International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 04 | April - 2021 ISSN: 2582-3930

Data-Driven Cyber Security approach for Twitter traffic classification

Rahul Suresh Ubale
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Rohan Kamble
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Anuja Sonawane
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Mohaseen Shaikh
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Mohini Devikar
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Shubhangi Khade
dept. of Computer Engineering
Modern Education Society's College of Engineering
Pune,India

Abstract*— Social network sites involve billions of users around the world wide. User interactions with these social sites, like twitter have tremendous and occasionally undesirable impact implications for daily life. The major social networking sites have become a target platform for spammers to disperse a large amount of irrelevant and fundamental in detecting and defending cyber-attacks. Twitter, it has become one of the most extravagant platforms of all time and, most popular microblogging services which is generally used to share unreasonable amount of spam . Fake users send unwanted tweets to users to promote services or websites that do not only affect legitimate users, but also interrupt resource consumption. Furthermore, the possibility of expanding invalid information to users through false identities has increased, resulting in malicious content. Recently, the detection of spammers and the identification of fake users and fake tweets on Twitter has become an important area of research in online social networks (OSN). In this Paper, proposed the techniques used to detect spammers on Twitter. In addition, a taxonomy of Twitter spam detection approaches is presented which classifies techniques based on their ability to detect false content, URL - based, spam on trending issues. Twelve to Nineteen different features, including six recently defined functions and two redefined functions, identified to learn two machine supervised learning classifiers, in a real time data set that distinguish users and spammers.

***Keywords*—** *Cyber security, Internet traffic analysis, machine learning (ML), social spam detection.*

I. INTRODUCTION (*HEADING 1*)

Online social networking sites like Twitter, Facebook, Instagram and some online social networking companies have become extremely popular in recent years. People spend a lot of time in OSN making friends with people they are familiar with or interested in. The expanded interest of social sites grants users to gather bounteous measure of data and information about users. Large volumes of information accessible on these sites additionally draw the attention of spammers. Twitter has quickly become an online hotspot for obtaining continuous data about users. Twitter is an Online Social Network (OSN) where users can share anything and everything, such as news, opinions, and even their moods. Several arguments can be held over different topics, such as politics, current affairs, and important events. At the point when a client tweets something, it is right away passed on to his/her supporters, enabling them to extended the got data at a lot more extensive level. With the development of OSNs, the need to ponder and break down clients' practices in online social stages has strengthened. Numerous individuals who don't have a lot of data with respect to the OSNs can without much of a stretch be deceived by the fraudsters. There is additionally an interest to battle and place a control on the individuals who use OSNs just for commercials and in this manner spam others' records.

Recently, the recognition of spam in social networking sites attracted the consideration of researchers. Spam detection is a difficult task in maintaining the security of social networks It is basic to perceive spams in the OSN locales to spare clients from different sorts of malevolent assaults and to protect their security and protection. These unsafe moves embraced by spammers cause huge demolition of the network in reality. Twitter spammers have different targets, for example, spreading invalid data, counterfeit news, bits of gossip, and unconstrained messages. Spammers accomplish their noxious destinations through promotions and a few different methods where they bolster diverse mailing records and consequently dispatch spam messages haphazardly to communicate their inclinations. These exercises cause unsettling influence to the first clients who are known as non-spammers. Furthermore, it likewise diminishes the notoriety of the OSN stages. Subsequently, it is fundamental to plan a plan to spot spammers so restorative endeavors can be taken to counter their malevolent exercises.

The ability to order useful information is essential for the academic and industrial world to discover hidden ideas and predict trends on Twitter. However, spam generates a lot of noise on Twitter. To detect spam automatically, researchers applied machine learning algorithms to make spam detection a classification problem. Ordering a tweet broadcast instead of a Twitter user as spam or non-spam is more realistic in the real world.

II. RELATED WORK

Literature survey is the most important step in any kind of research. Before start developing we need to study the previous papers of our domain which we are working and on the basis of study we can predict or generate the drawback and start working with the reference of previous papers.

In this section, we briefly review the related work on Spam Detection and their different techniques.

Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro [1] describe the "Aiding the detection of fake accounts in large scale social online services".in this paper, Sybil Rank, an effective and efficient fake account inference scheme, which allows OSNs to rank accounts according to their perceived likelihood of being fake. It works on the extracted knowledge from the network so it detects, verify and remove the fake accounts.

G. Stringhini, C. Kruegel, and G. Vigna [2] describe the "Detecting spammers on social networks" this paper help to detect spam Profiles even when they do not contact a honey-profile. The irregular behavior of user profile is detected and based on that the profile is developed to identify the spammer.

J. Song, S. Lee, and J. Kim [3] describe the "Spam filtering in Twitter using sender receiver relationship" in this paper a spam filtering method for social networks using relation information between users and System use distance and connectivity as the features which are hard to manipulate by spammers and effective to classify spammers.

K. Lee, J. Caverlee, and S. Webb [4] describe the "Uncovering social spammers: social honeypots and machine learning" in this System analyzes how spammers who target social networking sites operate to collect the data about spamming activity, system created a large set of "honey-profiles" on three large social networking sites.

Nathan Aston, Jacob Liddle and Wei Hu*[5] describe the "Twitter Sentiment in Data Streams with Perceptron" in this system the implementation feature reduction we were able to make our Perceptron and Voted Perceptron algorithms more viable in a stream environment. In this paper, develop methods by which twitter sentiment can be determined both quickly and accurately on such a large scale.

K. Thomas, C. Grier, D. Song, and V. Paxson [6] describe the "Suspended accounts in retrospect: An analysis of Twitter spam" in this paper the behaviors of spammers on Twitter by analyzing the tweets sent by suspended users in retrospect. An emerging spam-as-a-service market that includes reputable and not-so-reputable affiliate programs, ad-based shorteners, and Twitter account sellers.

K. Thomas, C. Grier, J. Ma, V. Paxson, and D.Song [7] describe the "Design and evaluation of a real-time URL spam filtering" in this paper, service Monarch is a real-time system for filtering scam, phishing, and malware URLs as they are submitted to web services. Monarch's architecture generalizes to many web services being targeted by URL spam, accurate classification hinges on having an intimate understanding of the Spam campaigns abusing a service.

X. Jin, C. X. Lin, J. Luo, and J. Han [8] describe the "Social spam guard: A data mining based spam detection system for social media networks" in this paper Automatically harvesting spam activities in social network by monitoring social sensors with popular user bases.Introducing both image and text content features and social network features to indicate spam activities. Integrating with our GAD clustering algorithm to handle large scale data. Introducing a scalable active learning approach to identify existing spams with limited human efforts, and Perform online active learning to detect spams in real-time.

S. Ghosh et al [9] describe the "Understanding and combating link farming in the Twitter social network" in this paper Search engines rank websites/webpages based on graph metrics such as PageRank High in-degree helps to get high PageRank. Link farming in Twitter Spammers follow other users and attempt to get them to follow back.

H. Costa, F. Benevenuto, and L. H. C. Merschmann [10] describe the "Detecting tip spam in location-based social networks" in this paper identifying tip spam on a popular Brazilian LBSN system, namely Apontador. Based on a labelled collection of tips provided by Apontador as well as crawled information about users and locations, we identified a number of attributes able to distinguish spam from non-spam tips

III. EXISTING SYSTEM

Lot of work has been done in this field because of its extensive usage and applications. In this section, some of the approaches which have been implemented to achieve the same purpose are mentioned. These works are majorly differentiated by the algorithm for spam detection.

IV. PROPOSED SYSTEM

In proposed system, the process of Twitter spam detection by using machine learning algorithms. Before classification, a classifier that contains the knowledge structure should be trained with the pre-labeled tweets. After the classification model gains the knowledge structure of the training data, it can be used to predict a new incoming tweet. The whole process consists of two steps: learning and classifying. Features of tweets will be extracted and formatted as a vector. The class labels i.e. spam and non-spam could be get via some other approaches. Features and class label will be combined as one instance for training. One training tweet can then be represented by a pair containing one feature vector, which represents a tweet, and the expected result, and the training set is the vector. The training set is the input of machine learning algorithm; the classification model will be built after training process. In the classifying process, timely captured tweets will be labelled by the trained classification model.

A. Advantages

1. The system implements a method that will use the ML mechanism to detect if the post is spam or not.

2. Implementation of system can also be hosted online for use and data will be archived and retrieved from the server.
3. The user with the maximum amount of spam can be blocked by the system.

B. *System Architecture*

Proposed system, we evaluate the spam detection performance on our dataset by using machine learning algorithm. The process of Twitter spam detection by using machine learning algorithms. Before classification, a classifier that contains the knowledge structure should be trained with the pre-labeled tweets. After the classification model gains the knowledge structure of the training data, it can be used to predict a new incoming tweet. The whole process consists of two steps: 1) learning and 2) classifying. First, features of tweets will be extracted and formatted as a vector. The class labels (spam or no spam) could be get via some other approaches (like manual inspection). Features and class label will have combined as one instance for training. One training tweet can then be represented by a pair containing one feature vector, which represents a tweet, and the expected result, and the training set is the vector. The training set is the input of machine learning algorithm; the classification model will be built after training process. In the classifying process, timely captured tweets will be labeled by the trained classification model.

V. CONCLUSION

We have presented a new research approach for DDCS and reviewed its application in social and Internet traffic analysis. DDCS shows the strong link among data, model, and methodology during the review of key recent works in Twitter spam detection and IP traffic classification. We have highlighted the challenges and future works about big traffic data, domain knowledge, and research methodology. Hopefully, this survey can provide new insights and ideas to push the research boundary of cyber security, in particular, social and Internet traffic analysis.

REFERENCES

[1] Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro, "Aiding the detection of fake accounts in large scale social online services," in Proc. Symp. Netw. Syst. Des. Implement. (NSDI), 2012, pp. 197 – 210.

[2] G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in Proc. 26th Annu. Comput. Sec. Appl. Conf., 2010, pp. 1–9.

[3] J. Song, S. Lee, and J. Kim, "Spam filtering in Twitter using sender receiver relationship," in Proc. 14th Int. Conf. Recent Adv. Intrusion Detection, 2011, pp. 301–317.

[4] K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: social honeypots + machine learning," in Proc. 33rd Int. ACM SIGIR Conf. Res.Develop. Inf. Retrieval, 2010, pp. 435–442.

[5] Nathan Aston, Jacob Liddle and Wei Hu*, "Twitter Sentiment in Data Streams with Perceptron," in Journal of Computer and Communications, 2014, Vol-2 No-11.

[6] K. Thomas, C. Grier, D. Song, and V. Paxson, "Suspended accounts in retrospect: An analysis of Twitter spam," in Proc. ACM SIGCOMM Conf. Internet Meas., 2011, pp. 243–258.

[7] K. Thomas, C. Grier, J. Ma, V. Paxson, and D. Song, "Design and evaluation of a real-time URL spam filtering service," in Proc. IEEE Symp. Sec. Privacy, 2011, pp. 447–462.

[8] X. Jin, C. X. Lin, J. Luo, and J. Han, "Socialspamguard: A data mining based spam detection system for social media networks," PVLDB, vol. 4, no. 12, pp. 1458–1461, 2011.

[9] S. Ghosh et al., "Understanding and combating link farming in the Twitter social network," in Proc. 21st Int. Conf. World Wide Web, 2012, pp. 61–70.

[10] H. Costa, F. Benevenuto, and L. H. C. Merschmann, "Detecting tip spam in location-based social networks," in Proc. 28th Annu. ACM Symp. Appl. Comput., 2013, pp. 724–729.