Volume 3, Issue 7, July 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Spam Detection System Using Hidden Markov Model

Vandana Jaswal, Astt Professor. Nidhi Sood

(Deptt of Computer Science & Engineering)

Bahra University, Shimla Hills, India

Abstract— ***Spams are the textual context of the system which can damage our system. E-mail is an essential communication tool that has been greatly abused by spam sender to disseminate unwanted information and spread malicious contents to Internet users. Spam filters provide better protective mechanisms that are able to design a system to recognize the spams. We propose an image spam detection system that uses detect spam words. We rely on filtering methods to detect stemming words of spam images and then use Hidden Markov Model of spam filters to detect all the spam images. In the first section of the paper analysis the Introduction of Spam. In the second section of the paper described the related work. In the third section analyzed the problem formulation. In the fourth section described spam detection techniques, different steps for spam detection. In the fifth section described methodology of spam detection and then the different spam feature extraction. Finally present the Conclusion & future works with the references.***

Keywords- ***Spam, Spam filter, spam detection, text reorganization.***

1. Introduction

Spam is a very annoying problem for email users. So far there is no perfect tool to determine if an incoming email is spam or good. Spam is anonymous, unsolicited bulk email from the recipients' point of view, it is unwanted Detritus that chokes up their inboxes. Spam has become a part of our everyday lives. It is indicative of what is happening to the global economy. When looked at as a whole it is clear that the themes and development trends of spam closely correlate to the global financial situation. More than 70% of global email traffic consists of spam. Dealing with spam incurs high costs for organizations, prompting efforts to try to reduce spam-related costs by installing spam filters. This is called as spam filter mechanism. The individual efficiency of a spam filter installation depends on the amount of spam that is received and on the level of knowledge about spam.[1,2,3] Spam filters are mainly categorized as list based and content based spam filters. List-based filters attempt to stop spam by categorizing senders as spammers or trusted users, and blocking or allowing their messages accordingly. The various types of filters in this category are Blacklist filters, Real time Black hole list and White list filters. Content Based Filters Rather than enforcing across-the-board policies for all messages from a particular email or IP address, content- based filters evaluate words or phrases found in each individual message to determine whether an email is spam or legitimate.

1.1 Spam detection and spam reorganization

Spam Detection has importance regarding finding the patterns, forecasting, discovery of knowledge etc., in different business domains. Spam Detection techniques and algorithms such as classification, clustering etc., helps in finding the patterns to decide upon the future trends in businesses to grow. Spam Detection has wide application domain almost in every industry where the Spam is generated that's why Spam Detection is considered one of the most important frontiers in Spam base and information systems and one of the most promising interdisciplinary developments in Information Technology. Spam text recognition aims to automatically identify the textual state of a human being from his or her voice. It is based on in-depth analysis of the generation mechanism of Spam signal, extracting some features which contain textual information from the speaker's voice, and taking appropriate pattern recognition methods to identify textual states. Like typical pattern recognition systems, our Spam text recognition system contains four main modules: Spam input, feature extraction, SVM based classification, and text output

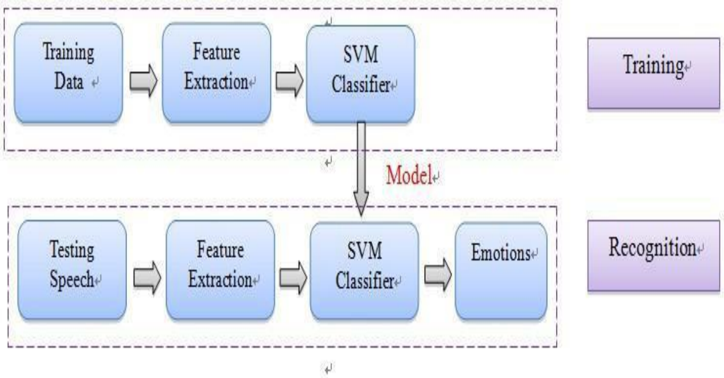

Figure 1 . Spam Text Recognition System

2. Related Work

Natarajan 2010 [4] provides a third party large-scale blacklist to decide which email is spam. A blacklist is a list of traits that spam emails have, and if the email to be tested contains any of those traits, it is marked as spam. It is possible to organize blacklist based on "From:" fields, originating IP addresses, the subject or body of the message, or any other part of the message that makes sense. A small-scale blacklist works fine if the user gets spam from one particular address. O' Brien J and Chiarella J (2003) [5] state that it is obvious problem that it is impossible to predict who is going to send email, and anyone previously unknown to the user will be filtered out. One way is to avoid this problem is to read through the filtered email regularly but there is no point in filtering if the user must view all of the email anyways. Jacobson M. Modeling [6] is described Phishing is a form of social engineering in which an attacker attempts to fraudulently acquire sensitive information from a victim by impersonating a trustworthy third party. It should detect online crime by that the phisher could then notify the victim of a "security threat." The context aware phishing an attacker would gain the trust of victims by obtaining information about their bidding history or shopping preferences their banking institutions or their mothers' maiden names. Yue Zhang, Serge Egelman, Lorrie Cranor and Jason Hong [7] is described that the anti-phishing tools that were examined in this study left a lot to be desired. Spoof Guard did a very good job at identifying fraudulent sites, but it also incorrectly identified a large fraction of legitimate sites as fraudulent. The performance of the other tools varied considerably depending on the source of the phishing URLs. Of these other tools, only IE7 was able to correctly identify over 60% of phishing URLs from both sources, but it still missed 25% of the APWG phishing URLs and 32% of the phishtank.com phishing URLs. The only tool we tested that is known to make no use of blacklists was Spoof Guard. While it was able to identify the majority of phishing sites using only heuristics, it still missed some phishing sites and it had a very high false positive rate. Adil Alsaid and Chris J. Mitchell [8] is described for secure web sites which uses the SSL/TLS protocol for server authentication. Mutual authentication support is provided by SSL/TLS which uses both server and client authentication. This feature of SSL/TLS is not used by most web sites because not every client has a certified public key. Instead user authentication is typically achieved by sending a password to the server after the establishment of an SSL-protected channel. Certain attacks rely on this fact, such as web spoofing and phishing attacks. This paper described the issue of online user authentication is discussed and a method for online user authentication using trusted computing platforms is proposed.

3. Problem Formulation

Spams are the textual context of the system which can damage our system . Our basic problem is to protect our system from such unwanted files. To save our system form such kind of failures we need to design a system which can recognize the spams and can let you know on the basis of a training system. Today, Spammers are exploring the advantages of electronic mail (email) .This is because of its efficiency, effectiveness and it is considered very cheap as they can send the same messages to many email users from addresses gotten by various means.

4. Spam Detection Techniques

I. Web Spam Detection as a Classification Problem

Web spam detection can be viewed as a binary classification problem, where a classifier is used to predict whether a given web page or entire web site is spam or not. The machine learning community has produced a large number of classification algorithms, several of which have been used in published research on web spam detection, including decision-tree based classifiers, SVM-based classifiers, Bayesian classifiers, and logistic regression classifiers. Many spam detection techniques have been proposed in recent years. Some methods were developed through competitions such as Web Spam Challenge and Discovery Challenge. Web spam is divided into two types: content spam and link spam [9].

4.1 Content Spam

The content spam is most widespread form of web spam between client & server because of the fact that search engines use information retrieval models based on a page content to rank web pages, such as a vector space model [9] and statistical language models [10].

5. Content-Based Filters

A. Word-Based Filters: A word-based spam filter is the simplest type of content-based filter. Generally speaking, word-based filters simply block any email that contains certain terms. Since many spam messages contain terms no often found in personal or business communications, word filters can be a simple yet capable technique for fighting junk email. However, if configured to block messages containing more common words, these types of filters may generate false positives. Also note that since spammers often purposefully misspell keywords in order to evade word- based filters, your IT staff will need to make time to routinely update the filter's list of blocked words. [11]

B. Heuristic Filters: Heuristic filters take things a step beyond simple word-based filters. Rather than blocking messages that contain a suspicious word, heuristic filters take multiple terms found in an email into consideration.[11,12]Heuristic filters scan the contents of incoming emails and assigning points to words or phrases. Suspicious words that are commonly found in spam messages, such as "Rolex" or "Viagra," receive higher points, while terms frequently found in normal emails receive lower scores. However, heuristic filters configured to be aggressive may generate false positives if a legitimate contact happens to send an email containing a certain combination of words. Similarly, some savvy spammers might learn which words to avoid including, thereby fooling the heuristic filter into believing they are benign senders. [12]

C. Bayesian Filters: Bayesian filters, considered the most advanced form of content-based filtering, employ the laws of mathematical probability to determine which messages are legitimate and which are spam. In order for a Bayesian filter to effectively block spam, the end user must initially "train" it by manually flagging each message as either junk or legitimate. Over time, the filter takes words and phrases found in legitimate emails and adds them to a list; it does the same with terms found in spam. To determine which incoming messages are classified as spam, the Bayesian filter scans the contents of the email and then compares the text against its two-word lists to calculate the probability that the message is spam. For instance, if the word "valium" has appeared 62 times in spam messages list but only three times in legitimate emails, there is a 95 percent chance that an incoming email containing the word "valium" is junk. Because a Bayesian filter is constantly building its word list based on the messages that an individual user receives, it theoretically becomes more effective the longer it's used. However, since this method does require a training period before it starts working well, you will need to exercise patience and will probably have to manually delete a few junk messages, at least at first.[12,13]

4.3 List-based web spam detection is a basic approach to detecting automatically generated web spam pages. Techniques like Trust Rank [10] minimize the impact of spam pages on ranking. This method can detect web spam pages without analyzing page contents.

A. Blacklist: This 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 [11].

B. Real-Time Black hole List: This spam-filtering method works almost identically to a traditional blacklist but requires less hands-on maintenance. This filter simply has to connect to the third-party system each time an email comes in, to compare the sender’s IP address against the list. Since the list is likely to be maintained by a third party, you have less control over what addresses are on or not on the list [12].

C. White list: A white list blocks spam using a system almost exactly opposite to that of a blacklist. Rather than letting you specify which senders to block mail from, a white list lets you specify which senders to allow mail from; these addresses are placed on a trusted-users list. Some anti-spam applications use a variation of this system known as an automatic white list. 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 white list [11].

D. Relist: A relatively new spam-filtering technique, relists take advantage of the fact that many spammers only attempt to send a batch of junk mail once. Under the relist system, the receiving mail server initially rejects messages from unknown users and sends a failure message to the originating server. If the mail server attempts to send the message a second time step most legitimate servers will take the greylist assumes the message is not spam and lets it proceed to the recipient's inbox. At this point, the greylist filter will add the recipient's email or IP address to a list of allowed senders [11].

6. Methodology Of Spam Dectection

STEP-1 First, to design a spam detection system.

STEP-2 To select a spam file either it is text file or it is excel file.

STEP-3 To select the file on the basis of spam detection.

STEP-4 Filter stemming words only from spam detection.

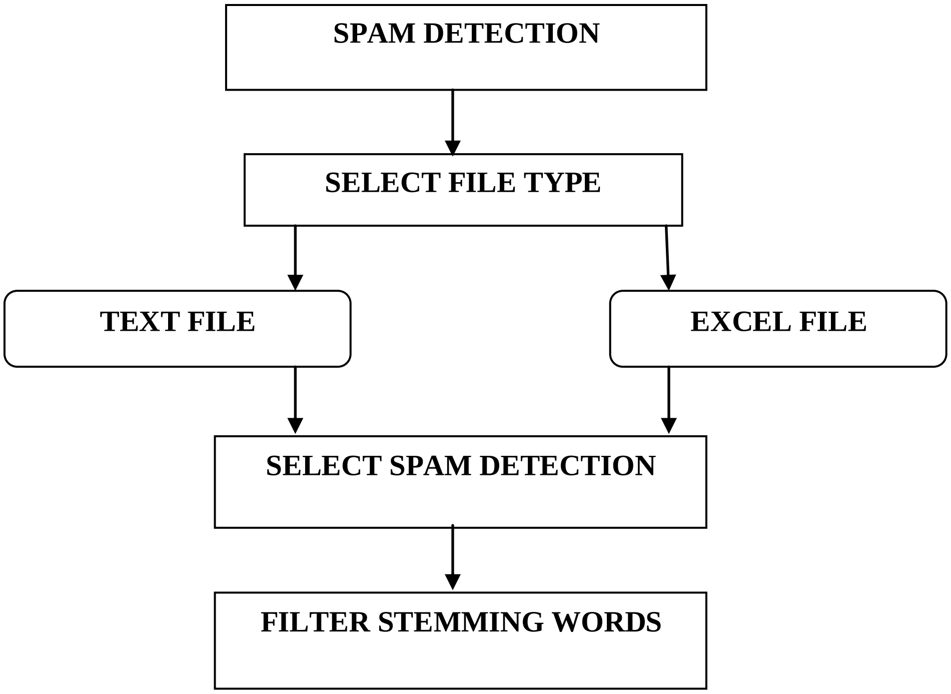

Figure 1: Spam detection system architecture.

7. Spam Feature Extraction

In recent researches, many common features are extracted, such as Spam rate, energy, pitch, formant, and some spectrum features, for example Linear Prediction Coefficients (LPC), Linear Prediction Cepstrum Coefficients (LPCC),Mel-Frequency Cepstrum Coefficients (MFCC) and its first derivative.

(1) Energy and Related Features

The Energy is the basic and most important feature in Spam signal. In order to obtain the statistics of energy feature, we use short-term function to extract the value of energy in each Spam frame. Then we can obtain the statistics of energy in the whole Spam sample by calculating the energy, such as mean value, max value, variance, variation range, contour of energy.

(2) Content and Related Features

The pitch signal is another important feature in Spam text recognition. The vibration rate of vocal is called the fundamental frequency F0 or pitch frequency. The pitch signal is also called the glottal wave-form; it has information about text, because it depends on the tension of the vocal folds and the sub glottal air pressure, so the mean value of pitch, variance, variation range and the contour is different in seven basic textal statuses.

(3) Linear Prediction Cepstrum Coefficients (LPCC) LPCC embodies the characteristics of particular channel of Spam, and the same person with different textal Spam will have different channel characteristics, so we can extract these feature coefficients to identify the texts contained in Spam. The computational method of LPCC is usually a recurrence of computing the linear prediction coefficients (LPC), which is according to the all-pole model.

(4) Mel-Frequency Cepstrum Coefficients (MFCC)

Mel frequency scale is the most widely used feature of the Spam, with a simple calculation, good ability of the distinction, anti-noise and other advantages. MFCC in the low frequency region has a good frequency resolution, and the robustness to noise is also very good, but the high frequency coefficient of accuracy is not satisfactory.

(5) Mel Energy Spectrum Dynamic coefficients (MEDC)

MEDC extraction process is similar with MFCC. The only one difference in extraction process is that the MEDC is taking logarithmic mean of energies after Mel Filter bank and Frequency wrapping, while the MFCC is taking logarithmic after Mel Filter bank and frequency wrapping.

The following properties are highly desirable for the

feature extraction unit:

• Efficient : The unit should be able to process incoming images very efficiently in order to match the throughput of targeted mail servers.

• Effective : Spammers typically add random "noises" to each spam image. For effective detection, the unit should produce features that are relatively insensitive to those added noises.

• Distinctive : To minimize false positive rate, the unit should generate features that can distinguish spam images from non-spam images.

8. Conclusion &Future Work

The EMAIL filter software is also designed to remove every form of flooding and illegal spoofing. Over time we have seen detector automatically spam messages from service provider because such messages wouldn't have been sent on a local INTERNET service provider's web application. The new intelligent system is designed to meet the local INTERNET providers' needs such as an automated view of activity logs of every action carried out by a user, deactivation and activation of clients, auto- train software with new words. Spam is one of the most annoying and malicious additions to global computer world. In this paper present different spam filters are available which are effectively work on their suitable scenarios. Some of list base filters and some of content based filters. Content based filters are more effective than list based filters. Based on this research, Bayesian filter is the most effective content based filter. The effectiveness of a Bayesian spam filter can be increased with pre-processing steps that are applied to the spam keywords training. These are used to increase the accuracy of the spam detection on the basis of pattern matching and stemming. In this paper, present an image spam detection system and spam feature extraction such as rate, energy, pitch, formant and spectrum features. In future to design a better system for spam detection of spam and planning to work on new feature extraction units for image spam filters that can improve the performance of the categories in which our current system does not perform well. Furthermore, since image spam is constantly evolving and finds new features that can effectively defeat new image spam techniques to increase the accuracy of the spam detection.

References

[1]. Wu, C. T., Cheng, K. T., Zhu, Q., Wu, Y. L., "Using Visual Features For Anti-Spam Filtering, " 2005 IEEE International Conference on Image Processing (ICIP2005), pp. 509–512, 2005.

[2]. postini: Email Monitoring + Email Filtering Blog. http://www.dicontas.co.uk/blog/quick-facts/emailspam-trafficrockets/ 65/.

[3]. Toshihiro Tabata, "SPAM mail filtering : commentary of Bayesian filter, " The journal of Information Science and Technology Association, Vol.56, No.10, pp.464-468, 2006.

[4]. Natarajan Arulanand (2010): Payload Inspection Using Parallel Bloom Filter in Dual Core Processor; Computer and Information Science: Vol. 3, No. 4; 2010. [5]. O' Brien J and Chiarella J (2003): AN ANALYSIS OF SPAM FILTERS; Available at; http://web.cs.wpi.edu/~claypool/mqp/spam/mqp.pdf on 9/10/2011.1

[6] Jakobsson M. Modeling and Preventing Phishing Attacks, Phishing Panel of Financial Cryptography, 2005.

[7] "Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong" "Phinding Phish: Evaluating Anti-Phishing Tools"" Vol 4 IJCSS 2011".

[8] "Adil Alsaid and Chris J. Mitchell" "Preventing Phishing Attacks Using Trusted Computing Technology" Information Security Group, Royal Holloway, University of London Egham, Surrey TW20 0EX, UK

[9] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. Commun. ACM, Vol.18, Nov. 1975.

[10] C. Zhai. Statistical Language Models for Information Retrieval. Now Publishers Inc., Hanover, MA, 2008.

[11]. http://www.cs.nmt.edu/~janbob/SPAM,Spam corpus, SMS corpus,

[12]. http://www.comp.nus.edu.sg/~rpnlpir/downloads/cor pora/smsCorpus/

[13]. Amayri O, Bouguil N (2009). Online Spam Filtering Using Support Vector Machines. IEEE., pp. 337- 340.

[14] Hyeon-Kyu Lee and Jin H. Kim, 'An HMM-Based Threshold Model Approach for Gesture Recognition', IEEE Transactions on Pattern Analysis and Machine Intelligence October 1999 (Vol. 21, No. 10). pp. 961-973