Toward Efficient Spammers Gathering in Twitter Social Networks

Yihe Zhang  
University of Louisiana at Lafayette,  
Lafayette, LA, USA  
Hao Zhang  
Oracle Corp.,  
Redwood Shores, CA, USA  
Xu Yuan  
University of Louisiana at Lafayette,  
Lafayette, LA, USA

ABSTRACT
This paper introduces a novel system, named pseudo-honeypot, for efficient spammers gathering. Different from the manual setup in the honeypot, the pseudo-honeypot takes advantage of Twitter users’ diversity and selects accounts with the attributes of having the higher potentials of attracting spammers, as the parasitic bodies. By harnessing a set of normal accounts possessing these attributes and monitoring their streaming posts and behavioral patterns, the pseudo-honeypot can gather the tweets that are far more likely of including spammer activities, while removing the risks of being recognized by smart spammers. It substantially advances the honeypot-based solutions in attribute availability, deployment flexibility, network scalability, and system portability. We present the system design and implementation of pseudo-honeypot (including node selection, monitoring, feature extraction, and learning-based classification) in Twitter networks. Through experiments, we demonstrate its effectiveness in term of spammer gathering.

CCS CONCEPTS
• Information systems → Spam detection.

KEYWORDS
Spammers gathering; Honeypot; Pseudo-honeypot

ACM Reference Format:

1 INTRODUCTION
Spammers have been pervasively annoying normal users since the inception of online social networks. By creating fake accounts or comprising benign ones, spammers can initialize social relationships or send unsolicited requests/messages, and then spread harm to the social users. There has been some efforts [1, 7] on exploiting the effective mechanisms to capture spammers. However, most of existing solutions are time-consuming and low efficient as they filter spams (or spammers) from a large set of blindly collected network contents (or accounts), but can only detect a small portion of spammers. The honeypot is proposed as an promising solution by manually creating accounts as lures for spammers [3, 4, 6, 8]. However, such category of solutions has the essential drawbacks on deployment flexibility, attribute variability, and network scalability, as it involves the considerable human overhead.

In this paper, we aim to propose a novel system, called pseudo-honeypot, for efficient spammer gathering. Instead of creating artificial accounts commonly used in honeypot-based solutions, we allow the pseudo-honeypot to take a normal user as the parasitic body and harness such a user to monitor spammer activities. Through this way, pseudo-honeypot can take advantage of users’ diversity in Twitter social networks and utilize them as the key resources to attract spammers. By carefully identifying the attributes that are more likely of attracting spammers, we can select a set of accounts that possess such attributes to be served as the pseudo-honeypot nodes. Pseudo-honeypot can harness these users and monitor their streaming posts/behavioral patterns transparently. From this point of view, the pseudo-honeypot can perform the similar functions as the honeypot in terms of attracting and trapping spammers but has much richer diversity. The existing Twitter’s API enables us to harness the selected pseudo-honeypot and monitor the associated accounts transparently, whereas Twitter API’s privacy policy terms will be strictly followed by us. Subsequently, we extract the features that can reflect spammers’ behaviors and leverage the machine learning classifier to identify each collected tweet as spam or non-spam intelligently and automatically. To validate the performance of our system, we create a pseudo-honeypot network with 1000 nodes and deploy them into Twitter social networks for spammer capturing. During a 500-hour experiment, we collect 1,227,708 tweets posted by 476,345 unique users which count as 1.3% of total active Twitter users, in which we capture a total of 371,981 spams and 53,857 spammers.

2 SYSTEM DESIGN OF PSEUDO-HONEYPOT
Our goal is to develop an efficient and effective mechanism to collect Twitter social network contents that are far more likely of including spammers’ activities so as to detect and remove them. As user diversity provides abundant resources, we propose a novel framework called pseudo-honeypot by utilizing such diversity to attract and trap spammers. Our solution screens a set of normal users that possess attributes of meeting spammer’s tastes. By constructing on the top of these users that may suffer spambly behaviors, pseudo-honeypots can collect the tweets that have a higher probability of including spammers’ activities. Notably, to comply with the Twitter privacy policy, the pseudo-honeypot is not allowed to affect the activities of normal users or make any change to their posts or behavioral patterns. By harnessing these accounts, pseudo-honeypot can monitor public streaming posts and behavioral patterns crossing these accounts in a real-time manner while keep invisible to
we use hashtag-based and trending-based attributes to select a set (e.g., Streaming API) that are available for developers enable the spammers. Here, we take the reverse engineering strategy. That is, we first select a large pool of prevalent attributes that have been widely studied in previous research and use them to select accounts serving as pseudo-honeypot nodes while creating a sampling pseudo-honeypot system. After running it for a period of time, we examine all attributes and refine the top ones that have the highest probability of capturing spammers. Then, we utilize the refined attributes to advise our design of a highly effective pseudo-honeypot system. As a start point, we mainly concentrate on a set of prevalent hashtag-based and trending-based attributes [5, 8], i.e., entertainment, business, tech, education, environment, social, trending-up topics, trending-down topics, popular tweets, and no-trend topics.

2.2 Constructing Pseudo-honeypot

We use hashtag-based and trending-based attributes to select a set of accounts with each account possessing at least one attribute. The pseudo-honeypot takes these types of accounts as parasitic bodies to monitor their posted tweets and behavioral patterns, thus are more likely of capturing spammers. Note here, the Twitter APIs (e.g., Streaming API) that are available for developers enable the pseudo-honeypot to monitor user accounts activities (only public information) while keeping invisible to them so as to follow Twitter’s privacy policy. Through this way, we have constructed the pseudo-honeypot that can perform similar functions as the honeypot in terms of attracting spammers without manually creating user accounts. Obviously, such pseudo-honeypot has salient advantages in the attribute availability, deployment flexibility, network scalability, while has no chance to be recognized by spammers when comparing to the traditional honeypot.

2.3 Pseudo-honeypot Monitoring

With the constructed pseudo-honeypot network, we can start to monitor tweets and users’ behavioral patterns. In this stage, we collect only the direct interactive behaviors instead of all streaming passing through the associated accounts to reduce the processing workload of the pseudo-honeypot network. In particular, we collect the “mentions” behaviors from the selected accounts as the “mention” activities are proactive behaviors that spammers commonly use to attract user’s attention and interact with victims.

2.4 Pseudo-Honeypot Spam Detector

It is unrealistic to perform manual checking among all collected tweets. Here, we employ machine learning-based classifier for pseudo-honeypot spam detector so as to automatically classify the collected data as spams or not. Our detector includes three components. First, we extract a rich set of features from account profiles, tweet contents, and users behaviors that can reflect tweets’ characteristics. Next, we manually label a ground truth dataset for training to cover a majority types of spammers. In the end, we employ the machine learning algorithms to classify the collected tweets and users.

3 SYSTEM EVALUATION

In this section, we implement the pseudo-honeypot system and use experimental results to demonstrate the advances of our system in gathering spammers.

3.1 System Implementation

Our system is implemented by selecting a set of accounts that include the attributes as discussed in Section 2.1. For hashtag-based attributes, we select the top 10 hashtags (from [2]) in each attribute and identify 10 accounts associated with each hashtag to serve as the pseudo-honeypot nodes. For example, in entertainment, we collect the top 10 popular hashtags and then select 10 user accounts for each hashtag. Hence, there is a total of 100 pseudo-honeypot nodes in entertainment attribute. Similarly, we also select 100 pseudo-honeypot nodes for business, tech, education, environment, and social, respectively. In total, we have 600 pseudo-honeypot nodes in the hashtag-based category. For the trending-based category, we select the top 10 topics that are satisfying each attribute from [2] and then identify 10 user accounts for each topic to serve as the pseudo-honeypot nodes. For the non-trending topic, we randomly select 100 pseudo-honeypot nodes that are not posting tweets with the topics in [2]. Thus, there is a total of 400 pseudo-honeypot nodes under the trending-based category. Overall, we have created a pseudo-honeypot network with 1000 nodes with the construction time less than 1 min. Such low time cost outperforms the traditional honeypot-based solution, which requires considerable human efforts to create and configure honeypots.

Our experiment is conducted with a total of 500 hours to evaluate the performance of the constructed pseudo-honeypot system. Within this period, we collected a total of 1,227,708 mention behavioral tweets which associated with a total amount of 476,345 unique accounts.

3.2 Ground Truth Labeling

We create a 100-node pseudo-honeypot network by randomly selecting attributes from hashtag-based and trending-based categories, and then run 300 hours to collect tweets to serve as the ground truth training dataset. We manually label a total of 1,290
We take Random Forest as our machine learning classifier to equip our pseudo-honeypot detector and perform classification in the all 500-hour collected tweets. There is a total of 371,981 tweets that are identified as spams and all remains are classified as non-spams. The 371,981 spams are associated with 53,857 unique accounts, so we have classified a total of 53,857 spammers. This result significantly outperforms the prominent honeypot solutions from Stringhini et. al. [6], Lee et. al. [4], Yang et. al. [8], and Yang et. al. [8]’s advanced system, which gathered spammers of 15,857 in 11 months, 36,000 in 7 months, 1,159 in 5 months, and 17,336 in 10 days, respectively.

Figure 2 shows the total number of collected tweets, classified spams and spammers under various hashtag-based attributes. This figure shows that social, technology, and business are top three attributes in our pseudo-honeypot system capturing most spammers and spammers under various hashtag-based attributes. This figure shows that spam percentages in all users with the values of attributes of spammers in various trending-based attributes. This figure shows the ratio of spammers over total tweets (i.e., spam capture rate) in term of gathering spams, i.e., trending up, popular, trending down, technology and entertainment. These five attributes are used to select pseudo-honeypot nodes, where 10 users accounts are selected of possessing each attribute. With such a 50-node pseudo-honeypot network, we run a total of 50 hours and gather 2,301 spammers. The efficiency of spammers gathering is 0.92 spammers per honeypot per hour. However, in Stringhini et. al. [6], Lee et. al. [4], Yang et. al. [8], Yang et. al. [8]’s advanced system, the efficiencies are only 0.0067, 0.12, 0.0034, and 0.087, respectively. This demonstrates the advantages of pseudo-honeypot system over the state-of-the-art honeypot-based solutions.

4 CONCLUSION

In this paper, we proposed pseudo-honeypot as a novel solution for spammers gathering in Twitter social networks. By taking advantage of user diversity, the proposed pseudo-honeypot system can substantially improve the deployment flexibility and enrich the feature availability while removing the risk of being recognized by spammers, when comparing the prominent honeypot-based solutions. We presented the system design of pseudo-honeypot and implement it in the Twitter social network for spammer gathering. Through conducting the experimental analysis on gathered data, we validated the effectiveness of the pseudo-honeypot system in spammers gathering.

REFERENCES