

# An Analysis of Microblogging Behavior on Sina Weibo: Personality, Network Size and Demographics

Lingyu Wang<sup>1,2</sup>, Weina Qu<sup>1</sup>, and Xianghong Sun<sup>1</sup>

<sup>1</sup> Institute of Psychology, Chinese Academy of Sciences, Beijing, China  
quwn@psych.ac.cn

<sup>2</sup> University of Chinese Academy of Sciences, Beijing, China

**Abstract.** In China, Sina Weibo is one of the most popular microblogging services. The relationship between Weibo users' microblogging behavior and their personality as well as demographics was addressed in current study. 498 users completed the online survey and their social networking data was downloaded based on the Application Programming Interfaces. The results showed that: 1) Weibo users' personality was associated with some aspects of microblogging behavior, especially, Extraversion and Openness significantly positively correlated with some aspects of microblogging behavior, while neuroticism significantly negatively correlated with that; 2) There were associations between some demographical factors and some aspects of microblogging behavior, for instance, the time spent in microblogging related to the age, educational level and monthly discretionary money, respectively. 3) The effect of gender on Weibo social networking size was significant, that was the amount of male Weibo users' followers was significantly greater than that of female users'.

**Keywords:** Personality, Microblogging, Sina Weibo, Demographics.

## 1 Introduction

As a new way to build relationship, Social Networking Services (SNSs) are being almost the first choice for countless users all around the world. For example, in the US, Facebook and MySpace have more than 117 million US visitors in 2009, which continue to attract wider range of users throughout the world [1]. Among all the SNSs, microblog has emerged as a new and popular platform for short, frequent communication which differs from the most SNSs in that the relationship of follower and followee doesn't require reciprocation [2]. One of the most popular microblog, Twitter, a microblogging service which established less than six years, has more than 510 million users in July 2012[3]. However, in China, where Facebook and Twitter were not available, one substitute for twitter is Weibo, a Twitter-like social networking service established in August 2009. At the very beginning, only 13.8percent of Internet users used Weibo in 2010[4]; in 2011, almost half the Chinese Internet population (48.7 percent) used the service—a total of 250 million users [5].

What motivates so many people to use this platform? A large amount of survey aimed at analyzing the motivations and usage of microblogging service. Most of the previous analyses involve two sorts of antecedents: 1) self-description, 2) demographical

factors [6]. Despite the increasing research on microblogging, the correlation between microblogging behavior and personality traits is not known very much. In the context of China, Sina Weibo is one of the most popular microblogging services. Based on this platform, we conducted this study to enhance the understanding of the relationship between microblogging behavior and personality traits as well as demographical factors.

## 2 Methods

### 2.1 Data Collection

In our study, we mainly focused on the general active users of Sina Weibo. In order to select the target group from the population, we conducted the screening process in two stages. At the first stage, we aimed at getting the Weibo users' ID as many as possible. Beginning with 100 users, we firstly downloaded the users' relationship network through breadth search by calling Sina Weibo's open application programming interface (APIs) in March 2012. Removing the repeated relationship, we got 15,767,158 users' IDs. Then we captured these users' profiles, including the amount of their updates, followees and followers. After preliminary analysis of these data, we found that the range of the amount of followers from 1000 to 3300 was very important. In this range, the amount of the users was relatively higher and the users' ID was less likely to belong to the same user. So we got 291,039 users' IDs of which the amount of followers fell in the range between 1000 and 3300.

At the second stage, we took the 291,039 users' IDs as the expanded seed for breadth search. We captured 1,116,408,085 users' IDs by following these users' relationship networking on April 12, 2012. After removing the repeated IDs, there were 96,497,290 users' IDs. Then we downloaded these users' profiles on April 18, 2012. On average, there were 136.65 updates ( $SD=788.87$ ) per user in this dataset. We defined active users as people who posted at least  $532(M+0.5SD)$ , in this dataset) updates and we excluded the highest 5 percent data. So the interval of the active users' total updates was [532, 4194]. According to this criterion, we got 6,047,966 users. And we downloaded the detailed information of 5,919,087 users successfully on April 25, 2012.

Also there were many star users and advertisers. These people should be excluded from the target group because they were not the general active users. According to the following standards, we screened the target group: 1) the amount of updates was equal or more than 532; 2) the last microblog was posted after 2012; 3) the interval between the time of the last microblog released and the time of Weibo ID registered was more than one month; 4) the average amount of microblogging was not more than 100 every day. Then we got 5,807,999 users' IDs. On average, there were 2.84 updates per user in this dataset every day. Using 40 as the up boundary of the average daily amount of microblogging, we got the range between 2.84 and 40. Finally, we screened out 1,953,485 users who fell in this interval as our target group. We randomly divided these people into 20 parts (19 parts including 200 thousands, the remained as the last part). Then we randomly selected 30 thousands users from the first part to take part in our study. We used the "@" function of Weibo to invite these users to participant an online survey and 505 users responded. After further screening according to the data integrity, 498 copies of data were remained.

Participants completed the online questionnaire from May 1 through July 22 in 2012. The questionnaire contained 3 parts, as following: 1) Demographics information; 2) Big

five personality measures; 3) Internet and Weibo use experience. And also we downloaded these users' profiles, including the amount of updates, favorites, followees and followers, as well as the social networking size which referred to the total of the user's followees and followers.

## 2.2 Measures

**Demographics Information.** Personal information, including gender, age, educational level (Edu.), occupation (Occu.), disposable money every month (Monthly Money), growth location, place of the current residence (Cur. residence) and time at current address (Time Add.), was collected in the first part.

**Big Five Personality Measures.** In the second part, we adopted the Big Five Inventory-44 (BFI-44) to measure the users' personality. This is a self-report personality inventory which is designed to assess the Big Five Factors of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience [7]. The BFI-44 has substantial internal consistency and retest reliability. For example, John et al stated that "in U.S. and Canadian samples, the alpha reliabilities of the BFI-44 typically range from .75 to .90 and average above .80; three-month test-retest reliabilities range from .80 to .90, with a mean of .85" [8]. We translated all the questions into Chinese and all participants answered this inventory according to a 5-point Likert scale (from 1= "strongly disagree" to 5= "strongly agree").

**Internet and Weibo Use Experience.** In the third part, we measured the users' general use of the internet and Weibo by asking the following 8 questions which all adopted a 5-point scale except the last two. The first two questions were used to measure the users' experience of internet use: (a) How long have you started to use the internet (i.e. duration of starting to use Internet, Dur.Internet)? (1= less than 1 year, 2= 2-3 years, 3= 3-5 years, 4= 5-10 years, 5= more than 10 years) and (b) How often do you use the internet now (i.e. frequency of using internet, Fre.Internet)? (1= sometimes when necessary, 2= less than 2 hours a day, 3= 2-4 hours a day, 4= 4-8 hours a day, 5= more than 8 years a day).

The next two questions were about the users' general Weibo use experience: (c) How often do you use Sina Weibo now (i.e. daily frequency of using Weibo, Fre.Weibo)? (1= seldom, 2= sometimes, 3= several times a week, 4= several times a day, 5= always) and (d) How much time do you spend on Sina Weibo every day at present (i.e. daily duration of using Weibo, Dur.Weibo)? (1= less than 15 minutes, 2= 15-30 minutes, 3= 30 minutes to an hour, 4= 1-2 hours, 5= more than 2 hours).

Another four questions were adapted to measure the idiosyncratic features of Weibo: (e) How many people do you know and are familiar with among your followees on Weibo (i.e. number of acquaintance among followees, Num.Followee)? (1= 0-50, 2= 50-100, 3= 100-200, 4= 200-500, 5= more than 500). (f) How many people do you know and are familiar with among your followers on Weibo (i.e. number of acquaintance among followers, Num.Follower)? (1= 0-50, 2= 50-100, 3= 100-200, 4= 200-500, 5= more than 500). (g) About how many times do you vote on Weibo every week (i.e. times of voting weekly, Times.Voting)? (1= 0 time, 2= 1-3 times, 3= 4-6 times, 4= more than 7 times). (h) About how many times do you use the "@" function of Weibo to mention others a day (i.e. times of using "@" daily, Times. @)? (1= 0 time, 2= 1-3 times, 3= 4-9 times,

4= more than 10 times). The participants answered the last two questions according to a 4-point Likert scale.

All these questions mentioned above could be seen as the index of the users' microblogging behavior.

### 3 Results

#### 3.1 Sample Distribution of Demographics

In our study, 498 participants finished the online survey. There were 189 males and 309 females (see Table1). And 14 percent of these users were under 18 years, 80 percent were 18-30 years and 6 percent were above 31. Among these users, 42.2 percent had a bachelor's degree, 22.1 percent had junior college degree, 30.5 percent had high school diploma and under. As for their occupation, 52.6 percent of the users were students. And 63.6 percent of these users were born in city and 84.2 percent now are living in the city.

**Table 1.** The sample distribution of the Weibo users' demographics (n=498)

Demographics	Frequency	Percent	Demographics	Frequency	Percent
Gender			Growth Location		
Male	189	38.0	Provincial level city	135	27.1
Female	309	62.0	Ordinary city	182	36.5
			Town	104	20.9
			Village	77	15.5
Age group			Occu.		
Under18	70	14.0	Worker	5	1.0
18-20	136	27.3	Civil servant	12	2.4
21-25	190	38.2	Medic & nurse	10	2.0
26-30	72	14.5	Office worker	93	18.7
31-35	16	3.2	Journalist	14	2.8
36-40	5	1.0	Student	262	52.6
41-45	6	1.2	Researcher & teacher	16	3.2
Above 46	3	0.6	Independent operator	18	3.6
			Unemployed	32	6.4
			Else	36	7.2
Edu.			Cur. Residence		
High school & Under	152	30.5	Provincial level city	226	45.4
Junior college	110	22.1	Ordinary city	193	38.8
Bachelor	210	42.2	Town	67	13.5
Master & Above	26	5.2	Village	12	2.3
Monthly Money			Time Add.		
Less than 500	116	23.3	1 year & Less	53	10.6
500-1000	115	23.1	1-3 years	123	24.7
1000-2000	126	25.3	3-5years	74	14.9
2000-5000	111	22.3	5-10year	58	11.6
More than 5000	30	6.0	More than10 years	190	38.2

### 3.2 Personality and Microblogging Behavior Correlations

Our results showed that Weibo users' microblogging behavior correlated with their personality traits. As table 2 showed, Extraversion positively correlated with the frequency of using Weibo, the number of acquaintance among followees and followers, the times of voting weekly as well as the daily times of using the "@" function. But Neuroticism negatively correlated with the number of acquaintance among followers. Conscientiousness was positively correlated with the times of voting weekly and Openness was positively correlated with the daily times of using the "@" function.

**Table 2.** The descriptive statistics of personality traits and Pearson correlation values between microblogging behavior and personality traits

	<i>M</i>	<i>SD</i>	Dur. Internet	Fre. Internet	Fre. Weibo	Dur. Weibo	Num. Followee	Num. Follower	Times. Voting	Times. @
Extra.	3.03	0.63	.07	.00	.09*	.02	.19**	.23**	.14**	.19**
Agree.	3.61	0.57	.12**	.03	.05	.03	-.02	.00	-.01	.05
Cons.	3.15	0.57	.06	-.03	-.08	-.08	.06	.04	.09*	.05
Neuro.	3.06	0.64	-.05	.02	.02	.03	-.05	-.09*	-.09	-.08
Open.	3.60	0.57	.13**	-.03	.02	.02	.09	.06	.05	.10*

Note: \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ , the same as following tables.

### 3.3 Relationship between Microblogging Behavior and Demographics

In table 3, the results showed that the users' microblogging behavior were not independent of the demographical valuables. For example, the frequency of using Weibo was different across the educational levels and occupation and the duration of using Weibo was different across the age, educational levels, as well as monthly disposable money. Also there was an association between the number of acquaintance among followees and age. And there was a significant difference in the times of voting weekly across the current residence. The daily times of using the "@" function related with gender, age and occupation too.

**Table 3.** Contingency table test for the users' microblogging behavior and demographics (Pearson Chi-Square values)

	Dur. Internet	Fre. Internet	Fre. Weibo	Dur. Weibo	Num. Followee	Num. Follower	Times. Voting	Times. @
Gender	1.63	1.04	7.81	7.37	9.06	5.48	7.01	9.15*
Age Group	170.44***	104.50***	17.59	53.90**	56.53***	27.95	22.06	33.54*
Edu.	75.51***	41.17***	36.28***	32.84***	12.23	14.50	16.79	11.12
Occu.	110.63***	104.96***	60.65**	48.17	45.06	47.98	28.60	53.15**
Monthly Money	126.96***	77.83***	21.02	39.35***	25.21	13.62	15.61	15.99
Growth. Location	32.38***	8.86	11.88	11.74	8.86	9.16	14.96	14.70
Cur. Residence	55.51***	22.06*	14.58	9.04	13.97	12.07	20.18*	10.50
Time Add.	27.79*	17.40	12.42	16.26	12.68	18.18	19.35	7.23

### 3.4 Differences in Demographics

The results in table 4 showed that the effect of gender and growth location on the amount of followers and Weibo social networking size was respectively significant. Specifically, the amount of followers and Weibo social networking size of the male users was significantly more than that of the female users. The users grown up in the town had significantly more followers and bigger social networking size than those grown up in the provincial level city, Ordinary city and Village.

**Table 4.** The significance test for the users' profiles under the demographics by independent sample T-test and ANOVA

	Gender (t-value)	Age group	Edu.	Occu.	Monthly Money	Growth Location	Cur. Residence	Time Add.
Amount of Followers	2.84**	0.34	0.38	0.40	0.41	3.13*	1.78	0.94
Amount of Followees	1.29	1.81	3.23*	1.59	1.81	1.98	0.47	1.30
Amount of updates	0.56	0.46	1.84	1.84	0.91	0.27	0.07	1.50
Amount of favorites	0.54	0.23	0.90	2.45*	1.89	0.84	1.50	1.21
Social networking size	2.87***	0.38	0.54	0.40	0.47	3.07*	1.59	1.04

And also there was significant difference in the amount of favorites across the users' occupation. And specifically, the amount of favorites of the journalist users was the highest and significantly more than that of other users except for the worker users. And the amount of followees was significantly different across the users' educational levels. The results of Post Hoc Tests showed that the users having junior college diploma had significantly more followees than those having high school diploma and under, as well as master degree and above.

## 4 Conclusions

In this study, first we picked out the active users of Sina Weibo and downloaded their users' profiles information, and then we invited the active users to take part in our online survey. We measured their personality and microblogging behavior and collected their demographical information.

These data were analyzed and the results showed that the users' personality was correlated with some aspects of the microblogging behavior, such as Extraversion positively correlated with the frequency of using Weibo, the number of acquaintance among followees and followers, the times of voting weekly as well as the daily times of using the "@" function. And there was a relationship between some microblogging behavior and the demographics, for instance, the frequency of using Weibo was different across the educational levels and occupation. Besides, the effects of some demographical factors on the users' profiles were significant. For example, the amount of followers and Weibo social networking size of the male users was significantly more than that of the female users.

**Acknowledgements.** We would like to thank Professor Tingshao Zhu and Lin Li, Ang Li, and Bibo Hao for their help in collecting data. This work was supported by Chinese Academy of Sciences “Strategic Leading Science and Technology Special Program” (No. XDA06030800).

## References

1. Eldon, E.: ComScore, Quantcast, Compete, Nielsen show a strong December for Facebook traffic in the US (January 22, 2010), <http://www.insidefacebook.com/2010/01/22/comscore-quantcast-compete-show-a-strong-december-forfacebook-traffic-in-the-us/> (accessed January 10, 2010)
2. Barnes, Böhringer: Continuance usage intention in microblogging services: the case of twitter. In: Proceedings of the 17th European Conference on Information Systems, Verona, Italy, June 8-10 (2009)
3. SemioCast: Geolocation analysis of Twitter accounts and tweets (2012), [http://semioCast.com/publications/2012\\_07\\_30\\_Twitter\\_reaches\\_half\\_a\\_billion\\_accounts\\_140m\\_in\\_the\\_US](http://semioCast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US)
4. China Internet Network Information Center: The 27th statistical reports on the Internet development in China 37, (2011), <http://www.cnnic.net.cn/hlwfzyj/hlwzxbg/201101/P020120709345289031187.pdf>
5. China Internet Network Information Center: The 29th statistical reports on the Internet development in China 4-5 (2012), <http://www.cnnic.net.cn/hlwfzyj/hlwzxbg/201201/P020120709345264469680.pdf>
6. Yu, Y., Zhu, Y.: Follow Me: An Analysis of Self-traits, Motivation, Microblog Usage and Attractiveness of One’s Microblog. In: Proceedings of the 2012 9th International Conference on Service Systems and Service Management, pp. 148–153 (2012)
7. John, O.P., Donahue, E.M., Kentle, R.L.: The Big Five Inventory—Versions 4a and 54. University of California Berkeley, Institute of Personality and Social Research, Berkeley, CA (1991)
8. John, O.P., Srivastava, S.: The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspective. In: Pervin, L., John, O.P. (eds.) Handbook of Personality: Theory and Research, 2nd edn. Guilford, New York (1999) (in press)