Efficacious Spam Filtering and Detection in Social Networks

U. V. Anbazhagu1*, J. S. Praveen2, R. Soundarapandian2 and N. Manoharan2

1Department of Computer Science and Engineering, Sri Lakshmi Ammaal Engineering College, Chennai, India; anbuveera@gmail.com
2AMET University, Chennai, India; praveenjs1985@gmail.com, rsoundar88@gmail.com, Directorresearch@ametuniv.ac.in

Abstract

In Online Social Networks, the internet mail server spam delivery is the most common issue. Email spam, also known as junk email or unsolicited bulk email (UBE), is a subset of electronic spam involving nearly identical messages sent to numerous recipients by email. In the Receiver Side, only most of the modern spam-filtering techniques are deployed. They may be effective in selection junk mail for clients, but junk mail communications however preserve losing World-wide-web bandwidth along with the storage space of email hosting space. In existing system, the Bayesian spam filters are easily poisoned by clever spammers who avoid spam keywords and add many harmless words in their emails. The detection system was proposed to monitor the simple mail transfer protocol (SMTP) sessions and email addresses in the outgoing mail messages from each individual internal host as the features for detecting spamming messages. Due to the huge number of email addresses observed in the SMTP sessions, Bloom filters are used to detect the spam messages and to increase efficiency.

Keywords: Bloom Filter, Spam Filtering, Social Networks.

1. Introduction

A social network is a social structure made up of a set of social actors (such as individuals or organizations). Most social network services are web-based and provide means for users to interact over the Internet, such as email and messaging. Web-based social networking services make it possible to connect people who share interests and activities across political, economic, and geographic borders. Some other social networks have further features, such as the ability to create groups that share common interests. Internet email is one of the most popular communication methods in our business and personal lives. Nowadays spam messages are becoming a continuous problem in email systems. Spam emails interfere with both email service providers and end users. Email spam, also known as junk email or unsolicited bulk email (UBE), is a subset of electronic spam involving nearly identical messages sent to various recipients by email. Blank spam may also occur when a spammer forgets or otherwise fails to add the payload when he or she sets up the spam run.

A spam filter should be personalized, and user-friendly. A more accurate filter generates less false positives and false negatives. False positives are legitimate emails that are mistakenly regarded as spam emails. False negatives are spam emails that are not detected. There are two primary types of spam filter attacks: poison attacks and impersonation attacks. In a poison attack, many legitimate words are added to spam emails, thus decreasing its probability of being detected as spam. In an impersonation attack, a spammer impersonates the identities of ordinary users by forging their IDs or compromising their computers. Spam filtering approaches can be mainly divided into two categories: content-based and identity-based. In the content-based category, emails are parsed and scored based on keywords and patterns that are typical in spam. The simplest identity-based spam filtering approaches are blacklist and white list, which check the email senders for spam detection. White lists and blacklists both maintain a list of addresses of people whose emails should not and should be blocked by the spam filter, respectively. One server-side solution records the number and frequency of
the same email sent to multiple destinations from specific IP addresses. Identity-based spam filters identify spam based on the identities of email senders. Most modern spam-filtering solutions are deployed on the receiver side. These filters are good at filtering spam words for end users, but the spam messages are wasting Internet bandwidth and memory storage.

To detect spamming bots, use the detection system to monitor the SMTP sessions and track the number and the uniqueness of the recipients email addresses in the outgoing mail messages from each individual internal host as the features for detecting spamming bots. Due to the huge number connected with email deals with affecting the SMTP periods, shop the deals with along with deal with these people effectively inside the Bloom filters.

2. Related Work

In 2005, Delany, and Cunningham have discussed about the spam filtering system that uses machine learning will need to be dynamic. Case-Based Reasoning (CBR) is a lazy approach to machine learning where it is delayed to run time. In this paper, the detailed description of such a system called evaluate design decisions concerning the case representation. It compares the performance with an alternative system that uses Naive Bayes. The initial stage of this research focused on identifying the most appropriate case base configuration for a case-based classifier for spam filtering. CBR as a lazy learner offers significant advantages; it provides capabilities to learn without the need for a separate learning process and facilitates extending the learning process over different levels of learning.

Yu, Kaminsky, Gibbons, and Flaxman in 2006 had discussed about the sybil attack, a malicious user can create multiple fake identities and pretends to be multiple, distinct nodes in the system. This paper presented Sybil Guard, a novel decentralized protocol for limiting the corruptive influences of sybil attacks, by bounding both the number and size of sybil groups. Sybil Guard relies on properties of the users’ underlying social network, namely that (i) the honest region of the network is fast mixing, and (ii) malicious users may create many nodes but relatively few attack edges.

In 2007, Haider, Brefeld, and Scheffer had proposed the problem of detecting batches of emails that have been created according to filter spam more effectively by exploiting collective information about entire batches of jointly generated messages. A sequential decoding procedure and it derives the corresponding optimization problem of supervised clustering. A sequential clustering algorithm and two integrated formulations for learning a similarity measure to be used with correlation clustering. Using the batch information, the email spam classification performance increases largely the information. The efficiency of the clustering algorithm makes supervised batch detection in enterprise-level scales, with millions of emails per hour and thousands of recent emails as feasible.

The problem of filtering spam messages for many users is each user receives messages according to an individual or unknown. A hierarchical Bayesian model also generalizes across users by learning a common prior which is imposed on new email accounts. It improves the performance of a personalized spam filter provided that the inbox contains sufficiently many messages. The Dirichlet-enhanced bias correction method estimates and compensates for the discrepancy between labeled training and unlabeled personal messages, learning from the new user’s unlabeled inbox as well as from data of other users.

3. Proposed System

In proposed system, the spam filtering techniques are going to deploy on the sender side itself. By this, the spam message cannot send to the receiver side. Spam email may also include malware as scripts or other executable file attachments. Before sending the mail it can able to filter the spam, normally some files come with an extension of .exe and encrypted file sent to the mails. Using filtering technique the junk mail can neglect the encrypted spam too. This system will improve bandwidth and memory storage. There are two techniques which are used to find the encrypted format text, Word Net dictionary and short message technique. The Bloom filters are used to find the junk mail, it have an advantage on other data structures for representing sets, such as self balancing binary search tree. This prevents text based spam filters from detecting and blocking spam messages. In addition, it is used to improve the performance of Online Social Network.

In previous system, spam mails are filtering on the receiver side. The sender mail before filtering, so spamming activities still exist, and spam messages still waste Internet bandwidth and the storage space of mail servers. Spamming bots may access web mail interfaces or deliver via secure SMTP for spamming. Since the packets are encrypted, the detection method cannot identify the spamming bots in this system. Bayesian spam filters need a considerable amount of time to adapt to a new...
spams based on user feedback. A Bayesian filter has a list of keywords along with their probabilities to identify an email as a spam email or a legitimate email.

4. Content-based Approaches

The basic approach of content-based spam filtering is the static keyword list, which however makes it easy for a spammer to avoid filtering by tuning the message. The second category of content-based approaches includes machine learning based approaches such as Bayesian filters, decision trees, Vector Machines, Bayes Classifiers and combinations of these techniques. In this approach, a learning algorithm is used to find the characteristics of the spam and of legitimate emails. Then, future messages can be automatically categorized as highly likely to be spam, highly likely to be legitimate emails, or somewhere in between. The third category of content-based approaches is collaborative spam filtering. Once a spam email is detected by one user, other users can avoid the spam later on by querying others to see if their received emails are spam or not.

4.1 Identity-based Approaches

Identity-based spam filters identify spam based on the identities of email senders. The simplest identity-based spam filtering approaches are blacklist and whitelist, which check the email senders for spam detection. Both Whitelists and blacklists maintain a list of addresses of people whose emails should not and should be blocked by the spam filter, respectively. The server-side records the number and frequency of the same email sent to multiple destinations from specific IP addresses. If the number of spam mails increases, the node with the specific IP address is blocked and the people cannot use the email account again.

A Bayesian filter has a list of keywords along with their probabilities to identify an email as a spam email or a legitimate email. It used to identify an email as spam depending on the probability based on user. The user can identify an email as junk mail or not.

4.2 Social Closeness-based Spam Filtering

When a person receives an email from another socially close person, the email has a low probability of being spam unless the email sender's machine is under an impersonation attack. Thus, the social closeness between individuals can be utilized to improve the accuracy of spam detection. Note that, in a social network, people treat others differently based on their social closeness. People impose different levels of interest, trust, or tolerance to the emails from others with different social closeness. People with close social relationship are willing to receive emails from each other. The emails containing spam keywords from senders that are socially far away.

4.2.1 Closeness Algorithm

1: Send a query message with TTL
2: if Receive a response from destinations then
3: Calculate its closeness with each node using Eq.
4: end if
5: if Receive a query initiated by node i then
6: Insert its closeness with node ito the message
7: TTL=TTL-1
8: if TTL!=0 then
9: Forward the message to its neighbors
10: else
11: Send the message to node i
12: end if
13: end if

It regards the spam as emails that receivers are not interested in. Therefore, it needs to differentiate emails from persons with different social closeness. SOAP loosely checks emails between individuals with high closeness and strictly checks emails between individuals with low closeness. The social closeness-based spam filtering module checks emails based on the closeness between the receiver and the sender. A smaller closeness rate leads to stricter checking, while a larger closeness rate leads to looser checking.
4.3 Bloom filter
A Bloom filter is a data structure optimized for fast and space-efficient. Bloom filters have a strong space advantage over other data structures for representing sets, such as self-balancing binary search tree. Stable Bloom filters as a variant of Bloom filters for streaming data. Stable Bloom filters continuously remove old information to make room for more recent elements. Since old information is ejected, the Stable Bloom filter introduces false negatives, which do not appear in fixed bloom filters. It shows that a tight upper bound of false positive rates is guaranteed, and the method is better to standard bloom filters in terms of false positive rates and time efficiency when a small space and an acceptable false positive rate are given. Bloom filters that can adjust dynamically to the number of elements stored, while assuring a minimum false positive probability rate. This technique is based on sequences of standard bloom filters with increasing ability and false positive probabilities, so as to make sure that a maximum false positive probability can be set earlier, regardless of the number of elements to be inserted.

4.4 Problem Analysis
Bloom filter used to refer the recipients’ domains to detect spamming bots with the assumption that the bots in the same botnet will target at similar domains and form a large cluster. The statement could be inaccurate due to the popularity of some mail services such as Gmail and Yahoo, which own a huge number of users. It is likely that the recipients’ domains are similar, but the REAs (Recipients Email address) are mostly different. Moreover, the spamming bots may not send spam to similar domains because they are not in the same botnet.

The spamming bot should deliver spam messages to a wide range of unique REAs for efficient spam delivery. It compares the number of total REAs, unique REAs and unique recipients’ domains in the spam messages from the spamming bots.

4.5 Detection Accuracy
The accuracy rate as the ratio between the number of successfully classified emails and the number of all received emails. The interest-based spam filtering component increases the spam detection accuracy by filtering out the emails in the receiver’s disinterests and accepting the emails that match the receiver’s interests.

A node can send a spam message to more friends in its social network as the nodes in RE within the same social network are likely to trust each other when the spammer sends the spam into the social network. A more accurate filter generates less false positives and false negatives. False positives are legitimate emails that are incorrectly regarded as spam emails. False negatives are spam emails that are not detected.

There are two primary types of spam filter attacks: poison attacks and impersonation attacks. In a poison attack, many legitimate words are added to spam emails, thus decreasing its probability of being detected as spam. In an impersonation attack, a spammer impersonates the identities of ordinary users by forging their IDs or compromising their computers. In contrast, Bayesian completely relies on data training and needs a significant amount of data and time to learn a new spam keyword.

The false negative (FN) rate is the number of false negatives divided by the total number of emails, and the false positive (FP) rate is the number of false positives divided by the total number of emails. A user usually recovers valid emails in the junk box by using the not spam function, and deletes spam emails without reading them.

4.6 Experimental Results and Performance Analysis
The performance of SOAP with bloom filter is used to analyze the false positive and false negative. The false negative (FN) rate is the number of false negatives divided by the total number of emails, and the false positive (FP) rate is the number of false positives divided by the total number of emails.

The False Negative rate of Bloom filter decreases as the sample size since more samples enable filter to learn more about the users and personal preferences. In SOAP, some emails that disinterest a receiver can be determined directly from the personal profile without training, thus

![Figure 2. Accuracy.](image-url)
the False Negative rate does not greatly vary as the sample size increases. Bloom component can learn those (dis) interest keywords, which enable SOAP to detect the spam that cannot be identified by the (dis)interest keywords checked by the poorly trained Bloom component.

The results imply that more social information helps the interest-based spam filtering component to infer more personal preferences for spam detection. Bloom has a higher False Negative rate than SOAP when the completeness is greater than 0 and the completeness does not have any effect on Bloom, since it does not believe social factors in spam detection.

5. Conclusion

In this paper, a Social network Aided Personalized and effective spam filter (SOAP) is proposed to meet the problem. In SOAP, nodes form a network for connecting the social friends. Each node uses SOAP to prevent spam separately. SOAP integrates a new component such as Bloom filter. The social closeness-based spam component prevents spam poison attacks, the social interest-based spam component helps to realize personalized spam filtering, the adaptive trust management component prevents impersonation attacks, and the friend notification scheme improves the power of a collective of socially close nodes to strengthen SOAP's ability to answer the impersonation attacks. Accurate spam filtering results function as input for Bloom filter used to have automatic training with reduced user effort to divide spam emails. The results of prototype based experiments show that SOAP improves on the performance of the basic Bloom filter in term of spam detection accuracy and training time. In future, the performance and security of the social network can be improved by SOAP.

6. References