International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER SSRN

IJIEMR Transactions, online available on 3rd Apr 2019. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-04

Title: A HYBRID APPROACH FOR DETECTING AUTOMATED SPAMMERS IN TWITTER

Volume 08, Issue 04, Pages: 57–63.

Paper Authors

Mr. V.RAVI KUMAR REDDY

Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

A HYBRID APPROACH FOR DETECTING AUTOMATED SPAMMERS IN TWITTER

Mr. V.RAVI KUMAR REDDY
VI[th] semester, dept of MCA , Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA
ravireddyv9666@gmail.com

ABSTRACT: Twitter is one of the most popular micro blogging services, which is generally used to share news and updates through short messages restricted to 280 characters. However, its open nature and large user base are frequently exploited by automated spammers, content polluters, and other ill-intended users to commit various cybercrimes, such as cyber bullying, trolling, rumor dissemination, and stalking. Accordingly, a number of approaches have been proposed by researchers to addressthese problems. However, most of these approaches are based on user characterization and completely disregarding mutual interactions. In this paper, we present a hybrid approach for detecting automated spammers by amalgamating community based features with other feature categories, namely metadata-, content-, and interaction-based features. The novelty of the proposed approach lies in the characterization of users based on their interactions with their followers given that a user can evade features that are related to his/her own activities, but evading those based on the followers is difficult. Nineteen different features, including six newly defined features and two redefined features, are identified for learning three classifiers, namely, random forest, decision tree, and Bayesian network, on a real dataset that comprises benign users and spammers. The discrimination power of different feature categories is also analyzed, and interaction- and community-based features are determined to be the most effective for spam detection, whereas metadata-based features are proven to be the least effective.

Keywords: Social network analysis,spammer detection,spambot detection, social network security.

I INTRODUCTION

TWITTER, a microblogging service, is considered a popular online social network (OSN) with a large user base and is attracting users from different walks of life and age groups. OSNs enable users to keep in touch with friends,relatives, family members, and people with similar interests, profession, and objectives. In addition, they allow users to interact with one another and form communities. A user can become a member of an OSN by registering and providing details, such as name, birthday,

gender, and other contact information. Although a large number of OSNs exist on theweb, Facebook and Twitter are among the most popular OSNs and are included in the list of the top 10 websites1 around the worldwide.

A. OSN and the Social Spam Problem

Twitter, which was founded in 2006, allows its users to post their views, express their thoughts, and share news and other information in the form of tweets that are restricted to280 characters. Twitter allows the users to follow their favourite politicians, athletes, celebrities, and news channels, and to subscribe to their content without any hindrance. Through following activity, a follower can receive status updates of subscribed account. Although Twitter and other OSNs are mainly used for various benign purposes, their open nature, huge user base, and real-time message proliferation have made them lucrative targets for cyber criminals and social bots. OSNs have been proven to be incubators for a new breed of complex and sophisticated attacks and threats, such as cyberbullying, misinformation diffusion, stalking, identity deception, radicalization, and other illicit activities, in addition to classical cyber attacks, such as spamming, phishing, and drive by download . Over the years, classical attacks have evolved into sophisticated attacks to evade detection mechanisms. A report2submitted to the US Securities and Exchange Commission in August 2014 indicates that approximately 14% of Twitter accounts are actually spam bots and approximately 9.3% of all tweets are spam. In social networks, spam bots are also known as social bots that mimic human behaviour to gain trust in a network and then exploit it for malicious activities . Such reports and findings demonstrate the extent of cyber crimes committed by spam bots and how OSNs are proving to be a heaven for these bots. Although spammers are less than benign users, they are capable of affecting network structure and trust for various illicit purposes.

The main contributions of this study can be summarized as follows.

• A novel study that uses community-based features with other feature categories, including metadata, content, and interaction, for detecting automated spammers.
• Six new features are introduced and two existing features are redefined to design a feature set with improved discriminative power for segregating benign users and spammers. Among the six new features, one is content based, three are interaction-based, and the remaining two are community-based. Meanwhile, both redefined features are content-based. When defining interaction based features, focus should be on the followers of a user, rather than on the ones he/she is followings.
• A detailed analysis of the working behavior of automated spammers and benign users with respect to newly defined features. In addition, two-tailed Z-test statistical significance analysis is performed to answer the following question: “is the difference between the working behavior of spammers and benign users in terms of newly defined features a random chance?”
• A thorough analysis of the discriminating power of each feature category in

segregating automated spammers from benign users.

II SYSTEM ANALYSIS

EXISTING SYSTEM

- Sahami et al. proposed textual and non textual and domain-specific features and learned naive Bayes classifier to segregate spam emails from legitimate ones. Schafer and proposed metadata-based approaches to detect botnets based on compromised email accounts to diffuse mail spams. Spam campaigns on Facebook were analyzed by Gao et al. using a similarity graph based on semantic similarity between posts and URLs that point to the same destination.
- Furthermore, they extracted clusters from a similarity graph, wherein each cluster represents a specific spam campaign. Upon analysis, they determined that most spam sources were hijacked accounts, which exploited the trust of users to redirect legitimate users to phishing sites. In and honey profiles were created and deployed on OSNs to observe the behavior of spammer. Both studies presented different sets of features to discriminate benign users from spammers and evaluated them on different sets of OSNs.
- Wang used content and graph-based features to classify malicious and normal profiles on Twitter. In contrast to honey profiles, Wang used Twitter API to crawl the dataset. Yang et al. Wang and Ahmed and Abulaish used content- and interaction based attributes for learning classifiers to segregate spammers from benign users on different OSNs.
- Yang et al. and Ahmed and Abulaish analyzed the contribution of each feature to spammer detection, whereas Yang et al. conducted an in-depth empirical analysis of the evasive tactics practiced by spammers to bypass detection systems. They also tested the robustness of newly devised features. In Zhu et al. used a matrix factorization technique to find the latent features from the sparse activity matrix and adopted social regularization to learn the spam discriminating power of the classifier on the Renren network, one of the most popular OSNs in China. Another spammer detection approach in social media was proposed by Tan et a.

PROPOSED SYSTEM

In the proposed system, the system proposes a hybrid approach for detecting social spam bots in Twitter, which utilizes an amalgamation of metadata-, content-, interaction-, and community-based features. In the analysis of characterizing features of existing approaches, most network-based features are not defined using user followers and underlying community structures, thereby disregarding the fact that the reputation of user in a network is inherited from the followers (rather than from the ones user is following) and community members. Therefore, the system emphasizes the use of followers and community structures to define the network-based features of a user.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

The system classifies set of features into three broad categories, namely, metadata, content, and network, wherein the network category is further classified into interaction- and community based features. Metadata features are extracted from available additional information regarding the tweets of a user, whereas content-based features aim to observe the message posting behavior of a user and the quality of the text that the user uses in posts. Network-based features are extracted from user interaction network.

III IMPLEMENTATION

Tweet Admin

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as View Users and Authorize(Give link on user to view Profile),View all Uses Friend Request and Response,Add Spam Filter name,View All spamming accounts with profile details and Block,View All Un Block request users details using decision tree format and Unblock by clicking user name ,View all User's Tweet Topic with Interactions and scores,View All Spam Account(Based on Virus,Malware) And Normal Account with Reasons based on Random Forest Tree, View All Spamming and Normal Behaviors based on Interactions by Filter Name and give link to show Number of both users in chart, View All Spamming and Normal Behaviors based on Tweet Meta Data by Filter Name and give link to show Numberof boith users in chart, View Numbe of Spamming Account and Normal Account in Chart

Friend Request & Response

In this module, the admin can view all the friend requests and responses. Here all the requests and responses will be displayed with their tags such as Id, requested user photo, requested user name, user name request to, status and time & date. If the user accepts the request then the status will be changed to accepted or else the status will remains as waiting.

User

In this module, there are n numbers of users are present. User should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like View Your Profile with community, Search Friends based on community, View Friend Request and Response,View My Friends based on community, Create Tweet Topic withtweet_postname,TAbout,TUses,tcontent desc,BrowseMetaData_desc,TweetURL,TDate and Time, TO wner,add TImage, Search Tweet Topic by keyword and give Your Interactions(increse score while viewing) and view URL to see web page, View all your Tweets Topic with other Interactions and scores, View all your Friends Tweet Topic with other Interactions and scores and give your Interactions, View All Similar Friend's Tweets Topic, show all Spamming behaviors friends Topics with profile.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IV SYSTEM DESIGN

SYSTEM ARCHITECTURE:

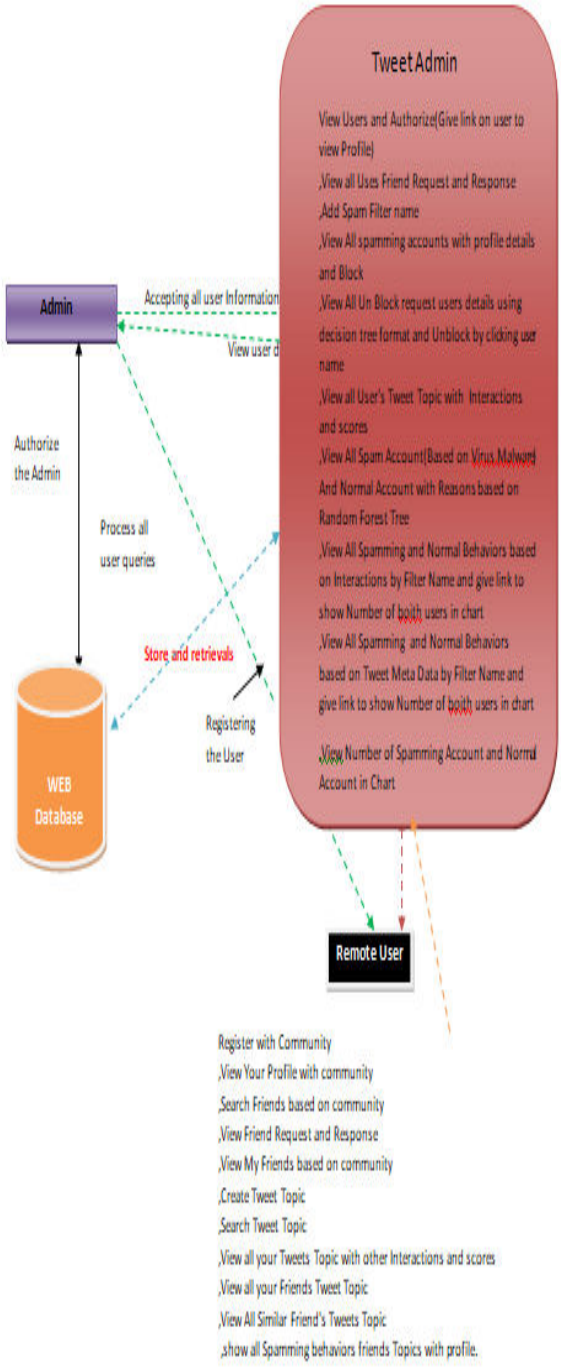

Figure 1: system architecture

DATA FLOW DIAGRAM:

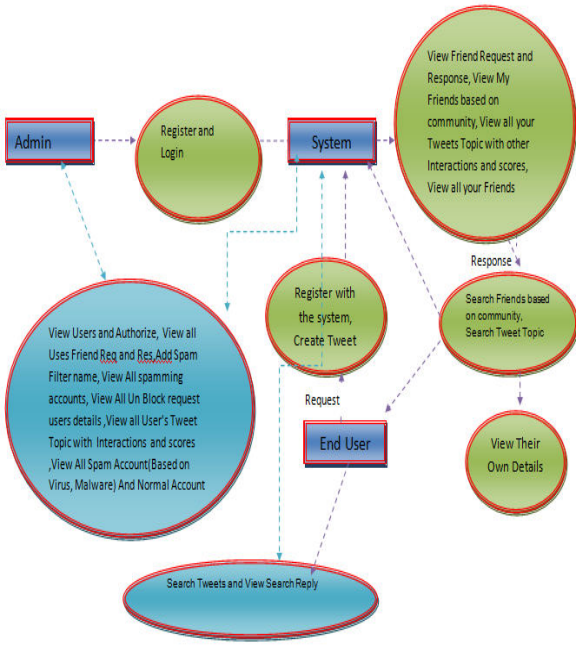

Figure 2: data flow diagram

V CONCLUSION

In this paper, we have proposed a hybrid approach exploiting community-based features with metadata-, content-, and interaction-based features for detecting automated spammers in Twitter. Spammers are generally planted in OSNs for varied purposes, but absence of real-life identity hinders them to join the trust network of benign users. Therefore, spammers randomly follow a number of users, but rarely followed back by them, which results in low edge density among their followers and followings. This type of spammers interaction pattern can be exploited for the development of effective spammers detection systems. Unlike existing approaches of characterizing spammers based on their own profiles, the novelty of the proposed approach lies in the characterization of a spammer based on its neighboring nodes (especially, the followers) and their interaction network.

This is mainly due to the fact that users can evade features that are related to their own activities, but it is difficult to evade those that are based on their followers.

On analysis, metadata-based features are found to be least effective as they can be easily evaded by the sophisticated spammers by using random number generator algorithms. On the other hand, both interaction- and community-based features are found to be the most discriminative for spammers detection. Attaining perfect accuracy in spammers detection is extremely difficult, and accordingly any feature set can never be considered as complete and sound, as spammers keep on changing their operating behavior to evade detection mechanism. Therefore, in addition to profile-based characterization, complete logs of spammers starting from their entry in the network to their detection, need to be analyzed to model the evolutionary behavior and phases of the life-cycles of spammers. But, generally spammers are detected when they are at very advanced stage, and it is difficult to get their past logs data. Moreover, it may happen that a user is operative in the network as a benign user, and later on, it start sillicit activities due to whatsoever reasons, and considered as spammer. In this circumstance, even analyzing log data may lead to wrong characterization.Analysis of spammers network to unearth different types of coordinated spam campaigns run by the spam bots seems one of the promising future directions of research. Moreover, analyzing the temporal evolution of spammers' followers may reveal some interesting patterns that can be utilized for spammers characterization at different levels of granularity.

VI REFERENCES

[1] M. Tsikerdekis, "Identity deception prevention using common contribution network data," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1,pp. 188–199, Jan. 2017.

[2] T. Anwar and M. Abulaish, "Ranking radically influential Web forum users," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 6,pp. 1289–1298, Jun. 2015.

[3] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "Design and analysis of social botnet," *Comput. Netw.*, vol. 57, no. 2, pp. 556–578,2013.

[4] D. Fletcher, "A brief history of spam," *TIME*, Nov. 2, 2009.[Online]. Available:http://www.time.com/time/business/article/0,8599,1933796,00.html

[5] Y. Boshmaf, M. Ripeanu, K. Beznosov, and E. Santos-Neto, "Thwarting fake OSN accounts by predicting their victims," in *Proc. AISec*, Denver,CO, USA, 2015, pp. 81–89.

[6] A. A. Amleshwaram, N. Reddy, S. Yadav, G. Gu, andC. Yang, "CATS: Characterizing automation of Twitter spammers,"in *Proc. COMSNETS*, Bengaluru, India, Jan. 2013, pp. 1–10.

[7] K. Lee, J. C. Lee, and S. Webb, "Uncovering social spammers: Socialhoneypots+ machine learning," in *Proc. SIGIR*, Geneva, Switzerland,Jul. 2010, pp. 435–442.

Author

Mr. V.RAVI KUMAR REDDY, VIth semester, dept of MCA , Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA .Email ID: ravireddyv9666@gmail.com