ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 12, December 2018

Survey on Malicious Posts Detection in Social Networks

Gurunath Aade[1], Rahul Kapse[2], Ekta Ukey [3]

Department of Information Technology, Pillai HOC College of Engineering and Technology, Rasayani, India[1]

Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, India [2,3]

ABSTRACT: Today, peoples are increasing amount of time in social networks. However, because of the popularity of online social networks, cybercriminals are spamming on these platforms for potential victims. Spams invite users to external phishing sites or malware downloads huge security issue online and undermined the user experience. However, current solutions do not reveal the Twitterspamming accurately and indeed. In this paper, we compared the performance of a wide range of conventional machine learning algorithms, with the aim of identify those that offer satisfactory detection and stability performance based on a large amount of true field data. With the objective in order to realize the real-time spam detection capability, we evaluated scalability algorithms. Performance the study evaluates the accuracy of the detection, the TPR / FPR and the F measure; stability analyzes the stability of algorithms using randomly selected training samples of different sizes. Scalability aims to better understand the impact of in reducing training time learning algorithms.

KEYWORDS: Machine learning, Twitter, spam detection, parallel computing, and scalability.

I. INTRODUCTION

Social networking sites such as Twitter, Facebook, Instagram and some enterprise of online social network have become extremely popular in the last few years. Individuals spend vast amounts of time in OSNs making friends with people who they are familiar with or interested in. Twitter, which was founded in 2006, has become one of the most popular micro blogging service sites. Around 200 million users create around the 400 million new posts per day the growth of spam. Twitter spam, which is referred as unsolicited posts containing malicious links that directs victims to external sites containing malware spreading, malicious link spreading etc. has not only affected a number of legitimate users but also polluted the whole platform. Consider the example as during the Australian Prime Minster Election in 2013 published an alert that confirmed its Twitter account @AusElectoralCom was hacked. Many of its followers received direct spam messages which contained malicious links. The ability to sort out useful information is critical for both academia and industry to discover hidden insights and predict trends on Twitter. However, spam significantly brings noise into Twitter. To automatically detect spam, machine learning algorithms have been applied by researchers to make spam detection as a classification problem. Classifying a streaming tweet instead of a Twitter user to spam or non-spam is more realistic in the real world.

Background:

1. In existing system, to detect Twitter posts spam, made use of account and content features, such as account age, number of followers or followings, URL ratio, and the length of posts to distinguish spammers and non-spammers.
2. Existing System can detect whether an account was compromised or not, but cannot determine the accounts which were created by spammers fraudulently.

Motivation:

1. System creates a big ground-truth for the research on maliciousposts detection.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 12, December 2018

2. System reports the impact of the data related factors, such as maliciousto non- maliciousratio, training data size, and data sampling, to the detection performance.
3. System extracts 12 lightweight features for streaming postsmaliciousdetection
4. System investigates machine learning algorithms to build up the postsmaliciousdetection model.

Objectives:

1. To categories the maliciousand Non-malicioustweets.
2. To categorize the tag based postsand link based posts.
3. To improve the classification and detection accuracy.

II. RELATED WORK

Guanjun Lin, Nan Sun, Surya Nepal, Jun Zhang,Yang Xiang, and Houcine Hassan describe the "Statistical Twitter Spam Detection Demystified:Performance, Stability and Scalability"In this paper, they compared the performance of a wide range of mainstream machine learning algorithms, aiming to identify the ones offering satisfactory detection performance and stability based on a large amount of ground truth data. With the goal of achieving real-time Twitter spam detection capability, we further evaluated the algorithms in terms of the scalability [1].

Q. Cao, M. Sirivianos, X. Yang, and T. Pregueiro describe the Aiding the detection of fake accounts in large scale social online services.SybilRank, 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 [2].

G. Stringhini, C. Kruegel, and G. Vigna describe the Detecting spammers on social networks.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 [3].

J. Song, S. Lee, and J. Kim describe the Spam filtering in Twitter using sender receiver relationship.A spam filtering method for social networks using relation information between users.System use distance and connectivity as the features which are hard to manipulate by spammers and effective to classify spammers [4].

K. Lee, J. Caverlee, and S. Webb describe the Uncovering social spammers: social honeypots and machine learning.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 [5].

Nathan Aston, Jacob Liddle and Wei Hu* describe the Twitter Sentiment in Data Streams with Perceptron.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 [6].

K. Thomas, C. Grier, D. Song, and V. Paxson describe the Suspended accounts in retrospect: An analysis of Twitter spam 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 [7].

K.Thomas, C.Grier, J.Ma, V.Paxson, and D.Song describe the Design and evaluation of a real-time URL spam filtering service Monarch is a real-time system for filtering scam, phishing, and malware URLs as they are submitted to

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 12, December 2018

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 [8].

X. Jin, C. X. Lin, J. Luo, and J. Han describe the Socialspamguard: A data mining based spam detection system for social media networks. 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 [9].

S. Ghosh et al describe the Understanding and combating link farming in the Twitter social network.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 [10].

H. Costa, F. Benevenuto, and L. H. C. Merschmann describe the Detecting tip spam in location-based social networks.Identifying tip spam on a popular Brazilian LBSN system, namely Apontador.Based on a labeled 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 [11].

III. PROPOSED SYSTEM APPROACH

Proposed System

Proposed system, I evaluate the malicious posts detection performance on our dataset by using support vector machine learning algorithm. The process of malicious posts detection by using machine learning algorithms. Before classification, a classifier that contains the knowledge structure should be trained with the prelabeled 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 non-spam) could be get via some other approaches (like manual inspection). Features and class label will be combined as one instance for training. One training posts can then be represented by a pair containing one feature vector, which represents a posts, 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 posts will be labeled by the trained classification model.

Advantages of Proposed System:

1. Extraction of 12 features and categories as Tag based features and URL based features.
2. The system implements a method which will use spot filter mechanism to detect whether the post is malicious or not.
3. The system implements application can also be hosted online for its use and the data will be stored and fetched from server.
4. User with maximum number of maliciousposts can be blocked from the system.
5. Performance evaluation done on Dataset by using TPR, FPR, Precision, Recall and F-measure

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 12, December 2018

Proposed System Architecture:

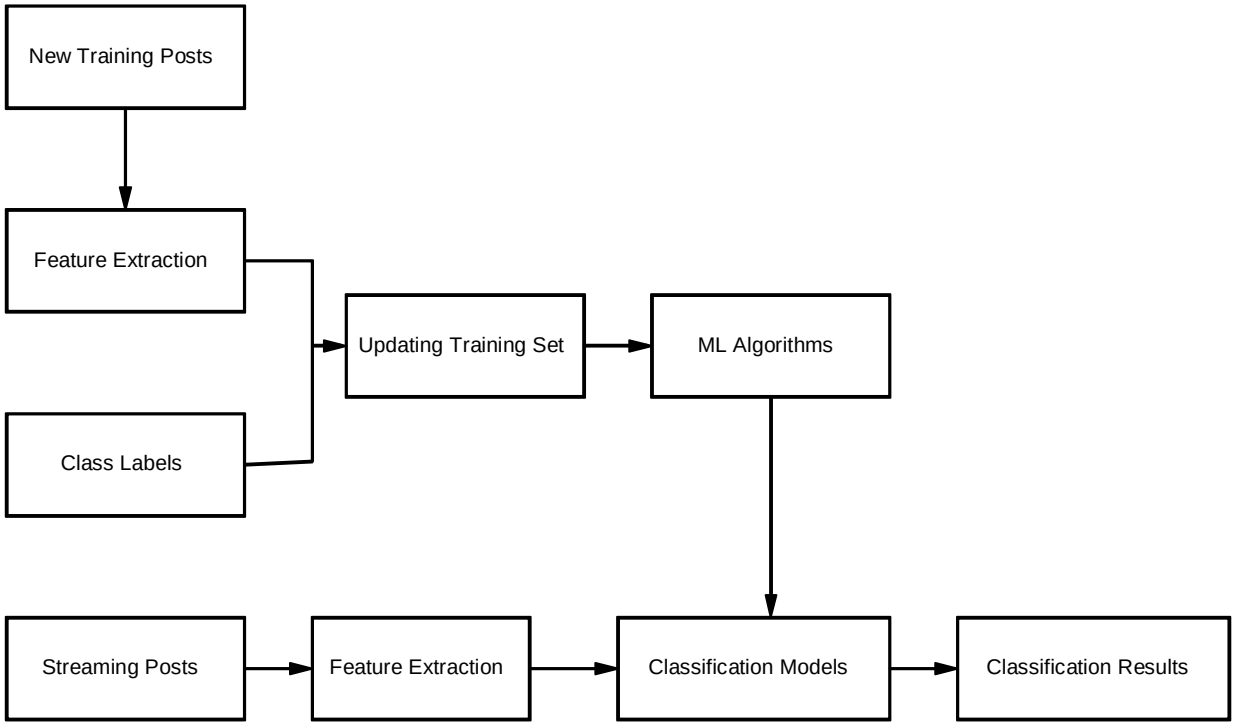

Fig 1. Proposed System Architecture

IV. CONCLUSION

In this Project, System found that classifiers ability to detect Twitter spam reduced when in a near real-world scenario since the imbalanced data brings bias. System also identified that Feature discretization was an important preprocess to ML-based malicious posts detection. Second, increasing training data only cannot bring more benefits to detect Twitter malicious postsafter a certain number of training samples. System should try to bring more discriminative features or better model to further improve detection rate.

REFERENCES

1. Guanjun Lin, Nan Sun, Surya Nepal, Jun Zhang,Yang Xiang, and Houcine Hassan"Statistical Twitter Spam Detection Demystified:Performance, Stability and Scalability" IEEE,2017.
2. 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.
3. G. Stringhini, C. Kruegel, and G. Vigna, “Detecting spammers on social networks,” in Proc. 26th Annu. Comput. Sec. Appl. Conf., 2010, pp. 1–9.
4. 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.
5. 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.
6. 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.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 12, December 2018

7. 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.
8. 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.
9. 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.
10. 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.
11. 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.