International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 04 - Issue 05 || May 2019 || PP. 01-06

Survey on Spam detection and Anomaly detection in Social Networking Sites

Sreelakshmi K.U[1], Abey Abhraham[2]
[1,2]*Department of Information Technology,*
Rajagiri School of Engineering and Technology,
Ernakulam, Kerala

Abstract: Social networking sites, SNS are one of thesuper colossal techniques used by the people to share their views, opinions, and interest so on. The wide acceptance and the ease of usage make it more agreeable for people to develop an interest towards these SNS. Not only legitimate users make use of the opportunities of SNS but also the attackers take this as an advantage to perform their indecorous activities which is a major threat to users. Spam detection is one of the major research areas in social network security. By detecting spam tweets we can efficiently manages the SNS activities and there by enhance the security. This survey paper focuses on spam detection and anomaly detection in social networking sites or social networking services. Anomaly detection inside the social networking communities, Group anomaly detection, semi supervised spam detection and detection of spam promoting campaigns these are the main area which are being discussed here.
Index Terms: Social Networking Sites, Social Networking Communities, Anomalies, Spam Detection

I. INTRODUCTION

The wide emergence of social media leads to a considerable growth in the usage of the social networking sites. This promotes an increase in communication between millions of people without any barrier of distances. Nearly every people in the world have profile on any of the social networking sites like Twitter, Facebook, Linkedln, Google+ etc. The psychology behind this is that it is difficult to connect with people geographically but it is easier to connect digitally. The risk in social networking sites is entirely depending on the amount of information in which people share. It is clear that if the users share more information without considering the privacy and security then it may leads to a great vulnerability.

The users of online social networks can create small virtual group inside the network. These virtual structures are called social network communities. Members of a social network community may not be known each other. They come under one cluster or community because of similar interest, opinions, views etc. In this scenario the trust is the major problem. Because there can be chance for spammers or anomalous people within the members of the group. Initially they may act as trusted users and make this trust as an opportunity to perform unacceptable activities which will affect risk factor of using social networks. So it is also important to detect spamming activities inside the social network communities. This study is all about different anomaly detection methods and spam detection in social networking sites.

Social networks can be represented as a graph consists of vertices and edges where vertices or nodes are the users and edge shows the relationship between them. Among these nodes some may possess unusual behaviour when compared to other nodes. These nodes are called anomalies or anomalous nodes. That is something that deviates from the standard behaviour or normal expectation is the anomalies. Anomalous users refer to the people who are deviating from the normal user behaviour. Initially the anomalous nodes behaves like a normal legitimate user but after gaining the trust and acceptability it starts performing unlawful activities which leads to serious security threats. Detection of anomalous activities is one of the key areas in the research of social network analysis.

Irrelevant or uninvited messages sent over the Internet which aims to reach typically a large number of users, for the purposes of advertising are known as spam messages. Nowadays there is a considerable increase in the growth of usage of SNS. The high click rate and the effective message propagation make social media as an attractive platform for spammers. Increase in spamming activities affects the people who are using social media adversely. So detection of anomalies and spam messages have equal importance in social network analysis. Most of the existing methods deal with the detection of anomalous users or spam nodes. But it is not a efficient method. Because the attackers can create multiple account and continue performing malicious activities.

II. ANOMALY DETECTION

An anomaly in social network is the unusual or abnormal activities of nodes when compared to other nodes in the network. It shows different kind of action when compared to

other nodes [1]. By definition anomaly detection aims to identify the suspicious nodes which exhibits malicious activities inside the networks. According to [2] anomalies or unusual activities can be classified into various categories like based on nature of anomalies, based on nature of graph structure, based on information available in the graph structure and based on behavior. Point anomalies, Contextual anomalies, Collective anomalies Horizontal anomalies, Labeled anomalies and unlabelled anomalies are some of the examples for these categories.

A. Anomaly Detection Methods

In social networks the node which doesn't obey the rules or similarity measures like centrality, betweeness, efficiency

[2] and deviates from normal behavior is need to detected in order to make the system more secure. Anomaly detection methods mainly classified as follows,

∑ Behavior based techniques ∑ Structure based techniques. ∑ Spectral based techniques.

Structural Based Techniques: These types of methods work on the principle of using structural properties which can be used to check the attributes of legitimate and anomalous users. Graph metric is calculated for different nodes in the network and nodes which show different structural values are considered as anomalous nodes or anomalous users. This is similar to supervised approach because here also there will be Behavior Based Techniques: In behavior based techniques it handles the behavioral properties of users of social network. Behavioral properties includes number and content of message, number of shared message content, likes and comments on a post, duration of conversation etc.

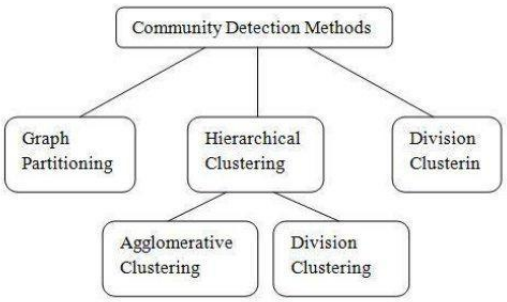
A predefined set of normal patterns which is already known and any deviation from that known patterns depicts the abnormal behavior or anomalous users. Structural based techniques can be used for dynamic anomaly detection.

Spectral Based Techniques: Social network can be represented as a graph consists of nodes and vertices which are described earlier. The node can have spectral characteristic in the spectral space of graph. Spectral based techniques detect anomalies by using these spectral properties

B. Anomaly Detection in Social Network Communities

Social network communities are small and dense groups which are created inside the social networks. These are virtual clusters which may have similar behavior. People who have similar interests connect with other people and create groups. The half of the people inside the communities doesn't have any face-to face interaction. So trust is the major concern to participate in these communities. At first the attackers acts as normal users. That is they spoof as legitimate users. After gaining the trust and acceptance from other members they start doing malicious activities such as spamming activities. Anomalous users in the communities are the one who shows any deviation from the normal behavior of legitimate users.

Community detection can be partitioned into three. [4]



There exist many community detection algorithms such as memory label propagation algorithm, Girwan-Newman algorithm, Clique Propagation algorithms, Greedy local algorithm etc. Depending on the application and structure of network any of these algorithms can be used to detect social network communities.

Anomalous behavior in these social network communities need to be detected in order to avoid the risk factors of using online social networks. False trust is the major issue here. One of the method identifies anomalies by plotting fitting curve and any deviation from the curve that means the nodes which is far from the fitting curve is considered as anomalies[1]. Degree and power of each node is calculated and which is used for plotting the fitting curve. Degree of a node is the number of connections to that node. Power is calculated using Political Independence Index, PII using the following equation,

$$\sum_{i=0}^{k} \beta^{i} [P(i)^{x} - N(i)^{x}]$$

β is the attenuation factor and P(i) and N(i) are the positive and negative distance of a node. Anomalous nodes are identified by calculating the distance from fitting curve. Usually anomalous nodes lie far away from the fitting curve.

C. Group Anomaly Detection

Anomalies may appear not only as an individual point, but also as a group [3]. In order to achieve a common goal the attackers create groups and transmit malicious data such as false product reviews or threat campaigns. This kind of anomalies referred to as group anomalies.

There are three major challenges in group anomaly detection. (i)Point-wise data and pair-wise relational data are the two forms of data exists in social media. The information which describe the properties of individual user is called point-wise data where as other describes the properties of social ties. It is important to consider both data during anomaly detection. (ii)Since the group anomalies are occasional when compared to point anomalies. Therefore most of the algorithms usually fail to detect group anomaly detection. (iii) User activities are dynamic in nature. Communications change constantly over time and it is very hard to determine groups beforehand.

II. SPAM DETECTION

Irrelevant or unsolicited messages which send over the internet in order to destroy the normal communication are referred to as spamming messages. Usually spam messages are sent as bulk messages which target a large number of users. Spamming messages need not be sent by human individuals it can be generated from the third party tools such as machines. Because it is difficult for an intruder to manage multiple account and send these kinds of message. This is done for avoiding early detection of spammers.

A large number or variety devices such as mobiles, laptops, tablet computers desktop computers comes into popular there by use of social networking is also increased. Spamming can be spread to new technologies rapidly. Micro-blogging services like twitter became a prominent platform for many activities like campaigning. Spam promoting campaigns are need to be detected.

A. Detection of Spam Promoting Campaigns.

Recently large numbers of campaigns which contain spam contents or promotion URLs are emerged. [5] Spam promotion using campaigns are more dangerous when compared to individual spam since they can infect more users. Twitter messages are short (140 characters long) so attackers have a tendency to insert URLs inside the tweets. This is used to reach more number of users.

Initially post normal or informative URLs in order to gain the acceptability. After that they post URLs which may direct to phishing website or malicious content which is very harmful for legitimate users. Large numbers of URLs are continuously sent over the social network in order to reach maximum users and to gain maximum visibility from users. Traditional methods identify spam nodes or accounts and block those from the network. But it is not an efficient method to control spamming activities. Because the attackers can create multiple account and continue the spamming activities.

There can be a group of accounts which manipulates these promotion activities which is referred to as promoting campaigns. It can be spamming promotion campaigns or normal campaigns. Number of methods is available for spam detection. One method is described here.

Account who post URLs similar purposes are identified first. Then extract candidate campaigns. Finally classifies them as normal, spam and promotion campaigns. URL method is one of the best methods because insertion of URLs inside the tweets is the one way to make it more convenience for people to visit. The URLs which are posted can be fall into two extreme opposite categories. Whether it can be spam or it is original URL of promotion activities. The duration between the similar URL posts is also considered. It is because the URLs from spam or promotion are posted with in shorter time interval. This is due to the fact that it is difficult for spammers to handle accounts manually so they will use tools such as botnet or some other programs to auto-

post the URLS. The programs post the URL may be at the same time or within the shorter time interval. Shorter the interval means it is from accounts who have similar purpose of posting the URL.

The amount of information in one URL can be finding using Shannon's Information Theory [5]. Using this values similarity between two accounts can be calculated. The similarity measure always lies in between 0 and 1. Finally plot an account graph consists of nodes and edges. An edge between two nodes or accounts exists if they are similar. Assume that these accounts are promoting campaigns and evaluate the URLS posted by them and using any suitable machine algorithm like, Decision tree, Naïve Bayes, Random Forest etc. the URLs can be classified into normal or un-normal and again un- normal URL s can be further classified into promotion or Spam.

B. Semi Supervised Spam Detection in Twitter Stream.

Semi Supervised Spam Detection (S^3D) is a framework to detect spam at tweet level. [6] This method consists of two main modules. Spam Detection Module and Model Update Module

Spam detection module operates in real time mode and model update module operates in batch mode. Spam detection module consists of 4 light weight spam detectors.

∑ Blacklisted URL detector ∑ Near Duplicate detector ∑ Reliable Ham detector

∑ Multi -classifier based detector

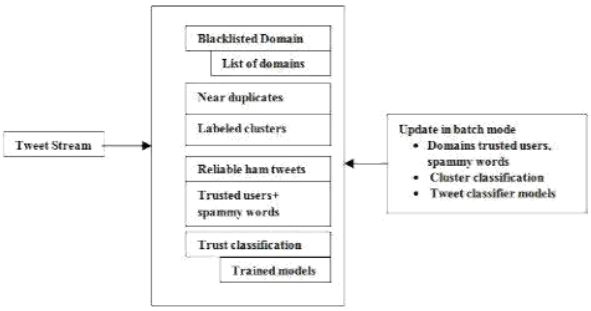


Fig 1. System overview of the S^3D framework

Blacklisted Domain Detector: This checks the URLs or links in the tweet stream is from a black listed domain or not. Spammers promote their services by spreading tweets containing links which is from a blacklisted domain. These kind of URLs are typically shortened URLs. Using java library HttpClient this shortened URLs can be convert into full expanded URLs. This is the dataset for these domain detectors.

Near duplicate detector: There exists a predefined set of clusters to detect the tweets are spam or ham. Ham tweets are nothing but it is the tweets posted by trusted users. If a tweet is near duplicate of other tweets classification can be done using this pre labeled clusters.

Reliable Ham detector: For tweet is said to ham if it satisfies two conditions. First, it should be posted by a trusted user. Second it should not contain any spammy words such as *followme.* Trusted users are those users who never posted anytweets containing spammy words and he/ she should post at least five confident tweets. Confidence tweets are the tweet doesn't contain spammy words. Suppose the total number of tweets containing word *w* to be *n(w),* the number of spam tweets containing the word *w* to be $n_s(w)$ and the number of ham tweets containing word *w* to be $n_h(w)$. The probability of words appearing in spam tweets $p_s(w)$ and the same for ham tweets is $p_h(w)$. Then $p_s(w)= (n_s(w)/n(w))$ and $p_h(w)=(n_h(w)/n(w))$. The word *w* is a spammy word if andonly if $p_s(w)>p_h(w)$.

Multi- classifier Based Detector: If a tweet is recognized by any of the above detector is need not be pass to the next level detector. Tweets that are not labeled by any of the detector that is none of the three detector is failed to label the given tweet then by using suitable machine learning algorithm that tweets can be

classified into spam or ham. So here a supervised learning method is applied. The features for tweet representation are hash tag information, fraction of words

containing spammy words, user information and URL information like URL from top 100 domains.

METHOD	ADVANTAGES	DISADVANTAGES
Anomaly detection in social network communities	Efficient method for large networks	Complexity and errors can be occur due to multiple iteration of algorithm and time complexity is more
Group anomaly detection	Group anomaly detection reduces risk than that of individual anomaly detection	Group anomalies are difficult to detect and more time complexity
Detection of spam promoting campaigns	Spam promoting campaigns are efficiently controlled inside social network with more accuracy	More attention is required because small errors may label real campaigns as spam promoting one
Semi supervised spam detection	Multiple spam detectors are available hence more accuracy	More time complexity

TABLE 1: Comparison between existing anomaly and spamdetection methods.

III. CONCLUSION

Anomaly detection and spam detection are one of the key research areas in social network analysis. Nowadays the people who are using social media are very large. Especially youngsters use social media as a best and most convenient platform to share their information. That is they give less importance to their privacy. Even though SNS uses new technologies and security policies there is no any guarantee that our information is safe. When new technologies are emerged not only legitimate users or researchers take this as an opportunity but also attackers utilize this. This paper is mainly divided into two sections. One for anomaly detection and the second one is spam detection. Latest methods are described in each section.

REFERENCES

[1]. Majunatha H C, Dr. R Mohanasundaram, "BRANDS: Big data Real-Time Node Anomaly Detection in Social Networks" 2nd International Conference on Inventive Systems and Control-2018
[2]. Sarbjeet kaur, Prabhjot Kaur "Review of different types of Anomalies and Anomaly detection techniques in Social Networks based on Graphs" International Journal of Computer Trends and Technology (IJCTT) – Volume 47 Number 2 - May 2017
[3]. Rose Yu, Xinran HE and Yan LIU, "GLAD: Group Anomaly Detection in
[4]. Social Media Analysis. ACM Transactions on knowledge discovery from data, Vol 10 No 2 Article 18, October 2015

[5]. Ruby, Dr Inderjeet Kaur "A Review of Community Detection Algorithms in Signed Social Networks". International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017)
[6]. Xianchao Zhang, Zhaoxing LI, Detecting Spam and Promoting Campaigns in twitter, ACM Transactions on the web, January 2016
[7]. Surendra Sedhai ,Aixin Sun, Semi-Supervised Spam Detection in Twitter Stream, IEEE Transactions on Computational Social Systems, March 2018.