ZLOC: Detection of Zombie Users in Online Social Networks Using Location Information

Jing Deng, Lixin Fu, Yi Yang

Department of Computer Science
University of North Carolina at Greensboro
Greensboro, NC 27410, U.S.A.
Email: jing.deng@uncg.edu, lfu@uncg.edu, y_yang6@uncg.edu

***Abstract*—Online social networks serve as a promising platform for social eliteness and financial gain. With such a promise or dream, zombie accounts, behind which stand no real users, become prevalent. The detection of such accounts has been games of cat and mouse, with more and more sophisticated methods used by zombie account managers. In this work, we propose a new zombie account detection technique called ZLOC, Zombie detection using Location information. ZLOC uses the location information of an account's friends or followers. More specifically, we investigate the follower accounts of suspected zombie accounts in SINA WeiBo, one of the two most popular microblogging websites in China. Our scheme is based on a natural social behavior that many of one's friends are usually in the vicinity of his/her location. Our analysis shows that the proposed ZLOC scheme has some salient features that help with zombie detection.**

Keywords–Online Social Networks; Zombie Detection; Fake Users; Detection Accuracy.

I. Introduction

With the prevalent usage of Internet, online social networks have quickly become the center of human interactions. Younger generations, as well as the old ones, use different kinds of online tools to communicate and to obtain information. Web pages supported by HTML 5.0 can be very dynamic and have various features, but Facebook, Twitter [1][2], LinkedIn, to name a few, are the place to make friends, share information or news, or establish professional connections.

For any online website with a large number of users, there also come the zombie accounts, or the faked users. Different from legitimate users who have real people behind to communicate, a software or a zombie manager may be controlling these accounts. The reasons for the existence of such zombie accounts can vary. They range from financial exploitation to phantom fame purposes. For instance, in most of daily deal websites, each deal is usually ranked by the number of thumbs-up and thumbs-down. With the control of a large number of zombie accounts, it is then possible to control the hotness of a deal by directing some of the zombie accounts to thumb up a particular deal (by a sponsored company or website) and/or thumb down another deal (maybe by a different company or website). Some online advertisement companies may exploit the effectiveness on the popularity of certain accounts to achieve financial gains. Even worse, the spammers may even steal information from profiles on network or direct users to phishing websites.

It is thus essential to detect such zombie accounts from online social networks. Unfortunately, the task can be hindered by the adaptive behavior of the zombie managers (the people controlling the zombie accounts). It is further complicated by the sometimes less than normal active behavior pattern by the real user accounts: the difference of an abandoned account and an occasionally used zombie account can just be tiny. Nevertheless, zombie detection has seen some interesting progresses in recent years, with techniques ranging from simply checking number of friends, to number of active posts or activity, to more sophisticated statistical analysis among different herds of users [3][4]. Interestingly, none of these prior arts has investigated the use of location information of the user accounts.

In this work, we focus on the detection of zombie accounts in SINA WeiBo, one of the two most popular twitter-like microblogging websites in China. Similar to Twitter, WeiBo allows each user to follow a number of other users. Since any real person is unlikely to be able to read large amount of feeds, it posts a limit of 2000 follows for each user. The number of followings, however, can be as large as millions, depending obviously on the popularity of the accounts.

Our work is based on the following natural social bond observation. Most people interact with their friends or relatives who live within the same city or the same province (state) where they live. While anyone may have friends or followers from other cities or provinces (states), the ratio of followers who are in the same city or province (state) should be at least higher than a certain threshold. Instead, zombie accounts tend to have followers from a very diversed geographical locations. Our proposed scheme ZLOC, Zombie detection based on Location information, is based on such an salient observation.

Our paper is organized as follows. In Section II, we discuss prior arts and their differences with our proposed scheme. We illustrate and explain our ZLOC scheme in Section III, followed by the performance evaluation in Section IV. In Section V, we conclude our work and point out several directions for future work.

II. Related Work

There have been some work in fake followers for online social networks in the technical literature. We review them in the following:

Rumors can propagate easily in social networks. Sun et al. [5] proposed an effective rumor classifier that categorized rumors into four types, one of which was text-picture unmatched event rumors. They designed special features that may be used to build a classifier to differentiate rumors from ordinary posts. Thomas et al. [1] researched on suspended Twitter accounts to find their lifetime events and behavior. McCord and Chuah [6] presented a detection method with user-based and content-based features and applied traditional classifiers such as Random Forest, Naive Bayesian, Support Vector Machine, and K-nearest neighbors.

In social networks, there are spammers, the detection of which is essential and can impact real users [7]. Hu et al. gave a unified model for detecting social spammers in microblogging by integrating both social network information and content information [8]. Lee et al. [9] deployed social honeyspots to harvest suspicious spam profiles and then classified them using machine learning. Lin et al. [10] collected a set of spammer samples by proactive honeypots and keyword based searching, and designed an online system for identifying spammers. They found three abnormal behaviors of the spammers: aggressive advertising, repeated reposting, and aggerssive following. Chu et al. mainly focussed on the detection of large-scale spam campaigns on Twitter rather than screening individual tweets [11]. They clustered the collected dataset of 20 million tweets into different campaigns according to their same final URLs. They presented a classification system based on a set of features generated from campaign data.

The detection of non-real users in social networks has shown to be tricky because of their evasive and ever-changing behaviors. Shen et al. [12] proposed a binary classifier to detect fake followers by their extracted major features in SINA WeiBo and presented their classifier's performance. Guo et al.[13] collected more than 20 million profiles of users and researched their posting behaviors. Marionette users like puppets are fabricated for fake popularity or financial gain. Wu et al. integrated both individual user tweeting behavior information and the social interactions among users to develop a semi-supervised probabilistic model in order to distinguish marionettes from normal users [14]. Due to lack of user-generated contents, it is difficult to capture the profiles of lurking users. Zhang et al. [15] presented a unified social context graph model and an algorithm to generate profiles of the lurking users to effectively detect them. Wang and Lu [3] introduced a star sampling method by taking all the neighbors as valid samples. They used it to identify ten thousands of top bloggers on Weibo. To analyze Twitter sphere, Black et al. [2] proposed an elegant architecture to perform Twitter studies. Jiang et al. [4] proposed CATCHSYNC that used and measured two suspicious behaviors: the first measure is "sync" behavior of zombies, that is, they often have similar behavior; another is "norm", that is, their behavior is different from other normal users.

Armed with a very large dataset with 54 million users and 1.8 billion tweets and a manually labeled collection of 1,065 users, Benevenuto et al. [16] carefully examined a large set of features, such as fraction of tweets with URLs, hashtags, and spam words, number of replies, number of followees and followers, account life time, number of tweets received, etc., to differentiate spammers from normal users. They also used Support Vector Machine and Chi-square method to classify and characterize the spammers. Liu et al. proposed ProZombie [17], a two-stage cascading model for detecting zombies. They also came up with new features to give a more refined description of Weibo users, improving the modeling efficiency without loss of accuracy. Zombies are essentially the same as Socialbots or Sybil accounts, which have received attention from the perspective of Turing machine/human classification [18][19][20][21].

TABLE I. Variables.

SAMEP	Ratio of followers who share the same province (state) with the user
SAMEC	Ratio of followers who share the same province (state) and city with the user
FER	Number of followers of a user
FING	Number of users that a user is following

To conclude, while quite some work has been performed on zombie detection in online social networks, the detection of such still remains inaccurate and/or requires too much extra information. None of the above prior arts considered user registration location specifically for zombie detection, although some used similar features in the big picture of user classification [11][17]. Instead, our work focuses on the use of such location information make accurate detection decisions.

III. The ZLOC Scheme

We describe the ZLOC scheme in this section. The ZLOC scheme requires the follower information of a suspected account. Once the list of all followers is obtained, the registration information of each of the followers will be retrieved through a simple web access. Such registration location information is then compared with the registration location information (such as Guangzhou, Guangdong). Then the following two numbers are generated: the number of followers of the suspected account having the same city and province (state) information as the suspected account, denoted as SAMEC; and the number of followers of the suspected account having the same province (state) information as the suspected account, denoted as SAMEP. Note that it is obviously true that SAMEC < SAMEP. We present all variables on Table I, as add "_TH" to denote the threshold in a comparison, e.g., FER_TH is the threshold for FER.

These two numbers are then compared to two thresholds, SAMEC_TH and SAMEP_TH, respectively. When a suspected account satisfies SAMEC < SAMEC_TH and SAMEP < SAMEP_TH at the same time, it is considered a zombie based on location information.

Due to the fact that some users may register under locations different from their real residential location, other perspectives should be checked to improve the detection accuracy. A simple and low-cost perspective is the number of followers, defined as FER, and the number of followings, defined as FING. At least in the early days, zombie accounts made large number of follows so that they might get followed back. Therefore, they usually have close-to limit FING numbers. On the contrary, it is hard to find real user accounts to follow them back, except that the zombie accounts can be directed to follow themselves (orchestrated by one zombie manager or even several zombie managers). It is thus easy to see that any real user account should have FING lower than a threshold FING_TH unless his/her FER number is greater than a threshold FER_TH (such as some highly popular accounts). Hence, the following additional detection rule:

When FER < FER_TH and FING > FING_TH, a suspected account is considered a zombie.

The reason to include a threshold for the number of followers is that some accounts may have so many followers and the users may prefer to follow (or counter-follow) a certain percentage of these followers. Therefore, it is possible that these real user accounts may have larger numbers of followings.

In fact, we use the above basic detection rule, termed FER-FING, to demonstrate the effectiveness of our location-based zombie detection strategy in Section IV. The basic FER-FING scheme is

$$\begin{aligned} &C1 : \text{FER} < \text{FER\_TH} \\ &C2 : \text{FING} > \text{FING\_TH} \\ &\text{Rule: } C1 \cdot C2 \text{==TRUE} \rightarrow \text{zombie} \end{aligned} \quad (1)$$

We describe the ZLOC scheme in the following:

$$\begin{aligned} &C1 : \text{FER} < \text{FER\_TH} \\ &C2 : \text{FING} > \text{FING\_TH} \\ &C3 : \text{SAMEP} < \text{SAMEP\_TH} \\ &C4 : \text{SAMEC} < \text{SAMEC\_TH} \\ &\text{Rule: } C1 \cdot (C2 + C3 \cdot C4) \text{==TRUE} \rightarrow \text{zombie} \end{aligned} \quad (2)$$

IV. Performance Evaluation

In this section, we present the performance evaluation of our ZLOC scheme.

A. Data Collection and Performance Metrics

First, we explain how we obtained our dataset. The data was collected through a web crawler that started with a real user account and found the followers and the followers' followers, and so on. For each user, we retrieved the list of all followers that are available from the WeiBo webpage. Our crawler stopped when the number of accounts reached 10,000. For all of these 10,000 accounts and all their known followers, we retrieved their registration location information (together with other information that we did not use, such as registration time/date and last post time/content).

Note that, in SINA WeiBo, only about 200 followers are now disclosed to anyone other than the user himself/herself. This might have been posted for privacy reason. It has an interesting effect on our evaluation. First of all, such a limit means that the follower list that ZLOC processes and makes decision upon is incomplete. Thus, the accuracy of zombie detection can be questionable. However, we argue that such a large snapshot of the follower list is already quite revealing, as demonstrated by our results. Secondly, it also affected the list of users that we crawled in a way that might have changed the ratio of zombie/real accounts in the dataset. Since the goal here is to evaluate how accurate the ZLOC scheme is, the ratio of the dataset does not actually matter.

The 10,000 accounts were then passed through the ZLOC scheme and the basic FER-FING scheme. A decision of either zombie or real account would be reached at the end. We randomly sampled more than 100 accounts and checked (through human reading) whether they were really zombies or real user accounts. We evaluated these two schemes in the following performance metrics (all based on the sampled pool):

Successful Detection Ratio: This is the ratio of the number of zombie accounts that are detected as zombie users divided by the number of all zombie accounts.

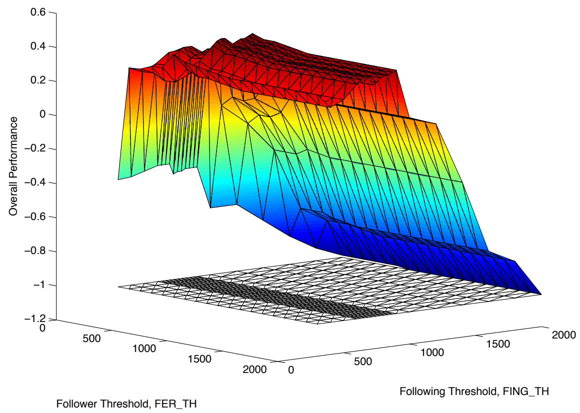

Figure 1. Overall performance of the basic FER-FING scheme with different FER_TH and FING_TH values.

Missed Detection Ratio: This is the ratio of the number of zombie accounts that are detected as real user accounts divided by the number of all zombie accounts.

False Alarm Ratio: This is the ratio of the number of real user accounts that are detected as zombie accounts divided by the number of all real accounts.

Overall Performance: This is computed as Successful Detection Ratio minus Missed Detection Ratio as well as False Alarm Ratio.

Among these metrics, the last calls for some explanations. In any classical detection problem, it is rather easy to increase successful detection ratio while ignoring missed detection and false alarm ratios, or vice versa. A practical scheme should indeed balance all three. In fact, different weights (positive or negative) for such ratios can be applied to these ratios and one can try to maximize the combined return. In this work, we choose the simple subtraction as the final return and leave more complex return weight to our future work.

B. FER_TH and FING_TH Selection

First of all, we need to choose the best FER_TH and FING_TH for the basic FER-FING scheme as well as our ZLOC scheme.

We present our investigation of the basic FER-FING scheme for its best FER_TH and FING_TH values (see Figure 1). When these thresholds are too low or too high, the overall performance of the FER-FING scheme is rather low. When they are in the range of 700-900, the basic FER-FING scheme works the best, at least for the data points that we sampled. Therefore, we will choose FER_TH and FING_TH values as 700 and 900, respectively. All simulations in Section IV-C are based on these values. Note that other similar parameters would produce similar comparisons, as shown below.

C. Accuracy of ZLOC

First of all, we plotted the SAMEP and SAMEC values of all sampled accounts (hence, we have manually checked whether they are zombie accounts or real user accounts) in Figure 2. All zombie accounts are represented by a red circle on its (SAMEP, SAMEC) position. Note that we have added a

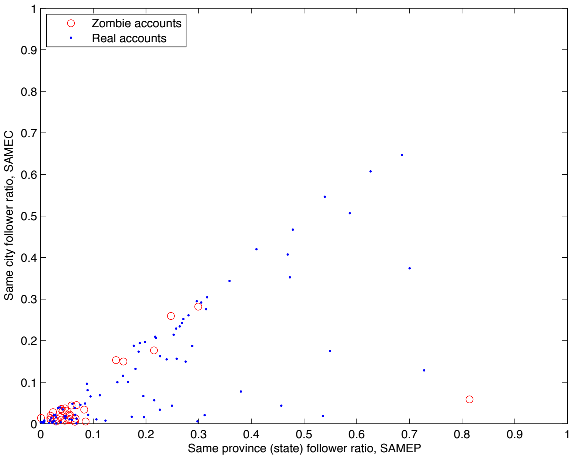

Figure 2. Distribution of sampled accounts. Here we plot sampled accounts on a 2-D surface using their SAMEP and SAMEC values. Zombie accounts and real user accounts are distinguished by the different symbols.

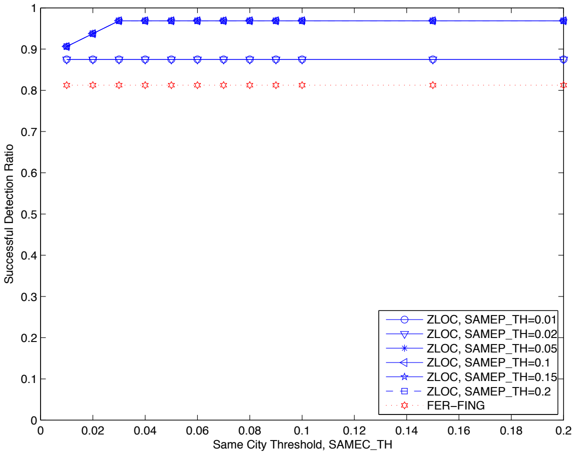

Figure 3. Successful detection ratio for the FER-FING scheme and ZLOC scheme.

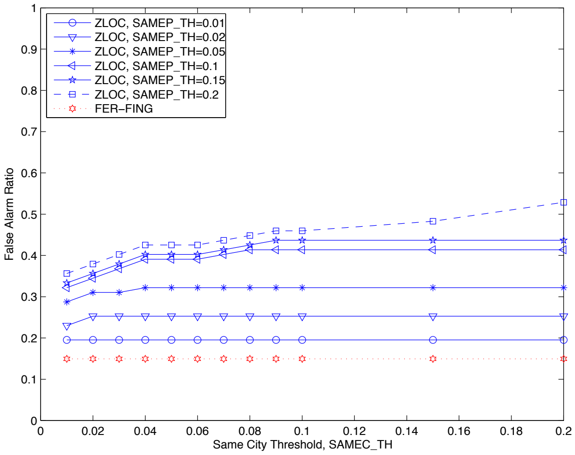

Figure 4. False alarm ratio for the FER-FING scheme and ZLOC scheme.

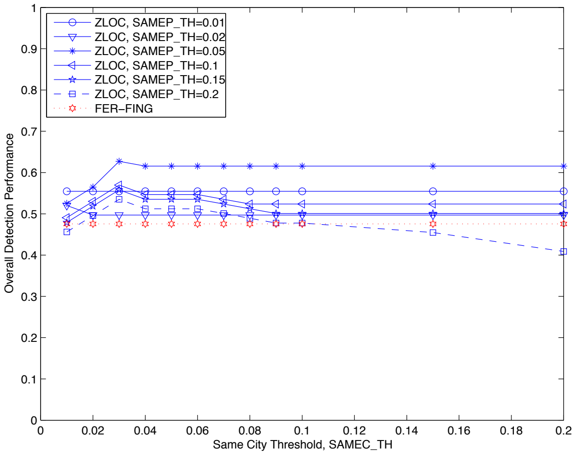

Figure 5. Overall performance comparison of the FER-FING scheme and ZLOC scheme.

slight perturbation for each point in order to show those at the exact same locations, which were caused by the small numbers of followers and hence, the same SAMEP values and the same SAMEC values.

From Figure 2, the pattern of zombie users is clearly demonstrated as most of them stay on the lower left corner of the region, except for a few data points. On the other hand, real user accounts are more diverse and vastly spread. Such an observation has served as the inspiration for our work. Also note that all data points satisfy SAMEC $<$ SAMEP.

We present the successful detection ratio in Figure 3. It can be seen that the ZLOC scheme generally has better successful detection ratios than the basic FER-FING scheme. Within the ZLOC scheme, a very small SAMEP_TH is rather ineffective. When SAMEP_TH reaches 0.1 and 0.2, however, the successful detection ratio remains the same. It might have been caused by the fact that none of the users have more than 10% of their followers within the same province (state). Except for SAMEP_TH=0.01, the successful detection ratio shows slight increases as SAMEC_TH increases, as more and more accounts become eligible to be declared as zombies.

The performance in false alarm ratio is shown in Figure 4. As SAMEP_TH increases, the false alarm ratio increases as more and more accounts fall within the detection threshold. Similarly, the increase of SAMEC_TH also raises false alarm ratio. In general, the basic FER-FING scheme has a lower false alarm ratio than the ZLOC scheme.

We compare the overall performance of the ZLOC scheme and the basic FER-FING scheme in Figure 5. The ZLOC scheme outperforms the basic FER-FING scheme except in a few places where the SAMEC_TH value is set to be too large. For several of the SAMEP_TH lines, an interesting up and down trend can be observed as SAMEC_TH increases, suggesting an optimum choice for SAMEC_TH. This is because of the joint impact of successful detection, false alarm, as well as missed detection. Overall, the results in Figure 5 suggest the best SAMEP_TH and SAMEC_TH to be 0.05 and 0.03, which are the parameters that we use in Section IV-D.

D. Zombie Ratios

Lastly, we present the ratio of accounts in our dataset that these two schemes detect as zombies. Among these 10,000 accounts, ZLOC detected 5,300 of them as zombies and the basic FER-FING scheme found 3,300 as zombies. The difference is significant, underlining the impact of our location-based detection approach.

V. Conclusions and Future Work

We have presented a new zombie detection scheme called ZLOC. The ZLOC scheme takes advantage of the fact that the location vicinity between a follower and the person whom he/she is following. ZLOC then compares the ratio of such followers of any suspected account. If the ratio is below a certain threshold, the account is more likely to be a zombie. Through our simulation studies, we have found that ZLOC could use two thresholds, one to compare with the ratio of the followers within the same city and the other to compare with the ratio of the followers within the same province (or state). With such additional ratios, the ZLOC scheme has been demonstrated to raise the successful detection of zombie accounts significantly. In addition, the overall performance, defined by successful detection minus false alarm ratio, as well as missed detection ratio, is also higher than other schemes.

In terms of applications, we believe that ZLOC can be used in combination with other techniques, such as those in [12], to further improve the accuracy of zombie user detections. Our scheme does not rely on detailed information of the suspicious user such as posting habits, timing, contents, etc., making it a great candidate for efficient detection.

We note that it is possible to adapt our ZLOC approach to detect phantom users in other online social networks such as Facebook, LinkedIn, and/or Twitter. For instance, in Twitter, many accounts show their current residential locations. Their followers can be checked as well and our ZLOC approach can be applied for detection. The story would be slightly different in Facebook, mainly because all connections are bi-directional instead of the directional following in microblogging websites. Instead of checking followers' location information, perhaps the location statistics of all friends of one account can be checked, although such friends are usually not viewable from a third party unless the privacy setting allows so.

In our future work, we will investigate the use of actual vicinity instead of hard-coded same city ratio. Therefore, followers in the neighboring cities will still be considered as close-by. The investigation of larger datasets from other social networks will be helpful as well.

Acknowledgment

J. Deng would like to thank the financial support from NSF, grant No. CCF-1320428.

References

[1] K. Thomas, C. Grier, D. Song, and V. Paxson, “Suspended accounts in retrospect: An analysis of twitter spam,” in Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, ser. IMC ’11. New York, NY, USA: ACM, 2011, pp. 243–258.

[2] A. Black, C. Mascaro, M. Gallagher, and S. P. Goggins, “Twitter zombie: Architecture for capturing, socially transforming and analyzing the twittersphere,” in Proceedings of the 17th ACM International Conference on Supporting Group Work, ser. GROUP ’12. New York, NY, USA: ACM, 2012, pp. 229–238.

[3] H. Wang and J. Lu, “Detect inflated follower numbers in osn using star sampling,” in Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ser. ASONAM ’13. New York, NY, USA: ACM, 2013, pp. 127–133.

[4] M. Jiang, P. Cui, A. Beutel, C. Faloutsos, and S. Yang, “Detecting suspicious following behavior in multimillion-node social networks,” in Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion, ser. WWW Companion ’14. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2014, pp. 305–306.

[5] S. Sun, H. Liu, J. He, and X. Du, “Detecting event rumors on sina weibo automatically,” in Web Technologies and Applications, ser. Lecture Notes in Computer Science, Y. Ishikawa, J. Li, W. Wang, R. Zhang, and W. Zhang, Eds. Springer Berlin Heidelberg, 2013, vol. 7808, pp. 120–131.

[6] M. McCord and M. Chuah, “Spam detection on twitter using traditional classifiers,” in Autonomic and Trusted Computing, ser. Lecture Notes in Computer Science, J. Calero, L. Yang, F. Mrmol, L. Garca Villalba, A. Li, and Y. Wang, Eds. Springer Berlin Heidelberg, 2011, vol. 6906, pp. 175–186.

[7] C. Sibona and S. Walczak, “Unfriending on facebook: Friend request and online/offline behavior analysis,” in Proc. on System Sciences (HICSS), the 44th IEEE Hawaii International Conference, January 2011, pp. 1–10.

[8] X. Hu, J. Tang, Y. Zhang, and H. Liu, “Social spammer detection in microblogging,” in 23rd International Joint Conference on Artificial Intelligence (IJCAI ’13), August 3-9 2013.

[9] K. Lee, J. Caverlee, and S. Webb, “Uncovering social spammers: social honeypots+ machine learning,” in Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval, July 19-23 2010, pp. 435–442.

[10] C. Lin, J. He, Y. Zhou, X. Yang, K. Chen, and L. Song, “Analysis and identification of spamming behaviors in sina weibo microblog,” in Proceedings of the 7th Workshop on Social Network Mining and Analysis, vol. 5, 2013, pp. 1–9.

[11] Z. Chu, I. Widjaja, and H. Wang, “Detecting social spam campaigns on twitter,” in Applied Cryptography and Network Security, ser. Lecture Notes in Computer Science, F. Bao, P. Samarati, and J. Zhou, Eds. Springer Berlin Heidelberg, 2012, vol. 7341, pp. 455–472.

[12] Y. Shen, J. Yu, K. Dong, and K. Nan, “Automatic fake followers detection in chinese micro-blogging system,” in Advances in Knowledge Discovery and Data Mining, ser. Lecture Notes in Computer Science, V. Tseng, T. Ho, Z.-H. Zhou, A. Chen, and H.-Y. Kao, Eds. Springer International Publishing, 2014, vol. 8444, pp. 596–607.

[13] Z. Guo, Z. Li, H. Tu, and L. Li, “Characterizing user behavior in weibo,” in Mobile, Ubiquitous, and Intelligent Computing (MUSIC), 2012 Third FTRA International Conference on, June 2012, pp. 60–65.

[14] X. Wu, Z. Feng, W. Fan, J. Gao, and Y. Yu, “Detecting marionette microblog users for improved information credibility,” in Machine Learning and Knowledge Discovery in Databases, ser. Lecture Notes in Computer Science, H. Blockeel, K. Kersting, S. Nijssen, and F. Zelezny, Eds. Springer Berlin Heidelberg, 2013, vol. 8190, pp. 483–498.

[15] Z. Zhang, B. Zhao, W. Qian, and A. Zhou, “Generating profiles for a lurking user by its followees’ social context in microblogs,” in Web Information Systems and Applications Conference (WISA), 2012 Ninth, Nov 2012, pp. 135–140.

[16] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, “Detecting spammers on twitter,” in Collaboration, electronic messaging, anti-abuse and spam conference (CEAS), vol. 6, 2010, pp. 1–9.

[17] H. Liu, Y. Zhang, H. Lin, J. Wu, Z. Wu, and X. Zhang, “How many zombies around you?” in Proc. on Data Mining (ICDM), IEEE 13th International Conference on, Dec 2013, pp. 1133–1138.

[18] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, “The socialbot network: When bots socialize for fame and money,” in Proceedings of the 27th Annual Computer Security Applications Conference, ser. ACSAC ’11. New York, NY, USA: ACM, 2011, pp. 93–102.

[19] T. Hwang, I. Pearce, and M. Nanis, “Socialbots: Voices from the fronts,” interactions, vol. 19, no. 2, Mar. 2012, pp. 38–45.

[20] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, “The rise of social bots,” CoRR, vol. abs/1407.5225, 2014, pp. 1–11.

[21] Z. Yang, C. Wilson, X. Wang, T. Gao, B. Y. Zhao, and Y. Dai, “Uncovering social network sybils in the wild,” ACM Transactions on Knowledge Discovery from Data, vol. 8, no. 1, Feb. 2014, pp. 2:1–2:29.