International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 05 | May-2016 www.irjet.net p-ISSN: 2395-0072

Review on Spam Detection in OSN using Integrated Approach

Bhagyashri Toke[1], Dinesh Puri[2]

[1] *Research Scholar, Department of Computer Engineering, SSBT's COET, Maharashtra, India*
[2] *Assistant Professor, Department of Computer Engineering, SSBT's COET, Maharashtra, India*

Abstract - *Online Social Networks (OSNs) are to a great degree well known among internet users. Tragically, in the wrong hands, they are likewise powerful for executing spam companion. It present an online spam filtering framework that can be conveyed as a segment of the OSN platform to access messages produced by users at real-time. The proposed system is to classify and analyze spam messages into spam and genuine messages separately. It gives an enhanced integrated approach measure by using a Bayesian Approach, List Filtering, Pattern Matching and Rule Based Filtering for detection of a spam messages in online social networking website. Thus, for the most part propose framework permitting OSN users to have an immediate control on the messages posted on their walls. It is conceivable to accomplish through an exploratory assessment a computerized framework known as Filtered Wall (FW) which has capability to filter undesirable spam messages from OSN user walls on the premise of both message content and the message poster connections and attributes of it.*

Key Words: Online Social Networks, Spams, Filtered Wall, List Filtering, Rule based Filtering, Bayesian Approach, Pattern Matching.

1. INTRODUCTION

OSNs become a new medium for dissemination of information, at the same time, they are also a playground for the spread of misinformation, fake news. OSN platform have two kinds of users namely, Spammers and Non-Spammers. Spammers, out of malicious intent, post either unwanted information or spread misinformation on OSN platforms. Such activities disturb the ham users, called Non-Spammers and also decrease the reputation of OSN platforms [1]. Spam is abuse of electronic messaging system to send unrequested bulk messages. Today large sizes of spam messages are causing serious problem for the users and internet services. Such as, It break down user search experience, It causes spreading of virus in network, It enlarge load on the network traffic, It expends carelessly the resources such as bandwidth, storage, and computation power, It wastes the user time and energy.

1.1 Spam Filtering Architecture

Spam filter reduces the amount of junk email. Email filtering is the processing of emails to arrange it according to specified criteria. Common use of mail filters is to arrange incoming mail, removing of spam emails, removing of computer virus.

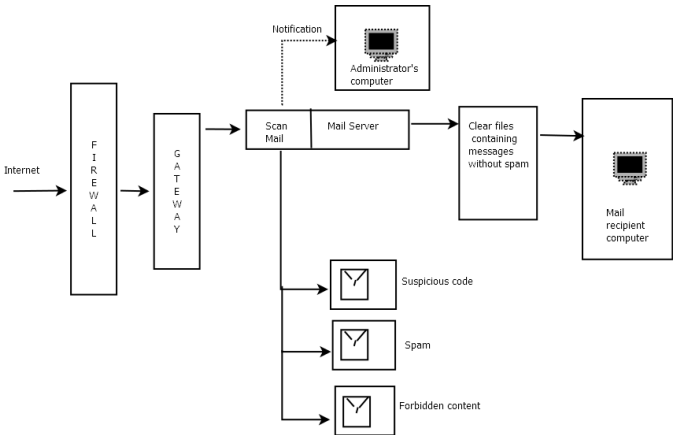

Fig -1: Spam Filtering Architecture

The above figure shows the typical architecture of spam filter. Spam filter applied at all layers, firewall exist just before of email server or at MTA(Mail Transfer Agent), Email server to give an combined anti-spam and anti-virus solution providing perfect email protection at the network perimeter level, before undesired or potentially unsafe email reaches the network. Spam filters can be installed as a service to all of their customers at MDA (Mail Delivery Agent) level also [2]. At email client user can have an individualized spam filter that then automatically filters mail according to the selected criteria.

1.2 PROBLEM STATEMENT

Online Social Network (OSN) is operated by millions of people to share information, But it also can share the fake news. To avoid the unwanted content to be display, OSNs is to permit users the ability to control the messages posted on their own private space. It will reduce the performance of server and waste of time in detecting messages manually.

It will also minimize the accuracy to detect the spammers and non-spammers. To detect spammers and non-spammers using the integrated approach, which improve the accuracy of detection of spammer and non-spammers in OSN as compared to existing systems.

2. LITERATURE SURVEY

Spam detection in OSN is based on different techniques. There are different spam filtering techniques such as clustering, SVM based spam Filtering, Bayesian spam filtering, Decision tree based filtering.

2.1 Background

All current OSN acquire the client-server architecture. The OSN service provider performs as the controlling entity. It stores and handles all the content in the system. Furthermore, the content is generated by users easily from the client side. The users can interact with others through a rich set of well-defined interfaces given by OSN service provider. Currently two famous ways of interaction exist. Facebook is an example of OSNs that acquire the interaction between a pair of sender and recipient as their primary way of interaction, although they also support other ways. Twitter is an example of OSNs that acquire broadcasting as their primary way of interaction.

2.2 Related Work

Konstantin Tretyakov in [3] have briefly described that clustering is basically an unsupervised learning method. Unlike Naive Bayes, a separate training data need not be produced for this algorithm. On the basis of same feature values (like similar kind of reply trend, similar usage of spam words in tweet), this algorithm could classify the whole set of accounts into two classes. One of this class was labeled as Spammer and another as Non-Spammer.

A. Niimi, H. Inomata, M. Miyamoto and O. Konishi in [4] have presented that support vector machine classification filter method for classifying whether a mail is spam mail or regular mail is as follows. The procedure separates into pre-processing (filter learning) and the classify processing (filtering).

- **Pre-processing(filter learning):**

1. Collect the spam mails and regular mails.
2. Separate all the mail to tokens.
3. Calculate appearance frequency of each token.
4. Define token code.
5. Make vector sets with the token code and its appearance frequency.
6. Construct filter(classification rules) by learning filter by SVM with vector sets and
7. label (the mail is spam mail of regular mail).

- **Classify processing(filtering):**

1. Separate a mail to tokens.
2. Make vector with token code and its appearance frequency.
3. Using SVM filter and this vector, classify the mail.

S. Chakraborty, B. Mondal in [5] have described the decision tree generated by C4.5 can be used for different classification problems. At each node of the tree the algorithm selects an attribute that can further split the samples into subsets. Each leaf node represents a classification or decision. At every level of the tree, the tree structure was prepared and the decision was made.

These decisions were made on the basis of feature values. The Twitter account to be classified was designated as root node and then after passing through a series of decisions, this node finally reaches at the leaf node, which then decides its class.

Jeremy Eberhardt in [6] has presented that Bayesian spam filter is considered to be a more advanced form of content based filter, uses the laws of mathematical probability to check which messages are regular and which are spam. The Bayesian filter technique for classifying whether a mail is spam mail or regular mail is as follows. This technique separates into pre-processing (filter learning) and the classify processing (filtering).

- **Pre-processing(filter learning):**

1. Collect the spam mails and regular mails.
2. Separate all the mail to tokens.
3. Calculate probability of spam for each token.
4. Store probability of spam in database.

- **Classify processing(filtering):**

1. Separate a mail to tokens.
2. Query probability of spam for the token.
3. Remove characteristic tokens, calculate composite probability.
4. If the composite probability is greater than establish threshold, this mail is classified as spam mail.
5. If the composite probability is lower than establish threshold, this mail is classified as regular mail.

3. PROPOSED WORK

The first step for the detection of spam messages in any OSN is data collection and necessary preprocessing to convert it into a form, which can be used by the learning algorithms. There are various different classification algorithms, which can be used to classify as spam or genuine message. The figure shows an overview of the complete process of spam detection. In which the preprocessed data is first classified using different learning algorithms to guess the spam or genuine message of all Facebook users. Fig-2 shows proposed approach of spam detection.

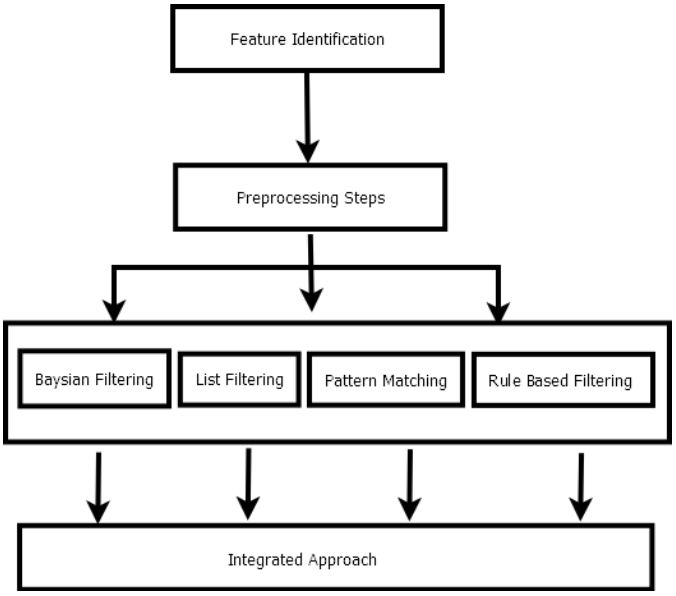

Fig -2: Proposed Approach of Spam Detection.

3.1 BAYSIAN FILTERING

Bayesian filter learn from both good and spam messages, result in effective anti-spam approach. Bayesian email filter take the advantage of Bayesian theorem is

P (spam/word) = [P (word/spam) P (spam)] / p (word)

Theorem in the context of spam, says that the probability that an email is spam, given that it has particular words in it, is equal to the probability of finding those particular words probability that any email is spam, divided by the probability of finding those words in any email.

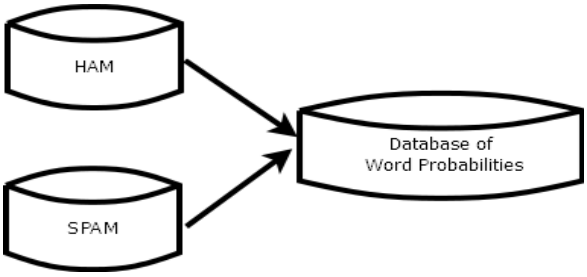

Fig -3: Creating Database for Filter

Fig-3 shows creating word database for filter. Particular words have particular probabilities of occurring in spam message and in legitimate message. The filter doesn't know these probabilities in advance, and must first be trained so it can build them up. To train the filter user must manually specify whether a new email is spam or not. For all words in each training email, the filter will adjust the probabilities that each word will appear in spam or legitimate email in the database of it.

3.2 LIST FILTERING

List based filter try to stop spam messages by categorizing senders as spammers or trusted users, and blocking or allowing their messages according to it.

3.2.1 Blacklist

This popular spam-filtering method attempts to stop unwanted email by blocking messages from a preset list of senders that you or your organization's system administrator creates. Blacklists are records of email addresses or Internet Protocol (IP) addresses that have been previously used to send spam. When an incoming message arrives, the spam filter checks to see if it's IP or email address is on the blacklist; if so, the message is considered as spam and rejected.

Though blacklists ensure that known spammers can't reach user's inboxes, they can also misidentify truthful senders as spammers. These so-called false positives can result if a spammer happens to be sending junk mail from an IP address that is also used by truthful email users. Also, since many clever spammers regularly change IP addresses and email addresses to cover their paths, a blacklist may not instantly catch the newest outbreaks.

3.2.2 White list/Verification Filter

A whitelist blocks spam using a system almost entirely opposite to that of a blacklist. Rather than letting you specify which senders to block mail from, a whitelist lets you specify which senders to allow mail from; these addresses are placed on a trusted-users list. Most spam filters let you use a whitelist in addition to another spam-fighting feature as a way to cut down on the number of genuine messages that unexpectedly get indicated as spam. However, using a very accurate filter that only uses a whitelist would mean that anyone who was not appreciated would automatically be blocked.

Some anti-spam applications use a modification of this system known as an automatic whitelist. In this system, an unknown sender's email address is checked against a database; if they have no history of spamming, their message is sent to the recipient's inbox and they are added to the whitelist.

3.3 PATTERN MATCHING

A spam filter is a program that is used to detect unrequested and unwanted email and prevent those messages from getting to a user's inbox. In this spam filter checks all incoming emails to your accounts against mail filter rules. Outline the tags and automatic parsing of the HTML tags and read the data from each tuple. Compare the tags with the keywords and URL's located in our repository using top down parsing keyword search algorithm. The email will be identified as the spam email if the data in the tags are matched. There is no particular algorithm for statistically determining whether or not a given e-mail spam message. The obvious method for pattern matching is just to check, for each possible position in the text at which the pattern could match, whether it does in fact match.

3.4 RULE BASED FILTERING

All these applications depend on a rule-based filtering procedure. This procedure consists in locating specific terms that are used to mention the concept that is being requested. A filter uses a set of complex rules, developed by an administrator, which make it possible to limit the semantic scope of terms. To develop a rule base an administrator has to find a trade-off between user's interests and actuality of corpus. There is a way to define rules of extreme precision that depend on particular syntactic relations or interior structure of documents. This filtering procedure uses a limited number of rules. These rules mention to a pattern involving a restricted set of words that have been selected by the administrator.

4. CONCLUSIONS

The Integrated approach provides more efficient spam detection. To detect spam messages and genuine messages, It gives the improved accuracy means it give the best results in terms of accuracy. In future different techniques will be combined to improve the performance of existing techniques.

REFERENCES

[1] A. Gupta, and R. Kaushal, "Improving Spam Detection in Online Social Networks," IEEE Cognitive computing and information processing, March 2015 , pp. 1-6.

[2] V.Christina, S.Karpagavalli, and G.Suganya, " Email Spam Filtering using Supervised Machine Learning Techniques," International Journal on Computer Science and Engineering, Volume 02, No.09, 2010, pp. 3126-3129.

[3] Konstantin Tretyakov, "Machine Learning Techniques in Spam Filtering," Data Mining Problem-oriented Seminar, May 2004, pp. 60-79.

[4] A. Niimi, H. Inomata, M. Miyamoto and O. Konishi, "Evaluation of Bayesian Spam Filter and SVM Spam Filter," May 2011.

[5] S. Chakraborty, B. Mondal, "Spam Mail Filtering Technique using Different Decision Tree Classifiers through Data Mining Approach - A Comparative Performance Analysis," Volume 47, No.16, June 2012.

[6] Jeremy Eberhardt, "Bayesian Spam Detection," UMM CSci Senior Seminar Conference, December 2014.