

Observation on Spammers in Sina Weibo

Yong Xu

Wenhui-xinmin United Press Group,
Shanghai, China
e-mail: xuy@wxjt.com.cn

Yi Zhou*, Kai Chen

Department of Electronic Engineering
Shanghai Jiaotong University, Shanghai, China
Shanghai, China
{zy_21th, kchen}@sjtu.edu.cn
* Corresponding Author

Abstract—Micro-blogging service has been developing and evolving rapidly in China which has led to a significant rise in social spamming attacks. However, little is known about these spammers. Thus, in this paper, we presented an observation on spammers in Sina Weibo, the biggest micro-blogging community in China. Specifically, we used program-controlled profiles to monitor, track and record spamming behaviors. We gave a detailed description of the experiment settings and then analyzed the spamming data collected by these profiles. We found that the spammers on Sina Weibo can be classified into two categories and they shared some distinguishing characteristics. These results are promising for the future study on automatically detecting and identifying spammers.

Keywords—Chinese micro-blog communities, Sina weibo, spammers

I. INTRODUCTION

In the past few years, micro-blog, also known as Weibo, has shown good development prospect ever since its emergence in China. According to a report by CNNIC in January 2012, the number of Weibo accounts has reached 250 million by the end of 2011[9]. However, the popularization of Weibo has also resulted in a great increase in spamming attacks. Successful defense against these spamming attacks is important to improving the quality of service and insuring the security of the users.

The issue of defending social spammers has long raised interest in researchers all over the world. Steve Webb and his partners provided the first characterization of social spammers in MySpace [2] in 2008, while Saptarshi Ghosh did a similar study on Twitter [8]. Moreover, detection methods have been proposed. Some methods focus on filtering spam messages [7] while others aim to filter malicious user profiles [11]. In the honeypot project of Kyumin Lee on Twitter [1, 3, 4, 5] and MySpace [2, 3, 4], machine learning techniques are used to build spammer classifiers. Besides such statistical analysis approaches, some researchers have been trying to apply semantic analysis to social contents [10], which gave a new angle of social spamming detection.

In this paper, we aim to make an observation on spammers in Sina Weibo who owns 57% of Chinese market share [6]. We did a two-month experiment for spamming observation, using program-controlled profiles to track and record spamming behaviors. After that, we carried out an investigation into the collected data. We finally found out

some distinguishing characteristics shared by the spammers in Sina Weibo which were helpful to the future study of detection and identifying spammers.

The rest of the paper will contain: (i) a detailed description of the experiment settings and program-controlled profiles, (ii) the analysis on the data harvested by our profiles and (iii) the conclusions and future plan.

II. OUR METHOD

Spammers on social network can be viewed as malicious profiles publishing spam messages. They post, comment and follow users just like legitimate user. In order to get samples of social spammers for further analysis, we deployed program-controlled profiles to monitor, track and record spamming behaviors.

2.1 Program-controlled Profiles

A program-controlled profile posts Weibo and follows users just like a human user does. It runs without human inspection and stores information of its fans. The followings are some detail of these profiles.

- Various personal information: Each profile has its own personal information different with others' which contains information such as name, birthday, and location.
- Random follow: The profiles scan the public news page and choose some users to follow randomly in order to get more fans and more chances to be followed by spammers.
- Auto post & repost: The profiles post and repost Weibo automatically with the help of a program which scans the web to fetch appropriate messages.
- Topic oriented: We varies the content topics of the profiles. Each profile has its own post theme favor just like a human user.
- Different action patterns: Each profile runs in its own pattern which is determined by several parameters: user number to follow, Weibo number to post and repost and launch time. All the profiles run at different time so as to avoid IP verification by Sina.

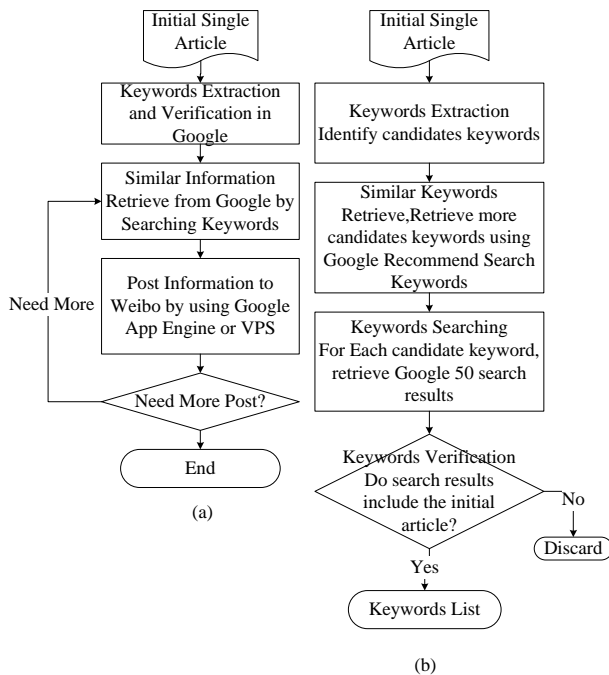


Figure 1. (a) Flow of Auto-Post Message to Weibo. (b) Flow of keywords extraction and verification.

The flow of our auto-post method shown in Fig. 1(a). Given a single initial article, keywords of the initial article are extracted and verified, then similar information are retrieved by searching Google using these keywords. Finally, retrieved results will be post to by artificial identities. To hide real IP address of artificial identities, GAE (Google App Engine) and VPS (Virtual Private Server) will be used. GAE is a platform for developing and hosting web applications in Google-managed data centers. In our method, a proxy-like web application can be hosted in GAE to redirect all commands to target sites. At the same time, Linode VPS is used in our method. Each Linode VPS is a Linux virtual machine with IP address in different regions, such as London in UK, Newark, Dallas, Atlanta and Fremont in US. In current experiment, each VPS hoses 50 identities. In the future, due to the excellent extensibility of VPS, more artificial identities can be hosted easily while difficult to be detected. Both GAE and VPS can hide the real IP address of artificial identities.

The flow of keywords extraction and verification is shown in Fig. 1(b). Firstly, candidate keywords of the initial article are extracted. Secondly, more similar keywords are retrieved by Google recommend search keywords. Thirdly, final keywords list will be identify by checking if search results using this keyword include the initial article.

2.2 Experiment Settings

8 program-controlled profiles were deployed and each of them ran four times a day. We set up a series of scheduled tasks to launch the programs automatically. The whole experiment last for 64 days from 20th Oct. 2011 to 23th Dec. 2011.

Status of the profiles was recorded by the program and stored to a database, for example, Weibo number posted, user number followed, fans number, friend number and some information of the fans such as user name and homepage address. We collected the information for a further analysis on attractiveness of profiles and the spammer features.

III. EXPERIMENT RESULTS AND ANALYSIS

3.1 Attractiveness of Program-controlled Profiles

The relationship status of the profiles and some more detailed information is shown in table 1 with fans number denoting the attractiveness of the profiles.

Table 1. Data of the experiment results

	Post	Follow	Fans	Friends	Post Per Fan	Non-friend Fans	Non-friend Fans Rate
1	1582	843	163	97	9.7	66	40.5%
2	1919	601	80	53	24.0	27	33.8%
3	1052	645	84	58	12.5	26	31.0%
4	1950	646	96	57	20.3	39	40.6%
5	1072	357	51	41	21.0	10	19.6%
6	1822	517	69	43	26.4	26	41.3%
7	4349	517	61	30	71.3	31	50.8%
8	2458	523	77	42	31.9	35	45.5%

From the above table, we made the following conclusions:

(i)Difference in topics leads to difference in attractiveness, which can be seen from the average Weibo posted to get a fan. (ii)A quite large part of users will follow strangers who follow them (seen in the response rate data).This can also be exploited by the spammers to get more targets. (iii)The attractiveness of a profile can be improved by choosing proper topics and being more active.

3.2 Spammers on Sina Weibo

We investigated the fans attracted by the profiles, picked malicious users out of the legitimate ones and divided them into the following two classes.

- Promoters: Promoters are those who publish advertising information for commercial purposes. They can be divided into two sub-classes. The first kind of promoters includes advertisements in their personal information such as portrait, self-introduction and blog address field. One distinguishing characteristic of these promoters is that they don't publish any Weibo because all the ad-info is in their personal information part. The other kind of promoters publishes similar or even identical Weibo containing URLs frequently. They also forward other promoters' Weibo with comments that have nothing to do with the original contents.

- Robot accounts: Evidences are found that this kind of accounts are programed and run automatically. They usually publish Weibo of several fixed topics with a fixed test format (title+content+url+picture). They tend to publish posts according to a preset temporal pattern such as one post every 20 minutes. The robot accounts seldom

publish spamming posts directly. They are more likely to forward spamming posts published by other users.

One distinguishing characteristic shared by the two kinds of spammers is that they have abnormal relationship status. They follow a huge number of users (usually more than 1500) and have a relatively low fans number (at most 800 with the following of more than 3000). This is because they are trying to attract fans by following and waiting for responses so as to expand their spamming range. Robot accounts usually have more fans than promoters because they provide useful contents and publish fewer advertisements. Moreover, there is usually no relationship between the fans of a spammer.

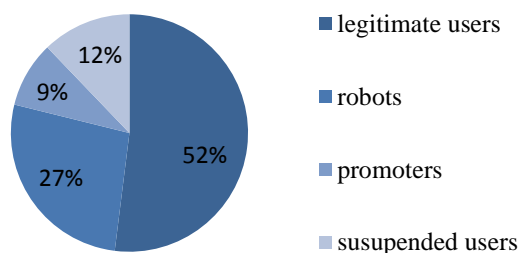


Fig. 2. The ration of different kinds of users harvested on Sina Weibo

The ratio of all kinds of users in our sample on weibo.com is shown in Figure 2 as above. Nearly half of the harvested users are malicious ones and only 12% have been detected and suspended by Sina. More robots exist than promoters because they are more difficult to detect and perhaps offend legitimate users less than promoters do.

Examples of spammers are shown in Figure 3. Spammer (a) contains ad-information in its portrait; spammer (b) publishes ad-Weibo; and spammer (c) is an example of robot accounts.



Fig. 3. Examples of spammers in Sina Weibo. (a) and (b) are promoters while (c) is a robot.

One distinguishing characteristic shared by the two kinds of spammers is that they have abnormal relationship status. They follow a huge number of users and have a relatively low fans number. This is because they are trying to attract fans by following and waiting for responses so as to expand

their spamming range. Moreover, there is usually no relationship between the fans of a spammer.

3.3 Useful Features for Detection

The experiments and analysis above have shown a brief picture of spammers on Sina Weibo. We finally listed several features that may be useful to distinguishing spammers from legitimate users.

- Follow and friend number: As was mentioned above, spammers usually follow a large number of users and have a relatively small fans number. The gap between follow and fans of a spammer tends to be bigger than that of a legitimate user.
- Friends relationship: Most spammers follow users randomly so most of their fans who establish relationship by responding their following don't know the other fans, while a legitimate user's friends tends form one or more groups whose members know each other.
- Personal information content: It's sometimes used for spamming. By checking the personal introduction and other information such as user portrait, we can tell spammers from legitimate users.
- Post content information: The similarity and context information of posts published can help a lot in detecting spammers. By statistical or even semantic analysis, we may find out spammers.
- URL information: URLs play a vital role in social network spamming for almost all spam messages contain a URL. The number and destinations of URLs in weibos can help a lot in detecting a spammer.

IV. CONCLUSION AND FUTURE WORKS

In this paper, we made an observation on spammers in Sina Weibo. We set up a two-month experiment, using program-controlled profiles to monitor and track spammers. By investigating the collected data and summarizing the characteristics of spammers on Sina Weibo, we made the following conclusions:

- It is an effective way to monitor, track and record social spamming behaviors with the help of program-controlled profiles.
- There are two kinds of spammers in Sina weibo: promoters and robots both of which showed some distinguishing characteristics which cannot be seen among legitimate users.
- Several features may be useful for the detection of spammers: Friendship status, post content and URL features.

The next step of our study would be focused on the identifying and classifying of malicious users or spammers on Chinese social network communities. We would like to upgrade the scale of our experiment so as to get more samples. Extended study on spamming features will be carried out and a statistical analysis is needed. We would like to find out an efficient way to judge whether an unknown user is a spammer.

ACKNOWLEDGMENT

The work is partially supported by the Shanghai Science and Technology Committees of Scientific Research Project (Grant No. 11dz1505502), and Key Lab of Information Network Security and Shanghai Science and Technology Committees of Scientific Research Project (Grant No. 11dz1505502).

REFERENCES

- [1] Kyumin Lee, Brian David Eoff, James Caverlee.: Devils, Angels, and Robots: Tempting Destructive Users in Social Media. In: AAAI Conference on Weblogs and Social Media ,2010.
- [2] Steve Webb, James Carverlee, Calton Pu.: Social Honey pots: Making Friends with A Spammer around You. In: 5th Conference on Email and Anti-Spam (CEAS), 2008
- [3] Kyumin Lee, James Caverlee, Steve Webb.: The Social Honey pot: Protecting Online Communities from Spammers. In: WWW 2010, April 26-30, 2010, Raleigh, North Carolina, USA.
- [4] Kyumin Lee, James Caverlee, Steve Webb.: Uncovering Social Spammers: Social Honey pots + Machine Learning. In: SIGIR'10, July 19-23, 2010, Geneva, Switzerland.
- [5] Kyumin Lee, Brian David Eoff, James Caverlee.: Seven Months with the Devils- A Long-term Study of Content Polluters on Twitter. In: AAAI 2011, 2011
- [6] Digimind.: How to 'Tweet' in China: Twitter vs Sina Weibo in Figures. 2011
- [7] Xin Jin, Cindy Xide Lin, Jiebo Luo, Jiawei Han.: A Data Mining-based Spam Detections System for Social Media Networks. In: Proceedings of the VLDB Endowment, 2011
- [8] Saptarshi Ghosh, Gautam Korlam, Niloy Ganguly.: Spammers' Networks within Online Social Networks: A Case-Study on Twitter. In: 20th World Wide Web Conference, 2011.
- [9] CNNIC.: The 29th Statistical Report on China Internet Development. January 16th, 2012.
- [10] Fabian Abel, Qi Gao, Geert-Jan Houben, Ke Tao.: Semantic Enrichment of Twitter Posts for User Profile Construction on the Social Web. In: ESWC'11 Proceedings of the 8th extended semantic web conference on The semantic web: research and applications.
- [11] Danesh Irani, Steve Webb, and Calton Pu.: Study of Static Classification of Social Spam Profiles in Myspace. In: proceedings of the fourth International AAAI Conference on Weblogs and Social Media. 2010.