International Journal of Pure and Applied Mathematics
Volume 117 No. 22 2017, 7-11
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

EXPERIMENTAL ANALYSIS ON SPAM FILTERING

Banumathy Rajesh[1], Hariharan Shanmugasundaram [2]
[1]Research and Development Centre Bharathiar University, Coimbatore- 641046, India.
[2]Professor, Department of Information Technology Vel Tech Multi Tech, Chennai, Tamilnadu, India
banulectit@gmail.com, mailtos.hariharan@gmail.com

Abstract-Internet has been a fastest and easy way of communication. E-mail is the predominant method for users to send information across networks and has attracted people of all ages. Unauthorized users make use of this mode for sending unwanted information which is called as SPAMS. Spammers also introduce new method of embedding the textual spam contents along with the images thereby causing several issues. Spam filtering tools and approaches aims at discriminating set of original image contents as ham and spam images. In this paper we present a study to analyze the content of email spam which includes text contents. The proposed study and experimental illustration focus on discriminating set of images taken from well known corpora as either images which are real images are as fake ones based on the distinguishing characteristic sets. The proposed framework seems to have good accuracy as compared to other baseline approaches.

Keywords -E-mail, Security, Spam, Images, Classification, Prioritization

1. Introduction

Internet technology and services are the core means in today's modern era. This has been used in daily life for effective communication with distinctive style. The concept of E-mail is used extensively for communication nowadays. This makes it possible to communicate with many people simultaneously in a very easy and cheap way. E-mail's are received by users without their desire in an unintentional manner. Spam mail (or, junk mail or bulk mail) is the general name used to denote these types of E-mail. Spam mails are defined as electronic messages posted to thousands of recipients usually for advertisement or profit. These spam mails increases day by day; hence they have to be treated immediately. The usage of email and associated spam contents have been a major concern today. It is found that about percentage of the incoming E-mail's to the network as spam increases by large amount. Fig 1 presents the overview of spam distribution across all domains.

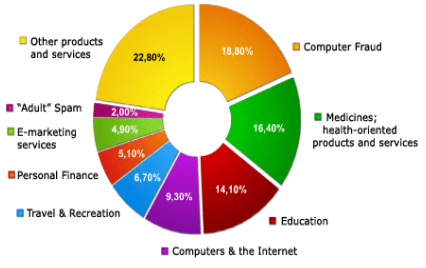


Fig 1: Spam distribution across all fields

The act of inserting text along with image is called as Image spam techn

Several tricks used by spammers to send spam includes IP address, sneaking and obfuscation. Several statistical spam filtering tools are available currently each with its own design characteristics [2, 18]. There are free software toolkit, which may used to conduct similar experiments [1]. Image spam tends to analyze the email's textual content performed by most of the spam filters like SpamAssassin, SpamBayes etc [3]. Fig 2 presents the distinctive ration of text and image spam.

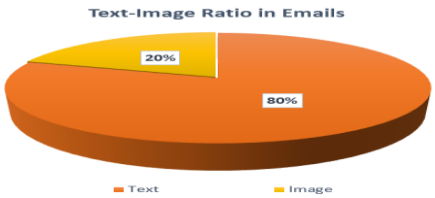


Fig 2: Text and image spam ratio

Image processing is quiet popular approach in converting an image into digital form. The approach tends to extract some useful features and information from it and derive the useful knowledge. Image processing is among rapidly growing technologies today, with lot of applications using it for business needs. Image processing is a core competent research area in engineering disciplines too. Image processing systems are becoming popular due to enhanced tools and techniques, applications [4]. Preventing text recognition using Optical Character Recognition (OCR) tools and imposing additional challenges in filtering reduces the spam in email contents. There exists several spam filters to meet out such demands and challenges [19].

Email spams are of major concern in modern world. Spams are defined as unsolicited bulk email or malware combined with the power of botnets, spammers are now able to launch large scale spam campaigns covering wide range of topics causing major traffic increase and leading to enormous economical loss. Filtering image email spam is considered to be a challenging problem because spammers keep modifying the images being used in their campaigns by employing different obfuscation techniques. Image spam are those as shown in Fig 3 & Fig 4. Fig 3 presents an example of image with clear description of the content. Fig 4 is obfuscated content with text and image interrelated or mixed together causing readability issues.

Fig 3: Clear image

Fig 4: Obfuscated spam images

Image visual features and associated low-level image features characterizes the image. Support vector machine is used widely to classify the images category as genuine or fake ones or not. Recent studies have proved that SVM classifier outperforms well with improved accuracy [5]. In particular, several authors investigated the possibility of recognizing image spam with obfuscated images by using generic low-level image features like number of colours, prevalent colour coverage, image aspect ratio, text area, image metadata [6].

Image spam is the have the format which resembles text spam including Subject, Message Body, and Receiver Address. Elimination of spam is done by matching the text with database that eliminates trained text which is considered as spam words. There exist several tips and tricks to send spam like IP address, obfuscation, URL address, encrypted message and others [8]. This paper presents to study the characteristics mainly based on obfuscation, when image spam in encountered. The text content and spam identification is done using text mining.

The rest of the paper is organized as follows. Section 2 presents some of the work related to the text and image spam. Section 3 presents the proposed framework, data set & characteristics, evaluation and discussion on the experiments carried out. Finally we present the conclusion and future work in section 4.

2.**Related Work**

This section presents some of the work relating to email spam detection of both text and image contents. Recent research studies inclined towards the characteristics of search engines and their performance, user interface design and analysis of adequate features of the WWW attract the users to access the web to deal with text documents as well as multimedia information like images, videos, sounds, graphics etc [12,13]. Earlier works have proposed a system based on Base64 encoding of image files and n-gram technique for feature extraction. Here normal images are transformed into Base64 presentation and then it used n-gram technique to extract the feature. Using SVM, spam images were detected from legitimate images leading to an approach shows resulting to improved performance in terms of time efficiency with the standardized training datasets based on two image features [9].

Work on image spam is considered to be an unsolved problem, mainly because of two reasons. One is due to the diversity of spamming trick and the other due to the evolving nature of image spam. Also it is analyzed and reported that new spam has been constantly emerging. To alleviate this problem, spam filters were designed. Study in the recent year's report that the effectiveness drop over a period of time interval [10]. The authors have presented an effective anti-spam approach to solve the two problems reported. First, a novel clustering filter is proposed. By exploring the density-based clustering algorithm, the proposed filter is found to be robust in nature to spamming. Then, subsequently a hierarchical framework is presented by combining the clustering filter with other machine learning based classifiers. Moreover, incremental learning mechanism is integrated to ensure the proposed framework be capable of adjusting itself to overcome new image spamming tricks. The proposed framework is evaluated on two public spam corpora. The experiment results have shown that the proposed framework achieves high precision along with low false

significance of search engines. The search engines have [illegible] user predominantly by web surfers. Information retrieval task is viewed as a problem of classifying items into one of two classes corresponding to interesting and uninteresting ones. The performance evaluation is defined in terms of classification accuracy, precision and recall metric. Measuring the information retrieval effectiveness of World Wide Web search engines is costly, since it involves human relevance judgments. Web search engines, since such search engines help their users find higher number of relevant Web pages with less effort. A comprehensive study evaluates the performance of three Web search engines. This is more important for identifying the spam information which would affect the performance [11].

Spam filter developers are confronted with the task of integrating their ideas in user-friendly products. Here, the authors have introduced Spamato, an open, extendable, and multi-faceted spam filter framework. Spamato provides fundamental services commonly required by filter developers to facilitate the implementation of new approaches. Furthermore, we support email clients with add-ons to enable users to intuitively collaborate with Spamato. Also, a variety of filters and exhibit an evaluation of URL-based techniques is also presented [14].

Today, the internet is the most powerful tools throughout the world. But the explosive growth of unsolicited emails has prompted the development of numerous spam filtering techniques. It needlessly obstructs the entire system. Spammers are creating new ways against anti-spam technology. The newest of which is image-based spam. In general words, image spam is a type of email in which the text message is presented as a picture in an image file. This prevents text based spam filters from detecting and blocking such spam messages. There are several techniques available for detecting image spam (DNSBL, GrayListing, Spamtraps, etc). Each one has its own advantages and disadvantages. On behalf of their weakness, they become controversial to one another. This paper includes a general study on image spam detection using histogram and hough transform, which are explained in the following sections. The proposed methods are tested on a spam archive dataset and are found to be effective in identifying all types of spam images having (1) only images (2) both text and images. The goal is to automatically classify an image directly as being spam or ham. The proposed method was able to identify a large amount of malicious images while being computationally inexpensive. [15]

Colour and gradient orientation histograms and utilizes this data on probabilistic boosting tree (PBT) to distinguish spam images from ham images. Each node of PBT contains colour or gradient orientation histogram data of corresponding part of images inside training dataset. In the proposed detection method postulated that spammers use the same template to send a lot of spam images and they add random noises to an image template in order to bypass filters [16].

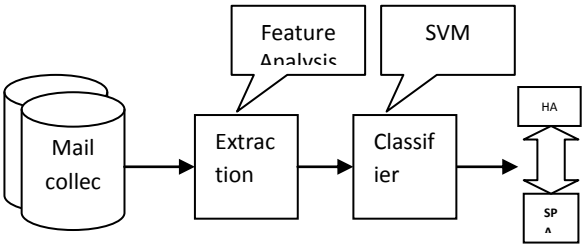
Image based spam is a recent trick developed by the spammers' community with the intention of bypassing the successful text based spam filters. Most of the traditional text based filters have been based on Naïve Bayes classification combined with text categorization methods. This work concentrates in developing a spam filtering system that accurately blocks image spam. The system analyzes images sent as attachments extracting both textual and visual features. The rationale behind employing a combination of both kinds of features is that spammers usually embed the payload in an image hidden by various obscuring methods. We used SVM classifier for the classification of low level features. The use of

[illegible] images also [illegible] rovides a final

measure of the spamminess of the images with its decision based on the classifiers. [7] To circumvent prevalent text-based anti-spam filters, spammers have begun embedding the advertisement text in images. Analogously, proprietary information (such as source code) may be communicated as screenshots to defeat text-based monitoring of outbound e-mail. The proposed method separates spam images from other common categories of e-mail images based on extracted overlay text and color features. No expensive OCR processing is necessary. Our method works robustly in spite of complex backgrounds, compression artifacts, and a wide variety of formats and fonts of overlaid spam text. It is also demonstrated successfully to detect screen-shots in outbound e-mail. [6] In recent years anti-spam filters have become necessary tools for Internet service providers to face up to the continuously growing spam phenomenon. Current server-side anti-spam filters are made up of several modules aimed at detecting different features of spam e-mails. In particular, text categorization techniques have been investigated by researchers for the design of modules for the analysis of the semantic content of e-mails, due to their potentially higher generalization capability with respect to manually derived classification rules used in current server-side filters. However, very recently spammers introduced a new trick consisting of embedding the spam message into attached images, which can make all current techniques based on the analysis of digital text in the subject and body fields of e-mails ineffective. A generic approach to anti-spam filtering was presented which exploits the text information embedded into images sent as attachments. Our approach is based on the application of state-of-the-art text categorization techniques to the analysis of text extracted by OCR tools from images attached to e-mails. The effectiveness of the proposed approach is experimentally evaluated on two large corpora of spam e-mails [17].

3. Proposed Work

Email spam filtering represents a major approach to combat spam. The goal of email spam filtering is to classify email messages into ham or spam. Content-based techniques inspect the body of an email searching for specific keyword(s) or features that are typically used by spammers or associated by certain spam campaign. Email body itself may be text, image, or both. Therefore, content-based filtering techniques usually deal with all these content types. This section presents an experimental study on some of the baseline methods as compared with the proposed approach. In our proposed scheme we investigate on the valid set of rules and the order of preferences. Figure 5 presents the framework for proposed spam filtering. Rule based approaches exists earlier [15], but it failed to capture the priority or importance of the rule order. In our proposed work, we have taken into account the importance or the priority of each such rule. The experiments were also compared with other baseline approaches using multilayer perceptron (MLP) neural network, Bayesian network and rule based approach. The experiments were investigated sing bench marked data set containing image spam data sources [18,19] collected from multiple email accounts organized into several groups.



In the proposed framework, the email contents collected are classification, the emails are identified as SPAM OR HAM. Extraction involves text analysis, which focus on review of text components. Table 1 presents the sample, outlining the sample set of SPAM words. The words clearly denotes the spam content, which could be removed during feature analysis. The subsections below present the algorithm used for investigation.

credit, loan, bills, info, money, investment, discount, win, order now, sign up, clearance, free gift, free samples dating, find, guess, statement, private, dear, partner, singles, fast cash, incredible deal, free info, satisfaction, buy direct call free, call now, camcorder, phone, cards, extra inches, cialis, viagra, spa, beauty, money back, click here, act now prize, guaranteed, claim, cash, no fees, limited time, life insurance, mortgage, amazing, 100% satisfied, 100% free

Table 1: Sample set of SPAM words

3.1. Multilayer perceptron (MLP) neural network:

The model delivers information by activating input neurons containing values labelled on them. Activation of neurons is calculated in the middle or output layer, as shown in equation 1.

$$a_i = \sigma\left(\sum_j W_{ij} O_j\right) \quad (1)$$

where α_i represent activation level of neuron i; j is neuron set of the previous layer;Wij is the weight of the link between neuron i and j; Oj is the output of neuron j and s(x) represents a transfer function.

3.2 Bayesian network:

A Bayesian network is an acyclic directed graph indicating probability distribution in a compressed way. A node in this graph shows a random variable, Xi. A directed edge between two nodes indicates potential interdependence between a variable shown by the parent node and another variable shown by a child node. Bayesian network used for classification including group C which indicates variable of class label and variable Xi which indicates features.

3.3 Rule based approach:

The work here was based on rules for proper scoring in terms of the efficiency of rules. The considered rules were provided in three forms: 1) email header information analysis, 2) keyword matching, and 3) main body of the message. A score was finally obtained for these rules. Table 2 presents a sample set of rule and its description.

Rule id	Description
R1	Email recipient
R2	Subject in email
R3	Content of email
R4	Unclear subject
R5	Unambiguous info

Table 2: Sample rules for spam identification

The proposed approach adopts the information in the above table, however, exhaustive set of rules which doesn't overlap each other is added. These rules are obtained by expert users who are knowledgeable and have deep study on the spamicity rules. We present our improved approach by modifying the weights by incrementing the weights as combined with each other rules. The higher the weight, the

ure 6 presents

ıe approaches

discussed in this section as compared with the proposed scheme. The proposed approach is more than other baseline schemes and has stronger influence which leads to reduction of spam.

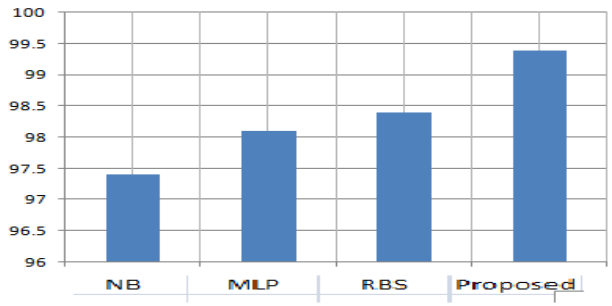


Fig 6: Accuracy of spam detection as compared with different schemes

4. Conclusion and Future Work

In this paper, we present a study on spam detection system. By examining the content of incoming mail contents, we perform an analysis for comparison of images that are near duplicates of known spam images (visually). The contents are analyses using text mining approach, our system can effectively analyze the spam content maintaining a low false positive rate. The proposed framework investigates on an efficient algorithms to identify spam mails. It is found that our proposed approach is significant in identifying the spam in an accurate manner as compared to other baseline schemes chosen for the study. Results presented is highly reliable and has potential importance in real world scenario. Our work can be extended to image spam, involving image processing. We also work towards presenting a integrated tool to identify the mails containing both text and images together.

References

1. GORDON V. CORMACK and THOMAS R. LYNAM, "On-line Supervised Spam Filter Evaluation", ACM Transactions on Information Systems (TOIS), Volume 25 Issue 3, Article No. 11, July 2007.
2. Le Zhang, Jingbo Zhu and Tianshun Yao, "An Evaluation of Statistical Spam Filtering Techniques", ACM Transactions on Asian Language Information Processing, Vol. 3, No. 4, December 2004, Pages 243-269.
3. Zin Mar Win and Nyein Aye, "Identification of Image Spam by Using Histogram and Hough Transform", International Journal of Science and Research, Volume 2 ,Issue 11,pp.310-314, November 2013.
4. Nida M. Zaitoun and Musbah J. Aqel, "Survey on Image Segmentation Techniques", International Conference on Communication, Management and Information Technology (ICCMIT 2015) , Procedia Computer Science 65 (2015) 797 – 806.
5. C. Burges. A Tutorial on Support Vector Machines for [illegible] overy, 2(2):121–167, 1998.
6. H.B. Aradhye, G.K. Myers and J.A. Herson, "Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on", Seoul, South Korea, South Korea, 2005.
7. Meghali Das, Alexy Bhomick and Y. Jayanta Singh, "A modular approach towards image spam filtering using multiple classifiers", Computational Intelligence and Computing Research (ICCIC), 2014 IEEE International Conference on, Coimbatore, India , 2014.
8. Thamarai Subramaniam, Hamid A. Jalab and Alaa Y. Taqa, "Overview of textual anti-spam filtering techniques", International Journal of the Physical Sciences Vol. 5(12), pp. 1869-1882, 4 October, 2010.
9. Christina V, Karpagavalli S and Suganya G, "A Study on Email Spam Filtering Techniques", *International Journal of Computer Applications (0975 – 8887, Volume 12– No.1, December 2010.*
10. Li Xiao Mang, HaRim Jung, Hee Yong Youn and Ung-Mo Kim, "AN INCREMENTAL LEARNING BASED FRAMEWORK FOR IMAGE SPAM FILTERING:, International Journal of Computer Science, Engineering and Applications (IJCSEA) Vol.4, No.1, February 2014.
11. AJAY and Olusola Olajide, "Performance Evaluation of Selected Search Engines", Computer Engineering and Intelligent Systems, Vol. 5, No.1, 2014.
12. G. Yan, Y. Ming, Z. Xiaonan, B. Pardo, W. Ying, T. N. Pappas, and A. Choudhary, "Image Spam Hunter," in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference, Las Vegas, Nevada, U.S.A. 2008, pp. 1765-1768.
13. G. Fumera, I. Pillai, and F. Roli. Spam Filtering Based On The Analysis Of Text Information Embedded Into Images. The Journal of Machine Learning Research, 7:2699–2720, 2006.
14. K. Albrecht, N. Burri, and R. Wattenhofer. Spamato-An Extendable Spam Filter System. 2nd Conference on Email and Anti-Spam (CEAS), Stanford University, Palo Alto, California, USA, 2005.
15. Zin Mar Win and Nyein Aye, "Identification of Image Spam by Using Histogram and Hough Transform", International Journal of Science and Research, Volume 2 ,Issue 11,pp.310-314, November 2013.
16. G. Yan, Y. Ming, Z. Xiaonan, B. Pardo, W. Ying, T. N. Pappas, and A. Choudhary, "Image Spam Hunter," in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference, Las Vegas, Nevada, U.S.A. 2008, pp. 1765-1768.
17. G. Fumera, I. Pillai, and F. Roli. Spam Filtering Based On The Analysis Of Text Information Embedded Into Images. The Journal of Machine Learning Research, 7:2699–2720, 2006.
18. http://www.cs.princeton.edu/cass/spam/
19.http://www.cs.northwestern.edu/~yga71/ML/ISH.htm# dataset