DETECTING BOT MESSAGE FROM SOCIAL MEDIA USING LANGUAGE MATCH ALGORITHM

[1]Jatin vaghehswari, [2]Saket Swarndeep
[1]PG Scholar - LJIET, [2]Ass.Prof. LJIET,
[1]Computer Engineering,
[1]LJ Institute of Engineering & Technology, Ahmedabad, India
[1]jatin.v.1994@gmail.com, [2]nandasaket1990@gmail.com

Abstract:- With the growing development of mobile internet and powerful smartphones, botnets have invaded the mobile domain. Social media, like Twitter, Facebook, and YouTube have created a new communication channel for spammers. Bot started to exploit social media for different fake activity, such as sending spam, recruitment of new bots, and botnet command and control.in the proposed system the user sent the tweet on twitter that is identify that sent tweet is sent by user or BOT .in proposed model language match that is train according to the language that is identify that sent tweet is human generated or bot generated.in this we have use our own algorithm for identify tweet. that is help to reduce the spam messages in social media like twitter.

Keyword: Online Social Network, Social Network Service, security, Botnet,Cyber security.

I. Introduction

The huge market for open source OS has opened the door for malicious writers to target the Vulnerable features of Open source OS. The survey that the 95% malware application targets the open sources like Android. Other than the mobile malware the new mobile botnet is now popular in the social media. The mobile malware is involving decrease the performance smart phones. Online social networking websites(OSNs), such as Twitter, Facebook and Sina Weibo have play an important part in people’s life. By using these social networking services, it is convenient for people to communicate with their friends easily, publish posts about their life freely, and follow hot topics immediate. We have survey that the many detection technology and algorithms are foud in to the detect the bot and Spammers. Social media bots are automatic or semi-automatic computer programs that mimic human behavior in OSN [3].For example, Facebot [4] is a mobile botnet where the bots steal information from the infected device and embed it steganographically in the profile picture of the user's account in Facebook. Bots encrypt this information with the bot master's public key to ensure that it is not accessible to anyone but the bot master. The bots join a group agreed upon by Facebot and the bot master, then the bot master scroll through new members of the groups and check their profile pictures for availability of stolen information.

In this paper, we develop a detection method to detect mobile botnets that make use of Twitter. The detection method is intended to be able to classify tweets as bot tweets and legitimate tweets, allow only legitimate tweets and block bot tweets from being sent. The proposed method use our own algorithm in language match.techniques are described in Section II. A full implementation of the proposed method is developed as an Android application that is used to evaluate the effectiveness of the method.

The rest of this paper is organized as follows: Section II discusses previous related work. Section III describes our detection method. Section IV shows the actual implementation of the proposed method. In Section V, test results for several scenarios are shown and discussed. Finally, Section VI presents our conclusions and directions for future work.

II. Related Work

Botnets are a new threat to the mobile world. Their threat have triggered several research work in detecting mobile botnets. The proposed method combines the correlation between tweeting and user activity, such as clicks or taps, and an Artificial Immune System detector, to detect tweets caused by bots and differentiate them from tweets generated by user or by user-approved applications. This detector creates a signature of the tweet and compares it with a dynamically updated signature library of bot behavior signatures. The proposed system has been fully implemented on Android platform and tested under several sets of generated Tweets[1]. detection mechanism that measures the causal relationship between network traffic and human activity, like mouse clicks or keyboard strokes. Communication with social media that is not assignably caused by human activity, is classified as anomalous. We explore both theoretically and experimentally this detection Mechanism by a case study, with Twitter.com as a Command and Control channel, and demonstrate successful real time detection of botnet Command and Control traffic[2]. first they extract the data from yeld database that is used as unlabeled data.label data used for training and testing this model.then it's come to the data preprocessing .in preprocessing they have apply the different method to remove the noisy,missing and inconsistent data that is used in to take decision[3]. they also use the Random forests Algorithm for classify the spam from the twitter.Random Forests algorithm is the combination of classification and regression that operate by constructing the tree at the training time.thay use bootstrap and bagging with random selection of feature.bagging is select the training data and fit in to the model.Bagging is the create the random forest from the train set classify then implemented using each model of random tree[6]. In these sentiment systems they where use the twitter API and then they identify the user and set their user id.after that it's extract the user network. then extract user profile with their previous behaviour.then it is store that in CV database and create the variable that is define the user aspect[4]. First section user web traffic using network proxy. Second section preprocessing of web usage data. Third section extract the features based on which model is created for user.Network proxy:- in the network proxy we connect the traffic that use network proxy setting in Smartphone. In this setting we can specify the proxy server address and port of the network[5].

III. PROPOSED WORK

In this proposed work the language match is presemt to detect the tweet from twitter.the flow chart of proposed work is shown in figure 1.

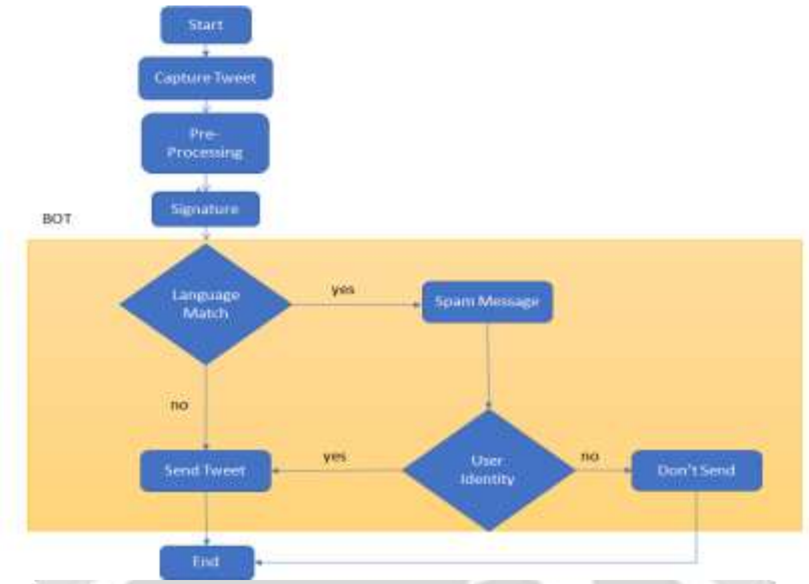

Fig 1 flow chart of proposed work

The proposed method working following :-

1. Start and Capture tweet from twitter.
2. Preprocess of that captured tweet.
3. After the preprocessing create the signature of that tweet.
4. After the signature it go to the language match .in the language match we have our own signature database.our algorithm is use to find the signature of tweet and signature of database compare.
5. If there is match in to the signature database then it is take as bot tweet and create pop-up for user approval.if user approve that tweet then sent it and delete from signature database.otherwise don't tweet.

A. Capture Tweet

In this module tweet sent to twitter from mobile device that is go to language match to decide the tweet is real user tweet or bot tweet.

B. Preprocess

In the preprocess extracted tweets.we were eliminates no longer needed or not useful data from that tweet.and reduce the tweets which may be spam tweet or bot tweet.then we analysis the tweet abd get the some useful data that is more help to identify that tweet is bot tweet or not.

C. Signature

We create signature from every tweet that is sent on twitter.the signature is created into digital form and tweet is converted in to the digital form.the characteristics of tweets are stored in numeric format.the signature is created from tweet from word like marketing word.then URL is use in that tweet.hashtage and mansions.

D. Signature database

Signature database is contain the signature that we have already created from the some random bot tweets.after all the detection signature is not bot tweet that is deleted and if new bot tweet signature that replace that signatures.

E. Data selection

Twitter has released a set of API functions that support user information collection []. The data used to train the AIS detector was collected from Twitter's Public stream API.

IV. EVALUATION AND RESULT

The evolution of proposed model we have test oue application against datasets of generated tweets from user generated tweets and bot generated tweets.we have arrange the datasets in three scenario.

Table 1:- summary of datasets for use in evaluation

Scenario	Tweets	From user	From Bot
1	240	210	30
2	250	200	50
3	200	10	190

Table 2:- Summary of result

scenario	Correctly detected	False positive	False nagetive	Detection rate
1	50	30	160	92.4%
2	40	45	165	94.9%
3	70	10	120	95.5%

V. CONCLUTION

In this proposed method we have developed a method to detect mobile botnets that use online social networks as their Command and Control channel. Our detection method works on mobile devices and detects botnets based on the existence of user activity in addition to checking the content of the tweet it self.The language match is detect that message is spam message or user message. The accuracy of the detection improves after the system has been trained with user input.

VI. REFERENCES

[1] Reham A. Al-Dayil, Mostafa H. Dahshan,” Detecting Social Media Mobile Botnets Using User Activity Correlation and Artificial Immune System” pp:-109-114 7th International Conference on Information and Communication Systems ©2016 IEEE journal.

[2] Burghouwt, Pieter, Marcel Spruit, and Henk Sips. "Towards detection of botnet communication through social media by monitoring user activity." In *International Conference on Information Systems Security*, pp. 131-143. Springer Berlin Heidelberg, 2011.

[3] Ahsan, MN Istiaq, Tamzid Nahian, Abdullah All Kafi, Md Ismail Hossain, and Faisal Muhammad Shah. "Review spam detection using active learning." In *Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual*, pp. 1-7. IEEE, 2016.

[4] John P. Dickerson, Vadim Kagan, V.S. Subrahmanian “Using Sentiment to Detect Bots on Twitter:Are Humans more Opinionated than Bots?” August 17-20, 2014, Beijing, China,pp:-1-8.IEEE journal 2014.

[5**]** Kilari, Vishnu Teja, Guoliang Xue, and Lingjun Li. "Host Based Detection of Advanced MiniDuke Style Bots in Smartphones through User Profiling." In *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1-6. IEEE,journal 2015.

[6] jatin vaghsheshwari,Hinal somani.”A survey on detecting Activity of fake user in social media by intelligent bot using Artificial intelligent” Internation Journal for Science and Advance Research in Technology pp:-405-408,2016.

[7] S. Soltani, S. Seno, M. Nezhadkamali and R. Budirato, "A Survey On Real World Botnets And Detection Mechanisms," International Journal of Information & Network Security (IJINS), vol. 2, pp. 116-127, April 2014.

[8] “Mobile botnets: The next big threat to take aim at smartphones tablets,” CyberTrend, 2015, [Online;accessed 1-June-2015]. Available: http://www.cybertrend.com/article/16969/mobile-botnets/ access on 3/10/2016 on 3:01 pm

[9] https://www.techopedia.com/definition/10282/information-security-is Access on[25/8/2016 at 5:31 pm]

[10] http://www.webopedia.com access on [20/8/2016 at 10:03 pm]

[11] https://qz.com/572763/the-best-twitter-bots-of-2015/s access on [21/8/2016 at 7:30 pm]