Detection and Blocking of Image Spammers

Vivek Khirasaria, Bhadreshsinh Gohil
GTU PG School, Network Security, Gujarat Technological University Ahmedabad, Gujarat, India.

ABSTRACT

Spammers continues to uses new methods and the types of email content becomes more difficult, text-based anti-spam methods are not good enough to prevent spam. Spam image making techniques are designed to bypass well-known image spam detection algorithms like optical character recognition (OCR) algorithm. As the use different methods to create image spam are increased, there must be an algorithm to prevent this type of spams. So using the different properties of images attached in the mail like width and height indicated in header of image, the aspect ratio of width and height, file size, image area, compression, owner, and colours and file format of image. We can develop an algorithm which is used to detect the spam based images using the properties of the attached images. Also it is maintains a database for email addresses of spammers and block the email address based on result of detection process.

Keywords—*Spam, Image Spam, Spammers.*

1. INTRODUCTION

As the use of email system is increased for communication and information sharing, the number of attacks also increased on users. One of these attacks is spam attack. Spam messages volumes are increased in past few years and spammers uses new techniques to target the system. Attackers performs many attacks like capturing private data of user, click frauds, and phishing[5,9].

Spam is inappropriate or undesirable messages sent over the Internet, typically to a group of users, for the purposes of phishing, advertising, sending malwares [12] [8]. Also popular as unsolicited bulk email (UBE), unsolicited commercial email (UCE), or junk mails[8]. Spamming cause wastage of bandwidth so it is significant challenge for system administrator to identify and block this type of emails[11].

So the new aim of spammers in e-mail spam is image spam. It is a method of spam in which the text is presented as a picture in an image file [9]. Aim of spammers behind it, is to bypass the filters whose analysis is based on only the emails textual content [5]. Lot's of techniques have been proposed to detect the image based spam like DNSBL, Spamtraps, Graylisting and many more. All the technique have same goal, trying to avoid image spams from entering inbox [5] [6]. Many anti-spam techniques also use combination of the whitelist, blacklist and greylist [11]. Image spammers are constantly uses new techniques to create image spams, so it is a constant battle to find new features that can effectively defeat new image spam techniques.

2. BACKGROUND

Often our inbox is full of unknown mails. So it is required to have a proper technique, tool or algorithm, using which we can avoid spams [6] [4]. If mail is an unsolicited mail then it would be send to the spam folder else if it is good then it would be send into the inbox [6]. Two basic spam filters are described here which are currently used.

2.1 Text Based Filter

Mostly there are two types of spam filters are available for text based spam detection first is content based spam filtering which generally consists the header's information which includes sender's mail address, receiver's mail address and content of the email and also consist the main body of the email [6]. Other one is rule based filtering uses rules to classify mails as spam or good. These rules may be applied on "Subject" From, or to fields of the header or the content of the mail [8] [4]. Different rules are like check the font size or check whether the sender's address in the receiver's address book or searching the content for words like "free", "lottery" and "sale". [8] [6] [4].

2.2 Image Based Filter

After the rise and spreading of text filtering techniques, the spammer's uses image spam to target people. The spamming text and the URL of advertisement is put into the images, it is useless to analyse the mail content

using text spam filters. It leads to development of image spam based filters [1]. These are based on the header information which contains all information regarding the mail. Content based filters also available which extract the content of the image and analyse the content of the image. Colour based filters also used which are used to get compress, perceptually related presentation of the colour from image. Some filters are also based on the structure of the image which includes the binary pattern, auto correlation, intensities, variation of pixels, geometric moment, And layout of the image, these properties are obtained by analysing the image in the layout or object view of an image [5] [2].

3. RELATED WORK

Research on image spam has not been as extensively as text based spam, however some recent research work have identify the image based spam filtering using the OCR techniques and SVM techniques [4] [10].

- Detection and Blocking of Spammers using Spot Detection Algorithm: Parvati Bhadre and Deepali Gothawal [1] have developed a system which uses a SPOT detection algorithm to detect the spammers and blocks the email ids of attackers. They use two methods PT (Percentage Threshold) and CT (Count Threshold). PT calculates the percentage of spam messages send by same user in one time window and CT counts the number of spam messages sent by a same user in one time window. If the percentage and the count is cross the predefine threshold value then that user is detected as spammer and blocked [1].

- Detecting Image Spam Based on K-Labels Propagation Model:Xiaoyan Qian,Weifeng Zhang, Yingzhou Zhang, Guoqiang Zhou, Ziyuan Wang [2] have proposed a method which used K-Labels Propagation Model to detect image spam. In this model they used two datasets classify-known images dataset and testing images dataset.Firstly, the images in the Classify known Images Dataset are clustered to get the information of cluster centers. Secondly, all images are labelled: if the image belongs to classify image dataset and is ham image, labelled 0; and if spam image labelled 1; if the image belongs to testing database, which is initialized and labelled 1; according to the information of the cluster center, the features of these images are clustered and then a testing feature library is created, the labels of the image in the testing image dataset are created and lastly these images are classified into Ham images and Spam images. They used two algorithms one is improved means clustering algorithm, which is used to extract the features of the images and making a dataset. Second algorithm is used to classify the image as the Ham image or Spam image [2].

- A Modular Approach towards Image Spam Filtering using Multiple Classifiers: Meghali Das, Alexy Bhomick, Y. Jayanta Singh and Vijay Prasad [3] have proposed a method which is based on modular approach. Modular design have two or more modules [3]. Each one is used to analyse any specific properties. This system have two modules. The first module extracts hidden text from the image and this is matched with the known dataset of the spam messages. The second module extracts properties of the image and give input for training a classifier. They design a spam detection system not using a single module. They combine the decisions of the two modules. For first module the OCR is developed witch is uses the words extracted from the image and make a template. This template have numerical, upper-case and lower-case letters and special characters. Second module is low level feature extraction which extracts Colour, texture and shape used as the low-level feature [3].

- Partial Image Spam E-Mail Detection Using OCR: V. Sathiya, M.Divakar and T.S. Sumi [4] proposed a method called Partial Image Spam Inspector (PIMSI) using OCR. It detects a spam from the body of the e-mail. Its first module classifies the content of mail as either text or image. Second module uses two datasets to find the spam image or spam word. First dataset is vocal dataset which contains the spam words and second dataset is image dataset contains dataset of spam images. [4].It uses OCR to separate the content from the mail body, its success rate is high as compared other methods, also it uses the predefine dataset so time taken to compare is less [4].

4. MOTIVATION

Spammers are the one of the key security thread now a day. Spammers send unsolicited mail in bulk to the targeted mail address and full the inbox of the target account. Spammers use this type of mail to collect the private data of the user and do phishing attacks using this type of spamming activities [13]. In detection of the spam mail, text-based filters have developed in sophistication and performance in filtering email spam. In other

hand spammers have developed a number of techniques to bypass these text-based filters. Currently, one of the most used spam development techniques uses embedding text into images and sending in either simple image form or a mix of images and text [13]. To prevent any spamming it is important to find out the spam images from the email body and block the sender from sending email in feature. Some of the image spam detection techniques are available there, which are either based on the OCR and the content of the email header [4]. OCR is Optical Character Recognition technique which is extracting the text information from the image file. And based on email header techniques only extract the header information like sender id, senders IP address and if any URL is there than check for it. And compare all this information to the existing database [6].

5. PROBLEM STATEMENT

Most of the current spam filters are based on the text based spam filtering techniques and based on the OCR techniques. There are many providers who provides the e-mail spam filtering softwares and techniques like Google and Yahoo [7] [8]. These software provides inbuilt filtering techniques still there are a major risk of mix image spam for a number of reasons like e-mail from unknown sources, link leads to an unknown web pages. The messages that are comes from any unknown sender must be prevented. The image spam can be classified in two categories: simple image and mixed image. simple image spam have only images and not detected by the text filters and the mixed image spam have image and a text message so it is partially detected by the text based filters. After the study of the current available filters we can identify the major issues related to image spam filtering.

- Design of an algorithm which detects the image based spam.
- Design of algorithm which blocks the spammers.

Keeping in mind the limitations of existing system and models of text and image based filtering system, we are proposing to develop and implement an algorithm which will contains following characteristics.

- Use image file properties to identify the image as spam image or ham image.
- Use header information of the e-mail to identify the sender and block it using this information.

6. PROPOSED METHOD

6.1 Design

The proposed system aims to detect image spams from the email body by using a statistical probability ratio formula. Fig 1 shows the flow work of the system.

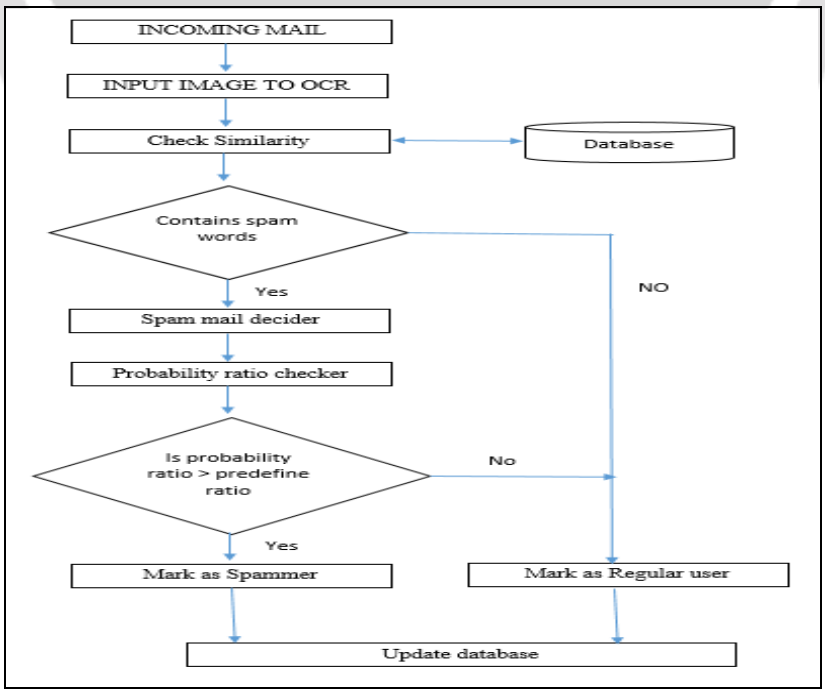

Fig 1 System Work Flow

6.2 Probability Ratio

- For each word, we calculate the ratio *b(w)* and *g(w) as below.*

$$b(w) = \frac{(\text{the number of spam mails containing the word w})}{(\text{the total number of spam mails})}$$

$$g(w) = \frac{(\text{the number of ham mails containing the word w})}{(\text{the total number of ham mails})}$$

- Using *b(w)* and *g(w)*, we can find out probability ratio *p(w)* using statistical probability ratio function.

$$p(w) = \frac{b(w)}{b(w) + g(w)}$$

- For new word probability ratio is defined as

$$p(w) = \frac{(s * x) + (n * p(w))}{s + n}$$

- S is the strength we want to give to word, generally it is given 1.
- X is our assumed probability for the word, which first time appeared in spam, generally it is given 0.
- N is the number of e-mail that contains word w.
- In our system user have to define threshold value (z).

6.3 PROPOSED SPAM DETECTION ALGORITHM

We proposed the bellowed algorithm using the statistical probability ratio formula which are described as above in equations.

1: An email arrives at system
2: Get IP address of email sender *u*
3: Check in database.
4: Let *w* be the message index
5: Let *p(w)* = 1 if message is spam, *p(w)* = 0 otherwise
6: **if** (*p(w)* == 1) **then**
7: *b(w)* = *b(w)* + 1
8: **else**
9: *g(w)* = *g(w)* + 1
10: **end if**
11: **if** (*p(w)* ≥ *Z*) **then**
12: User *u* is spammer. Test terminates for *u* and block u.
13: **else if** (*p(w)* ≤ *Z*) **then**
14: User *u* is normal. Test is reset for *u*.
15: *p(w)* = 0
16: Test continues with new observations
17: **end if**

7. CONCLUSION AND FUTURE WORK

The spam images are increased continuously. They waste the bandwidth and storage of the network. So there is need to employing an efficient method for detection and blocking of spammers. In this work the detection of spam images is based on the content of image file. It also blocks the email ids of the spammers and they have to create new mail ids for further communications. I hope my proposed solution will help to reduce the spamming attacks and provides more security to email.

8. REFERENCES

[1] Parvati Bhadre and Deepali Gothwal,"Detection and Blocking of Spammers using SPOT Detection Algorithm", IEEE conference on Network and Soft computing, pp. 97-101, 2014.
[2] Xiaoyan Qiqn, Weifeng Zhang, Yingzhou Zhang, Guoqiang Zhou and Ziyuan Wang,"Detecting Image Spam Based on KLabels Propogation", IEEE 10th conference on Web Information System and Application, pp. 170-175, 2013.
[3] Meghali Das, Alexy Bhomick, Y. Jayanta Singh and Vijay Prasad,"A modular Approach towards Image Spam Filtering
Using Multiple Classifier", IEEE Conference on computational Intelligence and Computing Research, pp. 1-8, 2014.
[4] V. Sathiya, M. Divakar, and T.S. Sumi,"Partial Image Spam E-Mail Detection Using OCR", International Journal of Engineering Trends and Technology, pp. 55-59, 2011.
[5] Abdolrahman Attar and Reza Moradi Rad Reza Ebrahimi Atani,"A survey of Image Spamming and Filtering Techniques" Springer International Journal of Artificial Intelligence Review., pp. 71-105, 2013.
[6] Sahil Puri, Dishant Gosain, Mehak Abuja, Ishita Kathuria and Nishtha Jatana,"Comparison and Analysis of Spam Detection Algorithm", IJAIEM Volume 2, Issue 4, 2013.
[7] Arushi Gupta, and Rishabh Kaushal,"Improving Spam Detection in Online Social Networks", IEEE Conference on Cognitive Computing and Information Processing, pp. 1-6, 2015.
[8] Meghali Das and Vijay Prasad,"Analysis of an Image Spam in Email Based on Content Analysis", IJNLC Volume 3, Issue 3, 2014.
[9] S. Dhanaraj, and Dr. V. Karthikeyani,"A study on e-mail Image Spam Filtering Techniques", IEEE Conference on Pattern Recognition, Informatics and Mobile Engineering, pp. 49-55, 2013.
[10] Rahul C. Patil and D.R. Patil,"Web Spam Detection Using SVM Classifier", IEEE Conference on Intelligent Systems and Controls, pp. 1-4, 2015.
[11] Peizhou He, Xiangming Wen and Wei Zheng,"A Simple Method for Filtering Image Spam", IEEE/ACIS International Conference on Computer and Information Science, pp. 910- 913, 2009.
[12] Abhinav Pathak, Sabyasachi Roy and Y. Charlie Hu,"A Case for a SpamAware Mail Server Architecture", CEAS Fourth Conference on Email and AntiSpam, 2007.