Study on the Relationship between Profile Images and User Behaviors on Twitter

Tomu Tominaga
Graduate School of Engineering Science,
Osaka University
Osaka, Japan
tominaga@nishilab.sys.es.oasaka-u.ac.jp

Yoshinori Hijikata
Graduate School of Engineering Science,
Osaka University
Osaka, Japan
hijikata@sys.es.oasaka-u.ac.jp

ABSTRACT

In recent years, many researchers have studied the characteristics of Twitter, which is a microblogging service used by a large number of people worldwide. However, to the best of our knowledge, no study has yet been conducted to study the relationship between profile images and user behaviors on Twitter. We assume that the profile images and behaviors of users are influenced by their internal properties, because users consider their profile images as symbolic representations of themselves on Twitter. We empirically categorized profile images into 13 types, and investigated the relationships between each category of profile images and users' behaviors on Twitter.

Categories and Subject Descriptors

H.5 [Information Interface and Presentation]: Group and Organization Interface—Web-based interaction; H.1 [Models and Principles]: User/Machine Systems—Human Factors

General Terms

Human Factors

Keywords

Twitter; User behavior; Profile images

1. INTRODUCTION

Twitter[1] is one of the most popular microblogging services that witnesses 284 million monthly active users[2]. It enables users to send and read short 140-character messages called "tweets". It provides users with a platform to both deliver (gather) information to (from) other people as well communicate with them. Compared to other social media services, the features and characteristics of Twitter are unique; therefore, it has attracted considerable research interest in recent years [1, 2, 10], which are known as user profiling studies.

When users join Twitter, they usually begin with setting up their user profiles, which includes a profile image, some introductory information and so on. Researchers have profiled users based on their introductory information or the content of their tweets [1, 2], or based on the name, location, and sex of the user [1, 4]. The reason behind using the aforementioned features for user profiling is that they are indicative of a user's internal properties. In this study, internal properties refers to the usage objectives, personal preferences, and other unique personality traits and characteristics of a person. However, to the best of our knowledge, no study has yet been conducted to study the relationship between profile images and user behaviors on Twitter. On the other hand, several user profiling studies based on profile images have been conducted for Facebook[3] [5, 7, 13]. They analyzed the relationship between internal properties such as self-construction [5], narcissism [7], self-presentation [13] and the corresponding user profiles on Facebook.

We assume that a profile image is influenced by a user's internal properties. A profile image is displayed alongside a user's tweet on their followers' timelines. We believe that users consider profile images as representative symbols of the people they follow on Twitter. This implies that users' profile images are influenced by their internal properties. Further, we assume that the internal properties of a person influences their behavior on Twitter. The reason behind our assumption is that users usually post tweets about interesting news or personal events on Twitter for various purposes such as advertising, promotions, or communication. However, it has not been extensively studied what type of internal properties affect a user's choice in profile images or a user's behaviors on Twitter. As a result, it is unclear what type of internal properties is a significant contributing factor in the users' choice of profile image and their behaviors. Therefore, we investigated the relationship between them in order to gain insight into users' internal properties.

We define user behavior as a user's usage history such as the number of people the user follows, the number of people the user is followed by, the number of tweets the user posts daily. When we observe a particular type of profile images, we determine how users belonging to that type use Twitter. Therefore, the findings in this study can contribute to future user profiling studies on Twitter.

[1]https://twitter.com
[2]https://about.twitter.com/company (2015.01.05)
[3]https://www.facebook.com/

WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742467.

2. RELATED WORK

2.1 User Characteristics

In the studies that extracted users' characteristics, researchers identified user characteristics through the tweet content of target users. To achieve this, the LDA[4] model has been extensively used. Pennacchiotti and Gurumurthy used LDA to associate Twitter users to a number of topics and based on the results recommended users with similar interests on the Twitter network to the target user [9]. Ramage et al. improved the accuracy of LDA based user recommendation by considering hashtags(#), replies (@), and emoticons in tweet content [10].

2.2 Link Characteristics

It is interesting to study users' following behaviors on Twitter. When we regard users as nodes, the following relationships can be observed as links. Java et al. reported that users' links can be categorized into three types by analyzing Twitter networks: (1) information source (a link from a hub user to others), (2) friends (a link between friends), (3) information seeker (a link from a user to a hub user) [6]. Cha et al. and Kwak et al. investigated the large-scale influential power of Twitter users, and stated that links on Twitter do not represent social relationships in the real world but represent information sharing relationships [3, 8].

3. WHY PROFILE IMAGE?

Figure 1 shows an example of a user using Twitter on a Web browser. In this figure, the tweets from users followed by the user (followees) are displayed on the right side in chronological order (timeline). It can be said that profile images serve as symbolic representations of users, and based on them users can identify the authors of tweets on their timeline. Therefore, a profile image may imply how a user wishes to be perceived by other users.

For the aforementioned reasons, it can be inferred that a user's choice of a profile image may be influenced by their internal properties. The manner by which users select profile images depends on how they want to express themselves on Twitter. A user's anonymity consciousness[5] and IT literacy[6] may also affect his/her choice in a profile image. Further, we also believe that a user's internal properties affects his/her behavior on Twitter.

In brief, users' choices in profile images and users' behaviors on Twitter are influenced by their internal properties. However, to the best of our knowledge, no study has yet been conducted to comprehensively clarify these relationships. Therefore, to address the same, in this study, we conducted a comprehensive analysis of profile images and users' behaviors on Twitter. Further, we categorized users' profile images based on the objects that appeared in the images and examined the behaviors of the corresponding users. In the next section, we explain the different categories of profile images established in this study, and show the validity and coverage of the categories.

[4]It stands for Latent Dirichlet Allocation, which is one of topic models in the area of natural language processing

[5]Fear of being maliciously targeted while using the Internet.

[6]The knowledge or ability to deal with some information, data, or service.

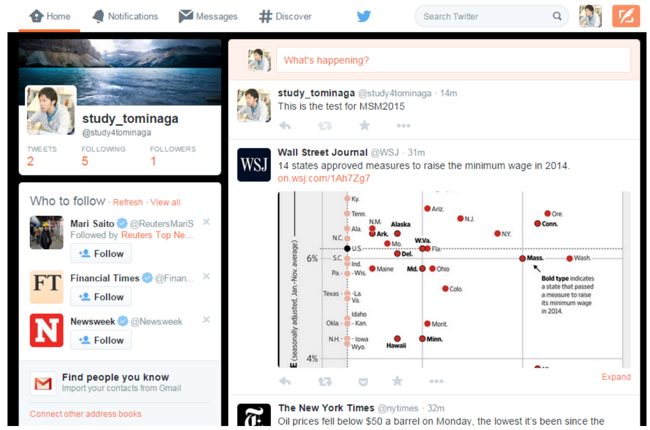

Figure 1: Timeline on Twitter

4. PROFILE IMAGE CATEGORIES

We established 13 types of categories based on the objects observed in users' profile images: oneself, self portrait, hidden face, associate, different person, letter, logo, otaku, character, animal, object, scene, and egg. These definitions are listed in Table 1. We proved the validity and coverage of these categories through two types of experiments.

In the experiment for verifying the validity of the categories, we prepared a set of 300 users' profile images that were randomly selected from 4,394,542 users on Twitter. We asked 10 test subjects to classify the 300 profile images into the 13 categories mentioned above. Next, we evaluated the consistency in the 10 results by Siegel coefficient [11]. The coefficient result obtained was 0.70, which indicates that the classifications were considerably similar.

Further, it is essential to verify the coverage of these categories. For verifying the coverage, we asked four test subjects to classify respectively 300 different profile images into 14 categories, which consisted of the 13 categories stated above and an additional category "others". In total, 1200 profile images that were randomly selected from Twitter were classified. When it is difficult for a test subject to classify a profile image into any of the 13 categories, he/she is asked to classify the profile image into "others". From the 1200 users, it was found that 113 users had already quit using Twitter prior to this experiment. This is an expected outcome as the 1200 users were selected from a considerably large pool of 4,394,542 users, which is likely to have a few users that are no longer active. Therefore, the classification results for the 113 users were disregarded. Finally, as per the classification results, 93 users were classified into the "others" category, which means the coverage is 91.4%.

5. INVESTIGATION AND RESULT

5.1 Method

We aim to investigate the behaviors of users on Twitter for each type of category. Users are categorized according to the 13 different categories of profile images defined in the previous section. We target the types of user behaviors that are commonly observed on Twitter. The most important user behaviors on Twitter presumably are the "follow" and

Table 1: The categories of profile images

Category	Explanation
Oneself (On)	Image or photo of the user himself/herself
Self portrait (Sp)	Illustration of the user's face
Hidden face (Hf)	Photo of the user but with some of the features of the face hidden
Associate (As)	Photo of the user and his/her friends together
Different person (Dp)	Image or photo of a person other than the user (usually a celebrity)
Letter (Le)	Image consisting only letters
Logo (Lo)	Image consisting a logo
Otaku (Ot)	Image of beautiful female characters from Japanese anime or manga
Character (Ch)	Image of famous cartoon characters other than female characters from Japanese anime or manga
Animal (An)	Image or photo of animals such as birds, cats, etc.
Object (Ob)	Image or photo of the user's possessions such as a ball, a bike, etc.
Scene (Sc)	Image or photo of a natural scenery
Egg (Eg)	Default image (egg icon)

"tweet" actions. Users engage in these actions to receive and deliver information.

First, we gathered 100 users for each category. Second, we obtained the number of followers, followees, and tweets for each user by using API supported by Twitter.

5.2 Result

To show our results, we use a boxplot, which is convenient for graphically depicting groups of numerical data through their median and variation[7].

5.2.1 Followers and Followed Users

We calculated the ratio of the number of followers to the number of followees (denoted as FF).

$$FF = followers \ / \ followees \quad (1)$$

The higher the value, the more a user is likely to gain new followers, even if they do not actively follow other users.

Figure 2 shows the boxplot of FF. The FF of logo users (Lo) is typically very high (median FF of 2.20). Users in this category are generally official groups such as a company or a department. It is usually easy for such users to attract the attention of other users especially since they may already be well known in the real world. Accordingly, it can be said that users who are familiar with a company or organization tend to follow their Twitter accounts, especially since they may consider the logo users' tweets important and relevant to their own interests.

The FF of letter users (Le) and oneself users (On) is also generally very high (FF = 1.25, 1.15). We found that letter users are usually student groups, private management companies, or automatic posting accounts[8]. Their objectives are promotion or advertisement similar to logo users; however, they are not as widely known in the real world as logo users. Some letter users employ the "following back" strategy [12] for acquiring many followers. Therefore, the FF of letter users (Le) yielded a lower value than that of logo users. The majority of oneself users (On) are scholars, journalists, entrepreneurs and so on. They typically post news or articles relevant to their specific fields; therefore, they are followed by users who are interested in the concerned domains.

[7]http://en.wikipedia.org/wiki/Box_plot

[8]bot, a system posting tweets automatically

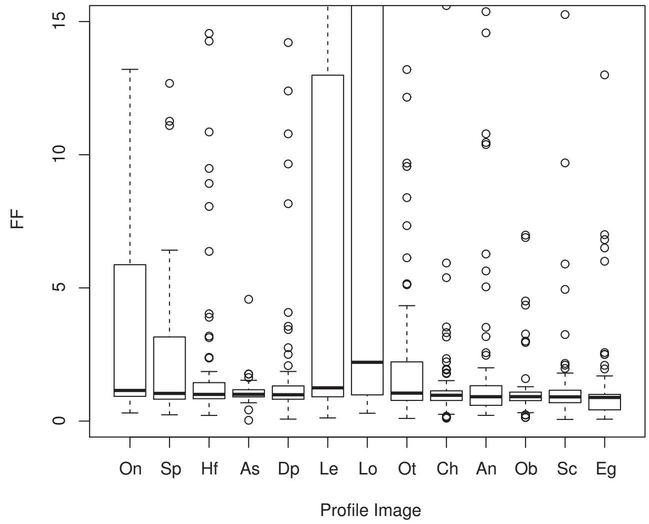

Figure 2: The ratio of the number of followers to the number of followees

5.2.2 Tweets

As we can obtain the total number of tweets in a user's lifetime (time period from the day the user has started using the service to the present day), we use it to determine the average number of tweets per day. We computed the average number of tweets in a day, R_{tw}, by dividing the total number of tweets (denoted as $Tweets_{All}$) by the length of the user's lifetime (denoted as *Span*), as the lifetime differs for different users.

$$R_{tw} = Tweets_{All} \ / \ Span \quad (2)$$

Figure 3 shows the boxplot of R_{tw} in each category. The median of otaku users (Ot) shows a surprisingly high value (R_{tw} = 25.40). Otaku users (Ot) post a large number of tweets mainly about their preferences and hobbies, such as Japanese anime, manga, and games. They typically prefer posting tweets in a one-way style of communication, i.e., they do not usually use the reply or mention actions. We believe these characteristics resulted in the high value.

The category of hidden face users (Hf) showed the second highest median (R_{tw} = 12.66). These users use pictures in which they appear facing away from the camera, have their faces hidden by their hands, or have their heads hung low. We believe that such users feel at ease to freely present their opinions or ideas as they know that they cannot be identified through their pictures.

The R_{tw} of logo users (Lo), letter users (Le), and oneself users (On) was low (R_{tw} = 2.80, 3.42, 3.69), which was contrary to the high FF values. Logo users (Lo) and letter

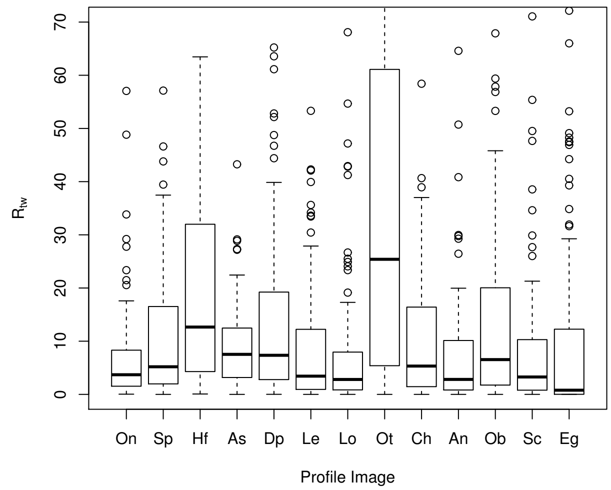

Figure 3: The average number of tweets a day

users (Le) usually post tweets for promotions or advertisements. They may be restricting their number of posts as a large number of promotions or advertisements related posts may displease followers. It can be assumed that oneself users (On) might be only posting about quality news articles or information as they prefer branding themselves and securing their credibility among their followers.

6. SUMMARY

In this study, we comprehensively investigated the relationship between users' profile images and behaviors on Twitter. We assume that a relationship exists between the type of profile images and the type of users' behaviors. In particular, we empirically divided profile images into 13 categories and examined users' behaviors in each category. We found the following distinctive behaviors on Twitter.

Letter users and logo users typically do not post a large number of tweets. The reason behind this may be that they avoid being hated by followers because of a lot of posting tweets for advertising their products or service. We found that otaku users posted tweets more frequently than other users. It can be assumed that otaku users post a large number of tweets about their interests and preferences. On the other hand, we were unable to identify unique user behaviors in other categories, such as animal, object, scene, and so on. We believe that they are general users.

In this study, we did not identify the types of internal properties that affect the selection of profile images and user behaviors. These relationships will help us better understand users' behaviors on Twitter. We will analyze users' internal properties in our future works.

7. REFERENCES

[1] S. Bergsma, M. Dredze, B. V. Durme, T. Wilson, and D. Yarowsky. Broadly Improving User Classification via Communication-Based Name and Location Clustering on Twitter. In HLT-NAACL, number June, pages 1010–1019, 2013.

[2] J. D. Burger, J. Henderson, G. Kim, and G. Zarrella. Discriminating Gender on Twitter. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, pages 1301–1309, 2011.

[3] M. Cha and K. P. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, volume 10, pages 10–17, 2010.

[4] E. Cunha, G. Magno, M. A. Gonçalves, C. Cambraia, and V. Almeida. He votes or she votes? Female and male discursive strategies in Twitter political hashtags. PloS one, 9(1):e87041, Jan. 2014.

[5] N. J. Hum, P. E. Chamberlin, B. L. Hambright, A. C. Portwood, A. C. Schat, and J. L. Bevan. A picture is worth a thousand words: A content analysis of Facebook profile photographs. Computers in Human Behavior, 27(5):1828–1833, Sept. 2011.

[6] A. Java, X. Song, T. Finin, and B. Tseng. Why We Twitter : Understanding Microblogging. In Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, pages 56–65, 2007.

[7] S. Kapidzic. Narcissism as a predictor of motivations behind Facebook profile picture selection. Cyberpsychology, behavior and social networking, 16(1):14–19, Jan. 2013.

[8] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a Social Network or a News Media? Categories and Subject Descriptors. In Proceedings of the International World Wide Web Conference, pages 591–600, 2010.

[9] M. Pennacchiotti and S. Gurumurthy. Investigating topic models for social media user recommendation. In Proceedings of the 20th international conference companion on World wide web, pages 101–102, New York, New York, USA, 2011. ACM Press.

[10] D. Ramage, S. Dumais, and D. Liebling. Characterizing Microblogs with Topic Models. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pages 130–137, 2010.

[11] S. Siegel. Nonparametric statistics for the behavioral sciences., volume 125. McGraw-Hill, New York, 1956.

[12] J. Weng, E.-P. Lim, J. Jiang, and Q. He. TwitterRank. In Proceedings of the third ACM international conference on Web search and data mining - WSDM '10, page 261, New York, New York, USA, Feb. 2010. ACM Press.

[13] C. Zhao and G. Jiang. Cultural differences on visual self-presentation through social networking site profile images. In Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11, pages 1129–1132, New York, New York, USA, 2011. ACM Press.