IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 21-24

Content Based Spam Filtering By Using Svm(Support Vector Machine)

Harshada Sonone[1],Neha Khedkar[2], ,Neha Patle[3] Prutha Khadse[4] Anjali Bhagat
UG Student, Dept. of Studies in CSE, TGPCET College of Engineering, Nagpur, Maharashtra, India1 Assistant Professor, G.Rajesh Babu Dept. of CSE,

Abstract: *The traditional anti-spam techniques like black and white list is not up to the mark in current scenario. The goal of spam classification is to distinguish between spam and legitimate mail message. But with the popularization of the Internet, it is the challenging to develop spam filters that can effectively eliminate the increasing volume of unwanted mails automatically before they enter a users 's mailbox .Many researchers have been trying to separate spam from legitimate emails using machine learning algorithms based on statistical learning methods. In this paper we evaluate the performance of non linear SVM based classifiers this algorithm classify the spam or not spam based on content of email .*
The experimental result show that this algorithms could improve the spam filtering accuracy rate effectively.
Keywords: *spam,support vector machine,machine learning ;*

I. Introduction

Email system is a standout amongst the best and regularly utilized sources of correspondence. The reason of the prevalence of email system lies in its financially savvy and quicker correspondence nature. Unfortunately, email system is getting compromised by spam messages. Spam messages are the excluded messages sent by some unknown users also called spammers with the intention of profiting. The email users invest the greater part of their valuable time in arranging these spam mails. Numerous copies of same message are sent commonly which influence an organization financially as well as bothers the getting users. Spam messages are barging in the client's messages as well as creating vast measure of undesirable information and consequently influencing the system's ability and utilization. In this paper, a Spam Mail Detection (SMD) system is proposed which will arrange email information into spam and ham messages. The procedure of spam sifting centers around three primary dimensions: the email address, subject and substance of the message.

Spammers are generally technically skilled persons that are hired by companies for sending spam. A third party is hired to prevent any legal action on the company itself. Spamming activity can cost attractively to a company, if done right.

E-mails are quick and cheap method for data sharing and correspondence in this day. Perusing inbox E-mails turns into the regular habit for the peoples. Email containing undesirable content irritates the user and possesses the half of the transfer speed of the inbox. These Emails are recognized as spam. The issues of spam mails are a horrid issue. Email spam alludes to sending different, erroneous and unconstrained email messages to various clients. The motivation behind these sends is attention, headway and dissipating indirect accesses or pernicious programs. The time spends by individuals in perusing and erasing the spam mail is waste. A spam mail can't just irritating yet in addition hazardous to beneficiaries. Tapping on connections contained in spam messages may send client to phishing and malware.

A spam mail cannot only be annoying but also dangerous to recipients. Clicking on links contained in spam emails may send user to phishing and malware. Machine learning approach has been widely studied and there are lots of algorithms can be used in e-mail filtering. They include Naïve Bayes, support vector machines, Neural Networks, K-nearest neighbor, Rough sets and the artificial immune system.

Naïve Bayes classifier is based on Bayes theorem with an assumption of strong independence. The classifier is a probability based classifier which computes the class probabilities of the given instances. The probability set is calculated by computing the combinational and frequency values of the data set. The class probability which is nearest to the rear end will be picked by the classifier. The Naïve Bayes classifier is a multiclass classifier and works efficiently with supervised learning approach.

E-mail Spam

Spam is the use of electronic messaging systems (in-cluding most broadcast media, digital delivery systems) to send unsolicited bulk messages indiscriminately. In this article it is considered the e-mail spam. E-mail spam, also known as junk e-mail or unsolicited bulk e-mail (UBE), is a subset of spam that involves nearly identical messages sent to numerous recipients by e-mail. Day by day the amount of incoming spam increase and, scammer attacks are becoming targeted and consequently more of a threat. When targeted attacks first

emerged five years ago, Symantec MessageLabs Intelligence tracked between one or two attacks per week. Subsequently, attacks have increased further from approximately 10 per day to approximately 60 per day in 2010 (Figure 1). By the end of 2010 MessageLabs Intelligence identified approximately 77 targeted attacks block- edeach day [1].

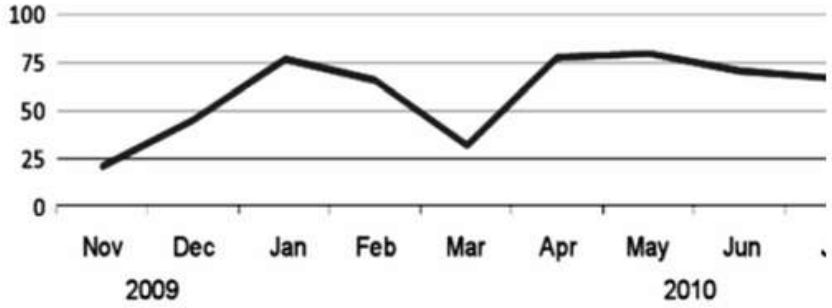

Figure 1.1 Targeted spam attacks 2009-2010 (Symantec. cloud Message Lab) [3].

By the Symantec MessageLab forecast spam will be- come more culturally and linguistically diverse, in 2011. The amount of spam sent from European countries will increase to 40% - 45% of all spam [1]. These facts statethat spam is a big problem for today and for tomorrow and it actually makes sense to investigate new effective methods against spam.

II. Background

In the literature work, various spam detection techniques are introduced. In linguistic approach natural language processing technique is used to identify similarity among multiple reviews. Feng et al. [3] uses n-gram and their composition. Some studies [2][4] Language modeling also include study for features between multiple reviews like capital words in statements. Lai et al. [5] proposes the probabilistic language modeling technique to find similarity between multiple reviews.This technique is based on metadata analysis of areview. Metadata includes user behavior and review behavior analysis. Feng et al. [6] proposes a technique that studies metadata of review based on distribution of user rating on different products. 36 different behavior analysis techniques are proposed by Jindal et.al [7] with supervised learning mechanism. [11] Indicates behavioral features show spammers` identity better than linguistic ones. Fi. Al [12] proposes machine learning method to identify spam reviews. Paper [13] investigates syntactic stylometry for deception detection.

Network based algorithms can be applied for spam detection. In this techniques heterogeneous network is established between reviews and users. Fei et al. in [8] proposed a network based Loopy Belief Propagation (LBP) algorithm to find burstiness in reviews to find spam reviews. Li et al. in [10] proposes a technique to analyze a review from multiple users from same IP address. For this heterogeneous network is established between users, reviews and user IPs.

The study of all categories is done independently. Netsapm[1] is the technique proposed by Saeedreza Shehnepoor, Mostafa Salehi, Reza Farahbakhsh, and Noel Crespi. In this technique simultaneous study of Behavioral (RB) Based, Linguistic (RL) Based and graph based approach is proposed. EuijinChoo, Ting Yu , and Min Chi [9] detects the spammer groups in review systems. This is done using sentiment analysis on user interactions and graph theory. It analyses user relationship graph and annotating the graph by sentiment analysis and then pruning is done. According to the studies in literature, a common platform is required that make thestudy of

III. Existing System

In the existing system, the mails are sent to the authenticated users who are intended to be received.This survey has explored almost all published fraud detection. It definesthe adversary, the types and subtypes of fraud, the technical nature of data, performance metrics, and the methods and techniques. After identifying the limitations in methods and techniques of fraud detection, this paper shows that this field can benefit from other related fields. Specifically, unsupervised approaches from counter terrorism work, actual monitoring systems and text mining from law enforcement, and semi supervised and game-theoretic approaches from intrusion and spam detection communities can contribute to future fraud detection research.

Disadvantages existing system are:

- Suspicious mails cannot be detected.
- Offensive users cannot be identified.

IV. Proposed System

In the proposed system the suspicious users are detected and the offensive mails are blocked. The Proposed system initially extracts some useful features such as "Suspicious keywords" and "non-suspicious indicators" from the e-mail message .Then the combination of those keywords and indicators are analysed. If suspicious keywords are present in an e-mailwithout any non-suspicious indicators ,the email will be detected as suspicious and the threat of a potential future terrorist event will be reflected .

V. Methodology Algorithm

K-Nearest Neighbor Classifier Method

The k-nearest neighbor (K-NN) classifier is consideredanexample-basedclassifier,that means that the training documents are used for comparison ratherthananexplicitcategory representation, such as the category profiles used by other classifiers. As such, there is no real training phase. When a new document needs to be categorized,thekmostsimilardocuments (neighbors) are found and if a large enough proportion of them have been assigned to a certain category, the new document is also assigned to this category, otherwise not. Additionally, finding the nearest neighbors can be quickened using traditional indexing methods. To decide whether a message is spam or ham, we look at the class of the messages that are closest to it. The comparison between the vectors is a real time process. This is the idea of the k nearest neighbor algorithm: Stage1. Training Store the training messages. Stage2. FilteringGiven a message x, determine its k nearest neighbours among the messages in the training set. If there are more spams among these neighbours, classify given message as spam. Otherwise classify it as ham.

The use here of an indexing method in order to reduce the time of comparisons which leads to an update of the sample with a complexity O(m), where m is the sample size. As all of the training examples are stored in memory, this technique is also referred to as a memory-based classifier [6]. Another problem of the presented algorithm is that there seems to be no parameter that we could tune to reduce the number of false positives. This problem is easily solved by changing the classification rule to the following l/k-rule:

If l or more messages among the k nearest neighbors of x are spam, classify x as spam, otherwise classify it as legitimate mail.

The k nearest neighbor rule has found wide use in general classification tasks. It is also one of the few universally consistent classification rules.

Support Vector Machine

Support Vector Machines are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships, the SVM modeling algorithm finds an optimal hyperplane with the maximal margin to separate two classes.

Input: sample x to classify training set T. T = {(x1,y1),(x2,y2),……(xn,yn)}; number of nearest neighbours k.

Output: decision yp ∈ {-1,1}

Find k sample (xi,yi) with minimal values of K(xi,xi) – 2 * K(xi,x)

Train an SVM model on the k selected samples

Classify x using this model, get the result yp Return yp.

VI. Conclusion & Further Work

Conclusion

This project proposes a fast content-based spam filtering with SVM classification algorithm. Unlike other spam filtering algorithm that k-means clustering algorithm is used to distinguish spam and normal mail i.e. it identify emails as spam or not spam. Where, this algorithm is use to compress data according to the similarity rules. Then support vector machine is used to train the classification model, and this model is better to deal with uncertain factors. This project model could improve the spam filtering algorithm from two aspects of reducing time consumption and increasing accuracy rate. We reviewed content-based spam filtering techniques based on Machine Learning methods propounded so far, highlighting the main ap- proaches and advancements gained by the approach. A quantitative analysis of the major reviews over the last decade was conducted. Overall the number and quality of literature demonstrates that remarkable advance- ments have been achieved and continue to be achieved. However some outstanding problems in e-mail spam filtering as highlighted above still remain. Till more improvements in spam filtering happen, anti-spam re- search will remain an active research area.

Further Work

Algorithm (GA) have been well incorporated in order

to increase or enhance the accuracy of spam detection (Mohammad and Zitar, 2011; Wu et al., 2011). The majority of researches focus on using spam methods to avoid the spam e-mail completely from entering the user's mailbox. Therefore, very difficult to prevent the spam mail from entering the user's mailbox due to the spammers changed their techn methods are not able to reduce overhead, bandwidth, processing power, time and memory used by spam (Tala Tafazzoli, 2009). Classification is very important and popular method to solve the problem of spam (Attri and Kaur, 2012).

The main objective and the output that proposed system outcome will contain is listed below:

- Preprocess dataset
- Perform Clustering
- Perform Training of dataset
- Testify the classification
- Experimental Analysis of Results

In future we can also work to make better tune the parameters in support vector machines to achieve better filtering results is important topic for future works.

References

[1]. Abdelmalek A., Zakaria E., and Michel S., "Evaluation of Text Clustering Methods Using WordNet," The International Arab Journal of Information Technology, vol. 7, no. 4, pp. 349- 357, 2010.

[2]. Alan P. and Hamblen J., "Computer Algorithms for Plagiarism Detection," IEEE Transactions on Education, vol. 32, no. 2, pp. 94-99, 1989.

[3]. Alberto C. and Paolo R., "Towards the Exploitation of Statistical Language Models for Plagiarism Detection with Reference," in Proceedings of ECAI Workshop Uncovering on Plagiarism and Social Software Misuse PAN, Greece, pp. 15-19, 2008.

[4]. Allan K., Kevin A., and Bruce B., "An Automated System for Plagiarism Detection Using the Internet," in Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, Chesapeake, pp. 3619-3625, 2004.

[5]. Francisco R., Antonio G., Santiago R., Jose L., Pedraza M., and Manuel N., "Detection of Plagiarism in Programming Assignments," IEEE Transactions on Education, vol. 51, no. 2, pp. 174-183, 2008.

[6]. Hermann M., Frank K., and Bilal Z., "Plagiarism -A Survey," Universal Computer Science, vol. 12, no. 8, pp. 1050-1084, 2006.

[7]. Wikipedia, available at: http://en.wikipedia. org/wiki/Plagiarism, last visited 2004.

[8]. Webconfs, available at: http://www.webconfs. com/stop-words.php, last visited 2006.

[9]. Jinan F., Alkhanjari Z., Mohammed S., and Alhinai R., "Designing a Portlet for Plagiarism Detections within a Campus Portal," Journal of Science, vol. 1, no. 1, pp. 83-88, 2005.

[10]. Juan A., Nicholas C., and Rafael C., "Applying Plagiarism Detection to Engineering Education," in Proceedings of School of Electrical and InformationEngineering University of Sydney, NSW, pp. 722-731, 2006.

[11]. Nathaniel G., Maria P., and Yiu N., "Nowhere to Hide: Finding Plagiarized Documents Based on Sentence Similarity," in Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, NSW, pp. 690-696, 2008.