the popular URL link domain among tweets?

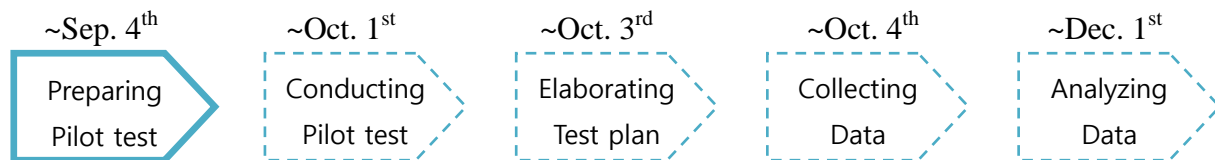
Methodology

I will collect data from the Twitter. I will select 20 twitter users based on the number of tweets they create over the public timeline. I should be careful to choose the users because of the existence of robots which generate tweets automatically. I will get rid of robots in the candidate list. I plan to gather all of the tweets that the candidate users make for a week. Then, I will focus on tweets which include URL links among the data set.

I try to figure out when users use URL links in their tweets. At first, I will count the number of words in the tweets before an URL links appears and also find the ratio of URL links tweets to non-URL links tweets in a specific user. After that, I will categorize tweets to find out why users use URL links in their tweets. There are several ways to classify tweets based on the purpose of tweets. Regarding of the purpose classification in [2], I can apply this criteria to classify tweets which contain URL links as Personal Updates, Directed Dialogue, Real-time Sharing and Information seeking. Another method in [3] can be used to classifying tweets as Daily Chatter, Conversations, Sharing information/URLs, and Reporting news. Moreover, I can also categorize tweets by the genre of what tweets URLs contain such as text, pictures, music, movie clips and maps. After classifying the tweets, I will figure out which URL domains are popular to use in Twitter.

Plan

- 1) September - Selecting candidate users and preparing a pilot test
- 2) October 1st week - Conducting the pilot test
- 3) October 2nd ~3rd week - Elaborating a test plan
- 4) October 4th week - Collecting data
- 5) November 1st~4th week - Classifying data and analyzing data
- 6) December 1st week – Presentation



Reference

- [1] Millan, M. How was 140 hatched as texting's upper limit? Not-so-scientific testing. *LOS ANGELES TIMES*(May 18 2009).
- [2] Westman, S. and Freund, L. *Information interaction in 140 characters or less: genres on twitter*. ACM, City, 2010.
- [3] Java, A., Song, X., Finin, T. and Tseng, B. *Why we twitter: understanding microblogging usage and communities*. ACM, City, 2007.