An Analytical Survey on Spam Account Detection in Twitter

1. Aliya Shaikh, IT department, AISSMS Institute of Information Technology, Pune
2. Niyanta Lad , IT department, AISSMS Institute of Information Technology, Pune
3. Simran Sayyed, IT department, AISSMS Institute of Information Technology, Pune

1

***Abstract* –Spam has been one of the biggest concerns for the majority of websites and social media outlets as it has the potential to alienate the users as it is highly frustrating and an annoying hurdle for a user. Spam is any message that is unrelated and unsolicited which is sent in bulk with the only intention to cause problems to the receiver. Spam is a very irritating concept and it is also used by people with malicious intent to cause harm by using phishing and other tactics. Spam is highly undesirable and can lead to a loss of trust between the user and the social media website as spam reduces the Quality of Experience for the user. As the central idea is to keep the user happy, spam negatively impacts this paradigm and it should be eliminated with the highest priority. This paper concerns with the various different methodologies that have been published by researchers over the years to combat spam. The related works are analyzed and scrutinized to identify their drawbacks and asses their performance. Most of the works fall short in their accuracy and this is the area that will be focused in the future researches.**

Keywords— Spam accounts in Twitter, Attacker, Machine learning, Natural Language Processing.

I. INTRODUCTION

The internet boom has changed the lives of human to a much larger degree. Before the advent of the internet, communication was primarily done through the painfully slow method of sending letters. The post office would transport the letter to the person you would like to contact, depending upon the distance between the sender and receiver, it could take somewhere between a couple of days to a couple of weeks for the receiver to get the letter. This is an extremely long wait for an information to be communicated to another person.

Therefore, a lot of researchers started working on various communication techniques that would reduce the amount of time taken for the information to be delivered. This has led to a lot of advances in the technology which has enabled faster and faster communication with the iterations each generation. The first iteration to faster communication came in the form of telegram. It required an extensive connection consisting of wires physically connecting the two remote locations together.

The telegram was one of the most innovative techniques and enabled a far better and faster means of communication. This helped a lot of people in communicating over large distances almost instantaneously. This age was also iterated when Marconi developed a technique to send information by utilizing electromagnetic waves, this meant the extensive framework consisting of wires making up the telegram network was no more needed. The waves could travel through the air and carry the message with them. This was revolutionary and subsequently paved the way for wireless communication we are aware of today.

The researches in the field of communication started progressing by leaps and bounds. This led to the development of the current wireless standards and the extremely fast and responsive optical fiber that currently forms the backbone of the internet through extensive connections to various parts of the world through gigantic Underwater fiber optic cables. The proliferation of the internet has led to a huge amount of increase in the number of webpages and their types all over the world.

This means a lot of communication can happen over the internet and over large distances almost immediately. Social media enables us to talk to loved one's thousands of miles away. But due to the ease of access, it has led to an increase in the number of cyber crimes on the internet. Most of this is due to unhealthy levels of spam found on the social media websites and public discussion forums. This is an ongoing challenge that the law-enforcement organizations are struggling to get under control.

Therefore, it is imperative to spam, which is defined as an abuse of any electronic messaging medium with unsolicited messages and irrelevant information being sent out in bulk with a malicious intent. It has been recognized since the time of the infancy of email and the attackers have improved the design to be compatible with modern messaging services. It is highly unwanted and leads to a significant number of losses that are incurred due to spam each year. Spam has the ability to completely change the landscape of the internet.

As the spam increases, it has an anility to mask the actual information and would lead to an impaired version of the internet. This would degrade the Quality of Experience for many users and would lead to decreased collaboration between individuals and ultimately lead to the internet becoming an immensely hostile place which would be highly difficult to navigate. This is the reason why spam must be controlled to keep the internet an information rich place which is very welcoming and safe for everyone.

This paper dedicates section 2 for analysis of past work as literature survey and section 3 concludes the paper with feasible statemement of the literature study.

II. LITERATURE SURVEY

This section of the literature survey eventually reveals some facts based on thoughtful analysis of many authors work as follows.

F. Concone [1] states that Online Social Networks (OSNs) have become increasingly popular both because of their ease of use and their availability through almost any smart device. Unfortunately, these characteristics make OSNs also target of users interested in performing malicious activities, such as spreading malware and performing phishing attacks. The authors addressed the problem of spam detection on Twitter providing a novel method to support the creation of large-scale annotated datasets. More specifically, URL inspection and tweet clustering are performed in order to detect some common behavior of spammers and legitimate users.

X. Wang presented a framework for finding abrupt shifts in twitter spam series. The drifted twitter spam classification technique depend on MDDT (multiscale drift detection test). The binary classification model on tweets is trained to decide whether they are spam or normal. Then K-L divergence is used for the representation of spam distribution and intuitively observed correlated drift patterns among twitter features including account age shift, the amount of followers and followings. Then, MDDT is adopted to check whether current data concepts differ from historical ones and if so, claims the drift time. Afterwards, drifted data after that time are utilized to update the model to enhance robustness. Finally, further data are input to verify performance improvement.

A. Eshmawi potrays the details of the roving proxy framework for SMS spam and SMS phishing (SMishing) detection. The framework aims to protect organizations and enterprises from the danger of SMishing attacks. Feasibility and functionality studies of the framework are potrayed along with an update process study to define the minimum requirements for the system to get used to the latest spam and SMishing trends. [3]

G. Xu elaborates that Online product review is becoming one of important reference indicators for people shopping, but the current product review site contains a lot of fraudulent reviews. Group review spamming, which involves a group of fraudulent reviewers writing a lot of fraudulent reviews for one or more target products, becomes the main form of review spamming [4]. However, solutions for group spammer detection are very limited, and due to lack of ground-truth review data, this problem has never been completely solved. The researchers propose a novel three-step method to detect group spammers based on Clique Percolation Method (CPM) in a completely unsupervised way, called GSCPM.

M. Li proposed a technique for feature extraction and the extracted features can describe comment spam effectively. For comment spam detection a gradient boosting tree classifier is combined with the extracted features. A GBT algorithm is used to generate comment spam detectors which achieve a higher accuracy.[5]

Rohit Kumar Kaliyar proposed a technique to classify the news article or other documents into certain or not. Various machine learning algorithm are explored for identification of fake news and predict the accuracy of different models and classifiers. A computational resources and models are produced for the duty of fake news detection. [6]
.

Bunyan Li proposed attention-depend LSTM-CNNs for uncertainty identification of social media texts named as ALUNI, which can indiscriminately focus on the words, regardless of cue-phrases or not, that have a decisive effect on the uncertain semantics without using extra knowledge or external NLP components [7]. The ALUNI contains three components: word representation, words encoding, and convolutional classifier. The convolutional neural networks of ALUNI capture the most important semantic information for uncertainty identification.

P. Hayati proposed a description for a new type of spamming boom called spam 2.0. Spam 2.0 is described as circulation of unsolicited, unidentified, mass content to penetrate legitimate web 2.0 applications. Spam 2.0 is different from the established spamming method since it has a parasitic nature and it's hosted on authorized web applications. One of the most famous tools utilized by spammers is web spambots. Web spambots can crawl the web, discover web applications and spread spam 2.0. [8]

A. Wijayanto introduces a fuzzy clustering approach to classify spam messages. It consists of three main phases: preprocessing, modelling, and evaluation phase. The modelling phase employs fuzzy clustering process FCM, where K-Medoid will be used as a benchmark due to its simplicity and popularity in clustering task. The evaluation phase consists of some validity index and also the evaluation of accuracy. The experimental simulation using public spam base data sets gives promising results in this process for implementing the fuzzy clustering approach in classifying spam email [9].

A. Alarifi [10] analyzed two web spam datasets, each with a different page language. First, a study of the distributions of eleven selected detection features is conducted, showing that most of these distributions vary according to the underlying language of the examined page. Using the decision tree classifier, then conducted several experiments using two, three, and four feature combinations to study the effect of the page language on the detection and false alarm rates. The experimental results showed that while there are a few common features that give almost similar results in both datasets, the performance of several other features vary depend on the language of the examined page.

III CONCLUSION

The paper has outlined the various different approaches and techniques for the detection and identification of various types of spam. Spam is highly undesirable and can be highly frustrating and irritating for a user. Spam also reduces efficiency and takes up valuable resources that could be utilized for doing something productive. Spam also tends to be utilized mainly for executing malicious intent through phishing among other attacks, which are detrimental to the user and can have a negative impact on the website. Therefore, this paper has enabled us to propose an accurate and effective spam detection technique that is based on K-Means Clustering and Linear Regression, which is assisted by the Hidden Markov Model and eventually classified through Fuzzy Classification. This proposed technique will be elaborated further in the upcoming research articles.

REFERENCES

[1] F. Concone, G. Re, M. Morana and C. Ruocco, “Assisted Labeling for Spam Account Detection on Twitter”, IEEE International Conference on Smart Computing (SMARTCOMP), 2019.

[2] X. Wang, Q. Kang, J. An and M. Zhou, “Drifted Twitter Spam Classification using Multiscale Detection Test on K-L Divergence”, IEEE Access (Early Access), 2019.

[3] A. Eshwami and S. Nair, “The Roving Proxy Framewrok for SMS Spam and Phishing Detection”, 2nd International Conference on Computer Applications & Information Security (ICCAIS), 2019.

[4] G. Xu, M. Hu, C. Ma and M. Daneshmand, “GSCPM: CPM-based group spamming detection in online product reviews”, IEEE International Conference on Communications (ICC), 2019.

[5] M. Li, B. Wu and Y. Wang, “Comment Spam Detection via Effective Features Combination”, IEEE International Conference on Communications (ICC), 2019.

[6] Rohit Kumar Kaliyar, “Fake News Detection Using A Deep Neural Network”, 2018 4th International Conference on Computing Communication and Automation (ICCCA), IEEE, 2018.

[7] Binyang Li et al, “Attention-based LSTM-CNNs for Uncertainty Identification on Chinese Social Media Texts”, Research Gate, December 2017.

[8] P. Hayati et al, “Definition of Spam 2.0: New Spamming Boom”, 4th IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2010).

[9] A. Wijayanto and Takdir, “Fighting Cyber Crime in Email Spamming: An Evaluation of Fuzzy Clustering Approach to Classify Spam Messages”, International Conference on Information Technology Systems and Innovation (ICITSI) 2014.

[10] A. Alarifi and M. Alsaleh, “Web Spam: a Study of the Page Language Effect on the Spam Detection Features”, 11th International Conference on Machine Learning and Applications, 2012.