Divya Arora al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 2, Issue 1, March 2015, pp. 135-137

A Study of Various Spam Filtering Techniques

Divya Arora[1], Gunjan Rehani[2]

[1]MTech. Student, SPITM, Sonepat;

[2]A.P.(CSE Depth.), SPITM, Sonepat;

Abstract: E-mail has become popular means for peculiar and professional announcement due to its fast and free availability but several people and companies misuse this facility to distribute unsolicited bulk messages that are commonly called as spam mails. In this paper we will present basics on which spam can be classified, various steps and method to determine whether the mail is spam or not. And we will also study the features of some of the main techniques with proper explain.

Keyword: Spam filtering, classifier, machine learning, fault tolerance, Independent, neural.

I. INTRODUCTION

E-mail has become popular means for personal and business communication due to its fast and free availability as well as low or free cost. But several people and companies misuse this facility to distribute unsolicited bulk messages that are commonly called as spam mails. Spam emails may include advertisements of drugs, software, Nigerian scam, adult content, health insurance or other fraudulent advertisements. Spammers can collect email addresses from chat rooms, some AOL profiles, public networking websites, customer lists, from white and yellow pages, newsgroups, worms etc. Sometimes little bit information about target system is enough to get the email address of him. Spam detection problem is becoming more serious now days. It consumes more than half bandwidth of mailboxes. Spam frustrates, confuse and annoy email users by wasting valuable resources and time. Spam even provides ways for phishing attacks and distributing harmful content such as viruses, Trojan horses, worms and other malicious code. Without a spam filter, one email user might receive over hundreds of mails daily and find that most of them are of spam category. The spam mails are with no use of email users. Due to this, serious attention has given to this issue in mailboxes. Several technical solutions like commercial and open-source products have been used to alleviate the effect of this issue. Spam filtering can be of two types: Non-machine learning based and Machine learning based. Early anti-spam techniques like white list, blacklist and set of keywords like "you have won" fall into non-machine learning based techniques. White list contains list of safe senders whereas blacklist contains list of blocked systems or users. As these methods are dependent on lists so these can be easily resolved by spammers. These methods also require manual update and sometime these methods misclassify legitimate mail as spam mail which is more dangerous than no filtering. The British Computer Society (BCS), concluded that misclassification of mails may waste over five million working hours a year by users.On the other hand, machine learning techniques first analyze the message content and then perform classification of mail as spam or ham. Various machine learning anti-spam methods are:

- Support Vector Machines
- Memory based learning
- Ripper rule based learning
- Boosting Decision based learning
- Bayesian Classifiers
- Fuzzy similarity [1]

Although previous study has reported promising detection accuracies but still false positive rate is high. Various techniques have been developed to combat the problem of spam but still effective and efficient technique is required which will have very low false positive and false negative. As single technique is not sufficient to combat this issue so multiple methods should be used. Neural network is information processing system that works on biological nervous system. In this system, large numbers of processing elements are connected together work to resolve a specific problem. Neural networks build a model that states complex relationship between inputs and outputs. Features of neural network are:

- They are extremely powerful computational devices.
- Massive parallelism makes them very efficient.
- They can learn and generalize from training data – so that there is no need for enormous feats of programming.
- Neural networks are fault tolerant it means graceful degradation in biological systems.
- They are very noise tolerant so they can cope with situations where normal symbolic systems would have difficulty.
- In principle, neural network can do anything a symbolic/logic system can do and more.
- Each neuron in neural network does some amount of information processing.
- It derives input from some other neuron and in return gives its output to other neuron for further processing. [6]

II PREVIOUS WORK

In this it has been provided a short overview of existing work and entry points to the literature. Since spam has becoming major threat to usability of electronic mail, research community has developed some techniques and solutions over the past few years. Some technical anti-spam approaches are:

- **Access Filtering:** It verifies and authenticates header information of an email.
- **Economic Filtering:** Two main categories of economic solutions are computing-time-based systems and money-based systems. Computing-time-based systems stimulate spammer to spend considerable computing resources to send a single spam message. Money-based systems charge a small amount of money from every email sent.
- **Content-based Filtering:** The systems that implement content-based filtering perform filtering when message is fully received. These systems can use rule-based filtering, Naive Bayesian classification, memory based approach and checksum methods.
- **Characteristics based Filtering:** It finds out distinct characteristics between good emails and spam.

III. ANTI-SPAM TECHNIQUE STEPS

As content features of spam keep on changing with time so system must be adaptable. This technique uses feature detection algorithm for extracting same features as spam. When feature selection is decided, various machine learning algorithms will be used. Following steps will be performed:

- **FEATURE SELECTION** Several methods have been used for selection of text documents: Word stemming, stop terms, mutual information feature selection.
- **WORD STEMMING** This algorithm improves accuracy of text classification. This removes suffixes from words so lessen the complexity of feature selection.
- **STOP TERMS** A stop lists contains those words which have not been used for feature selection such as "a", "as", "the" etc. As they are common in all mails so they can be excluded.
- **MUTUAL INFORMATION BASED FEATURE SELECTION** This will extract out most valuable features from documents. All features require computational time and space. So information gain of feature is measured. The feature which will have highest MI value will be selected for use in feature vector. It will be used for classification.

IV. TYPES OF SPAM FILTERING TECH

These are main techniques of spam filtering

Deep Belief Networks for Spam Filtering DBNs are feed forward networks with many hidden layers.It uses a greedy layer-wise unsupervised algorithm to initialize the weight of deep neural network based on use of RBM. Results of DBN are much better than SVM but technique for selecting number of hidden layers is still required.

Other techniques are:

A rule-based approach has developed for spam detection. It uses training and testing phases of data [2]. This approach improves efficiency of spam filtering as compared to previously proposed techniques but time complexity is higher due to rules generation and their execution. So digests were used in this approach to detect spam mails. A social network has to construct based on email exchanges between various users. Spammers are identified by observing abnormalities in the structural properties of the network. Another novel approach has been proposed which creates a Bayes network out of email exchanges to detect spam. Bayesian classifiers scan the contents of the email to calculate probability distributions for every node in the network.

Many research communities consider following three main families of techniques:

1. It fight against mail servers that are responsible for the generation of spam messages

   Blacklist: It contains list of misbehaving servers that are collected by several sites.

   Whitelist: It contains list of addresses of servers which are trusted not to propagate spam.
2. It considers content of messages and exploits the fact that spam typically falls within predefined categories and it is possible to distinguish spam based on its content. These methods are rule-based filters and Bayesian word distribution filters.
3. Same information is sent to many users, though spammers try to disguise it by creating a specific version of the message for each user.

Another approach detects spam emails by using Bayesian calculation for single keyword sets and multiple keyword sets along with its keyword contexts to improve the spam detection. This approach needs a large amount of memory and much hardware for execution, so workload increases.

Query Based Anti-Spam Technique A query based cross layer approach to detect spam follow some steps:

1. **Analyze the mail content:** This approach analyze the mail content and sender mail address of the mail, then cross analyze and compare the content and sender address of the previous spam mails. If content and sender address are already present then it declares that mail as a "spam".
2. **Trusted Knowledge Base:** In trusted knowledge base, database of trusted sender is stored over the inbox based on the frequency of the communication of mails. If sender is not the

trusted sender then following steps are needed to execute to identify the spam mails.

3. **Keywords knowledge base:** This stores the spam keywords. When any mail is received by system, this approach analyzes the keywords of mails with keywords knowledge base of spam. Then spam is declared as spam or useful mail on the basis of result.
4. **Sender mail address:** It verifies mail address of sender using mail header.
5. **Sender Location:** This approach finds the location of mail server and compares the location with spam mails location.
6. **Misbehavior of incoming mail:** Artificial Neural Network is used to predict any misbehavior of incoming mails [4].
7. **Cross Validation:** In this step, system will verify the sender that sender is a genuine human user or machine generated user using some cross request.

This approach needs a large amount of memory and much hardware for execution, so workload increases.

Content Independent Technique Other content-based filters provide some temporary relief from spam. But these filters are not robust enough against spammers. Spammers can easily fool these filters. So content-independent filter is needed [3]. A distributed, content independent, spam classification system called trinity has been proposed. Trinity is based on the following observation: Bots send a large number of e-mails in a short amount of time. If an e-mail is received from an unknown source that has sent many e-mails in a short period of time, then the likelihood of this being spam is high. But in this approach, right tradeoff between the security and weight of the protocols is needed. Trinity must have large number of peers to become effective. PCA has also been used for spam detection. SpamNET has been introduced for effective spam detection which uses heuristic rules, PCA and neural networks [5]. This program is able to adapt itself according to environment in which various users send mails. This program retrains itself after every seven days. SpamNET has used Bayesian classifier as well as neural networks so processing power is high.

Extractor module in SpamNET extracts out words which are common in spam mails. Then mail is passed to PCA. PCA is dimensionality reduction technique. It selects input vectors that are correlated to each other. PCA contributes to decrease error rate and increase efficiency. Then neural networks have been used which declare mail as spam or ham mail. Short-circuit rule is also available in SpamNET which directly declare mail as spam if specific features are available in mail and it will directly declare mail as ham mail if user has already responded to that mail address. In this program, 1.5 % false negatives are recorded and 1.3 % false positives are recorded. Various types of neural networks have been used to detect spam [6]. Neural Networks are able to detect features which can be detected by human. They state complex relationships between input and output. They make system adaptable so that system can adjust it according to changing environment. A huge number of techniques and solutions have proposed to detect spam but every technique has some pitfalls.

V. CONCLUSION AND FUTURE SCOPE

In this paper we presented a review of various reasons why emails are classified as spam and various criteria on basis of which they are classified. Later on a brief review of literature is given in the form of blueprint. And then the various general steps in spam detection are defined with various types of method to determine the spam and filter them from genuine email. The research can further be extended to the study of extractor module of PCA and also neural network discussed in introduction section.

VI. REFERENCES

[1]. Gaurav Kumar Tak and Shashikala Tapaswi, "Query Based approach towards spam attacks using artificial neural network", International Journal of Artificial Intelligence & Applications, October 2010

[2]. Grigorios Tzortzis and Aristidis Likas, "Deep Belief Networks for spam filtering", 19th IEEE International Conference on Tools with Artificial Intelligence, GR 45110, Ioannina Greece (2007)

[3]. Alex Brodsky (Canada) and Dmitry Brodsky (USA), "A distributed content independent method for spam detection".

[4]. A.Hyvarienen and E.Oja, Independent Component Analysis and Applications, Neural Networks 13(4-5):411-430, 2000.

[5]. Ann Nosseir , Khaled Nagati and Islam Taj-Eddin," Intelligent Word-Based Spam Filter Detection Using Multi-Neural Networks",IJCSI, Vol. 10, Issue 2, No 1, March 2013

[6]. Martin, Spam Filtering using Neural Networks, an internet draft http://www.web.umr.edu/~bmartin/378Project/report.html

