QUERY BASED APPROACH TOWARDS SPAM ATTACKS USING ARTIFICIAL NEURAL NETWORK

Gaurav Kumar Tak and Shashikala Tapaswi

ABV- Indian Institute of Information Technology and Management
Gwalior (M.P.), INDIA
gauravtakswm@gmail.com, stapaswi@hotmail.com

ABSTRACT

Currently, spam and scams are passive attack over the inbox which can initiated to steal some confidential information, to spread Worms, Viruses, Trojans, cookies and Sometimes they are used for phishing attacks. Spam mails are the major issue over mail boxes as well as over the internet. Spam mails can be the cause of phishing attack, hacking of banking accounts, attacks on confidential data. Spamming is growing at a rapid rate since sending a flood of mails is easy and very cheap. Spam mails disturb the mind-peace, waste time and consume various resources e.g., memory space and network bandwidth, so filtering of spam mails is a big issue in cyber security.

This paper presents an novel approach of spam filtering which is based on some query generated approach on the knowledge base and also use some artificial neural network methods to detect the spam mails based on their behavior, analysis of the mail header, cross validation. Proposed methodology includes the 7 several steps which are well defined and achieve the higher accuracy. It works well with all kinds of spam mails (text based spam as well as image spam). Our tested data and experiments results shows promising results, and spam’s are detected out at least 98.17 % with 0.12% false positive.

KEYWORDS

Artificial neural network, Spam, Scam, Cross Validation, Virus, Worms & Trojan

1. INTRODUCTION

Now days, Email (Electronic mail) communication plays a great role in the human life due to its fast and free availability, lower or free cost. It is more useful for many corporate because of some features like newsletters, business correspondence, Email marketing, Advertisements etc.

Like Freelancer.com Support use email service for business correspondence to send the emails and messages to its authorized members. Google news alerts use it for the news letter. Naukri.com, DevNetworkIndia.org and etc. use email service for the new jobs advertisements massively. Inkfruit, ZoomIn, Fashnvia.com (India) and etc. use email service for their product marketing and their advertisements. Many times, these mails like Product advertisements, job advertisements, news alerts are meaningful for the email users but sometimes, they generate spam mails over the mail-inbox.

Today, Email and chat services are the most common, instantaneous and successful Internet applications, which are threatened by spam mails and spam chats. These Service can be accessed using mobile internet or low speed internet. Spam mails can be an advertisement or notification of porn website, porn video, phishing website, Nigerian scam, medicines advertisements, adult content etc.
Spammers collect e-mail addresses from chatrooms, public networking websites, customer lists, newsgroups, and worms, viruses which harvest users' address books, and are sold to other spammers. They also use a practice known as "e-mail appending" or "ependning" in which they use known information about their target (such as a postal address) to search for the target's e-mail address. Much of spam is sent to invalid e-mail addresses. Spam averages 78% of all e-mail sent [14].

The spam detection problem seems more serious over mailboxes today. Without a spam filter, one email user might receive over hundreds of mails daily and find that most of them are of spam category. Spam mails consume unnecessary traffic over the internet as well as email service provider. Moreover, receiving spam mails are with no use for email users.

In the employed system, a highly simplified architecture of artificial neural networks is used to detect the misbehaviour of incoming mails.

An artificial neural network is a mathematical model which works on the principles of biological neural networks. Generally it is referred as neural network (NN). Using neural network model; we can easily map the complex inputs with the complex outputs.

Some of the silent features of ANN are as follows,

- They represent a highly connected network of neurons - the basic processing unit.
- They operate in a highly parallel manner.
- Each neuron does some amount of information processing.
- It derives inputs from some other neuron and in return gives its output to other neuron for further processing.
- This layer-by-layer processing of the information results in great computational capability.
- As a result of this parallel processing, ANNs are able to achieve great results when applied to real-life problems.

A typical architecture of neural network is depicted in figure1.

![Figure 1. Architecture of Neural Network](image-url)
Neural network performs its operations in two phases: learning phase and testing phase.

**Learning Phase:** In the proposed methodology, we have taught several SQL attacks to the network in a supervised manner. We entrust the system with several variants of any attack and assign it a particular label. Thus we can see that system learns by feeding various patterns of the same attack.

During the *training process of neural network*, matrix of inbox mails and spam mails is used as input matrix to the neural network. In the proposed methodology, the input matrix is updated after defined time interval.

Any neural network adjusts the weights of attacks in order to learn in a supervised or unsupervised manner.

In our method of learning, each candidate attack taught to the network is associated with a weight matrix. Weight matrix associated with the $k^{th}$ spam is assigned the label $W_k$. Weight matrix is updated with the progress of the learning of the spam mail. This matrix is initialized to zero when learning phase starts. An input pattern corresponding to the spam is taught to the submitted to the network.

According to information compiled by Commtouch Software Ltd., E-mail spam for the first quarter of 2010 can be broken down as follows[15].

<table>
<thead>
<tr>
<th>Topic</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmacy</td>
<td>81%</td>
</tr>
<tr>
<td>Replica</td>
<td>5.40%</td>
</tr>
<tr>
<td>Enhancers</td>
<td>2.30%</td>
</tr>
<tr>
<td>Degrees</td>
<td>1.30%</td>
</tr>
<tr>
<td>Casino</td>
<td>1%</td>
</tr>
<tr>
<td>Phishing</td>
<td>2.30%</td>
</tr>
<tr>
<td>Weight Loss</td>
<td>0.40%</td>
</tr>
<tr>
<td>Other</td>
<td>6.30%</td>
</tr>
</tbody>
</table>

Figure 2. Spam e-mails distribution by topic
Due to following characteristics, currently the identification process of spam mails is a difficult problem [3].

[1] Spam heterogeneity
[2] Spam definition

Figure 3. Represents the spam distribution over various countries.

By continent, Asia continues to dominate in spam, with more than a third of the world's unsolicited junk email relayed by the region. Asia covers 34.8% spam mails over all the spam mails. The breakdown of spam relaying by continent is as follows [6]:

Figure 4. Spam distribution over various regions.

2. SPAMMER APPROACHES AND THEIR ATTACK

There are many techniques adopted by the spammer or attackers to collect and store the email addresses or personal information etc. Some of those approaches are from posts to UseNet with
email address, from mailing lists, from web pages, from various web and paper forms, via an
Ident daemon, from a web browser, from IRC and chat rooms, from finger daemons, from AOL
profiles, from domain contact points, by guessing & cleaning, from white & yellow pages, from
a previous owner of the email address, by having access to the same computer, using social
engineering, from the address book and emails on other people's computers, buying lists from
others, by hacking into sites and etc.

![Figure 5](image.png)

With a marketing service, a person can arrange his contacts by certain demographics so that he
can create custom mailing lists. This means that he can have some newsletters that go to all
customers while also having some that only go to women or men or people with a history of
shopping in a particular category. These tailored mailing lists ensure that your messages are
only received by customers who may be interested in the subject matter, keeping those who
likely would not be from feeling as though they are being spammed and unsubscribing.

Currently, a lot of social networking sites exits over WWW. Some sites are really useful but
some creates spam mails over the mailbox. With social networking sites, when a person joins
some social networking websites (like shstyle.fm, yaari.com, indiarocks.com, mycantos.com,
facebook.com, tagged.com etc.), then these social site use some script to approach contacts (contact mail list) of that person and send invitation to his contacts to join the same social site.
Many times they fill spam mails in peoples' inbox using this approach. There are also many
several attacks over the mailbox by the spammers. Some spammers generate spam mails over
the mailbox using the manual script but some use machine generated scripts to generate the
spam mails.

3. RELATED WORK

In literature, there are many techniques described for the detection of spam and mail filtering.
Some of the techniques are described as follows:

In [16], A Rule approach has been proposed for the detection of spam mails. The discussed
approach uses the training and testing phases of data. Moreover, the stale and obsolete spam
rules suspend during the training. This action is used for improving the spam filtering
efficiency. However, the time complexity is higher due to the rules generation and their execution. E. Damiani et al. discussed some basic properties of the spam mails. They focused on the reasons of the popularity of spam mails. The uses of the digests in the proposed approach to identify spam mails in a privacy-preserving way is a fundamental technique for collaborative filtering[3]. A social network is constructed based on email exchanges between various users in [11][12]. Spammers are identified by observing abnormalities in the structural properties of the network. Many times spammer uses the public social sites for increasing their mail list database. However, it is a reactive mechanism since spammers are identified after they have already sent spam. In [13] a novel approach has been discussed, which creates a Bayesian network out of email exchanges to detect spam. Though Bayesian classifiers can be used for detecting spam emails, they inherently need to scan the contents of the email to compute the probability distributions for every node in the network. Since many times it is not possible, to detect spam mails for the particular inbox and its requirement for filtering the spam mails [4].

Nitin Jindal et al. discussed an approach of review spam. Review spam is quite different from Web page spam and email spam, and thus requires different detection techniques[17]. There is an effective technique to detect the spam mail that is ‘Fast Effective Botnet Spam Detection’. It uses the header information of mails to detect the spam mails. It is useful for both ‘Text based spam’ as well as ‘image based spam’. It analyzes the sender IP address, sender email address, MX records and MX hosts [1].

One approach is also described to detect the spam mails, it use the Bayesian calculation for single keyword sets and multiple keywords sets, along with its keyword contexts to improve the spam detection [5].

4. PROPOSED METHODOLOGY

Before proposing a new methodology for spam detection, we are aware of this fact that most of time spam mails and scams are spread out using the machine generated script. In this paper, we are proposing a new query based cross layer approach for the above that is based on the above facts and some other spam features.

Our system uses some knowledge base and query generation using the history of previous mails and spam mails which is specific for the each user or its mailbox. Using the knowledge base, detection of spam mails is performed. It also maintains some keywords list, which can easily be pointed out as some words or content in the incoming mail, then perform the detection operation.

Many times when a person clicks a URL which is present in his mailbox, (that URL has been provided by the spammers) then mail address of the person is captured by the spammer and is easily inserted in spammer’s database.

Proposed spam detection approach, follow the few steps to indentify the spam mails which are as follows:

1) **Analyze the mail content:** Firstly, proposed 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 both are already present in any of the previous spam mails then it directly declares the mail as “a spam” (a spam is already present with the same sender and same mail content).
If the some fraction of incoming mail content matches with the any previous spam mail then mail is filtered using the spam threshold value ($S_t$). The spam threshold value can be defined as the mathematical value which decides the performance and accuracy of spam detection system. It can be different for various systems. It is used to indentify the spam mails with the partially matching case.

If $S_t=0.7$ and matching fraction of the content of mail matches with the previous declared spam mails is greater than equal to 0.7, then the mail is declared as “a spam”.

Matching fraction of the content = \( \max (NM_1/N_1, NM_2/N_2, \ldots, NM_P/N_P) \)

- $NM_p$: Total number of exactly matched words of incoming mail with the p-th spam mail.
- $N_p$: Total number of words in p-th spam mail.
- $P$: The total no. of recent mails which are available in the spam mail list corresponding to that user.

Using the analysis step, following mail from PHP-classes is detected as spam mail because it was already present in the spam folder and user never communicated with the sender mail id.

![Figure 6. A Spam mail from the PHP Classes](image)

2) Trusted Knowledge Base: Knowledge Base is always a good, efficient and faster approach to give the results based on historical data. It is used some queries to execute the results. It also follow some update operation to make the result efficient based on the system requirements.
In the Trusted Knowledge Base, database of trusted sender is stored over the inbox based on the frequency of the communication of mails. The Knowledge Base is also updated upon the requirement of inbox or threshold count of incoming mails.

This Knowledge Base is responsible to the detection of spam mails when sender of incoming is already kept in the trusted zone.

If the sender is not the trusted sender then next steps would be executed to indentify the spam mails.

3) Keywords knowledge Base: To execute this step, A knowledge base is maintained at mail server for each user which stores the spam keywords (already defined by the specific user). During this step, proposed approach analyzes the keywords of mails with the keywords knowledge base of spam which is prepared by the particular user for detection of spam. Using the result it decides that incoming mail belongs to the spam category or not. If incoming mail has not been declared as “spam” then execute the other steps to indentify the spam mails.

4) Sender mail address: Our proposed methodology extract the sender mail address using the mail header (check the from field or reply-to field to get the sender email address) and analyze it to indentify the spam. Using the sender email address, system finds that have any communication been done previously between receiver and sender or not? If receiver has already communicated with that mail address, then mail is declared as “not a spam”. But if receiver has never communicated, then system explores the contact list of the receiver.

If the sender mail address already present in the contact list then the mail is declared as “not a spam”.

This step is very useful with the public networking site because many times networking sites send invitation using someone contacts.

In the given figure, It is shown that we have received the spam mail in the inbox of vikas@decenttechnologies.org from Skoot.com server and our proposed approach is able to detect the spam easily.
In the above example, user payal (receipt email address: payal@decenttechnologies.org) has sender user himanshi.s@gmail.com in her contact list. So the received mail will be declared as "not a spam".

5) **Sender Location**: This step is useful when mail user receive a mail from the another country which already belongs to the spam mail country. Our approach finds the sender mail server location and then compares the location with the spam mails location. Using this step, we are able to filter out some lottery spam and some Nigerian scams too.

Using this step following mail is easily detected as spam mail because nation of mail inbox is INDIA and incoming mail server exists in US and receiver has never communicated with the US mail sender so it can be detected as spam mail.
Figure 9. A Lottery Spam to capture personal information of mail user.

Figure 10. Lottery Spam header to find out the sender mail location.
Many mail servers use the sender location approach to identify the spam so they ask to the user's country and location at the time of mail registration.

6) **Misbehaviour of incoming mail**: This step is executed using the artificial neural network. Artificial Neural Network (ANN) is a scientific discipline that is concerned with the design and development of algorithms that allow computers to adapt their behaviour based on data. ANN automatically learns to recognize complex patterns and makes intelligent and efficient decisions based on data.

In the spam filtering ANN learns the complex pattern of mails and makes intelligent, efficient decisions based on the incoming mail.

Proposed methodology executes training phase testing phase using sample set of the mailbox to complete this step.

During this step, we are able to predict any misbehaviour event of incoming mails; Machines generated mails, flood of mails over inbox. Misbehaviour can be predicted using the time factor, some sender mail address, some attacks.

To detect the Misbehaviour, training phase is executed after each threshold value of incoming mail over inbox.

7) **Cross Validation**: During this step, system will verify the sender that sender is a genuine human user or machine generated user using some cross request.

If the incoming mail is machine generated email, it implies that sender is not human user. So the machine generated mails are not able to validate their identity. Most of the spam mails are detected during this step.

5. **IMPLEMENTATION AND ANALYSIS**

We have conducted the analysis of spam mails using the proposed methodology on some inboxes of different peoples. We have created the environment using some web technologies HTML, script languages, AJAX, XML and MySQL tools for implementing the methodology. We also applied some basic concepts of PHP, AJAX, MySQL and JavaScript from the references [7] [8]. **Figure 8** represents the diagrammatic representation of the proposed methodology.
Figure 11. Diagrammatic representation of the proposed methodology.
1) Extract Mail Content

- Analyze the matching pattern and calculate the matching fraction with the previous spam mail and then compare the matching fraction with the spam threshold value.

\[ \text{If (matching fraction} \geq S_0) \]
\[ \text{then mail} = \text{'spam'}; \]
\[ \text{Exit;} \]
\[ \text{else} \]
\[ \text{Go to step2;} \]

2) To find the sender belongs to the trusted zone of the specific user then it performs some query operations. The Trusted Knowledge Base is responsible to maintain the status of the sender user. This Knowledge base is created using some frequent and recent received and sent mails

\[ \text{If (sender exists in trusted knowledge base)} \]
\[ \text{then mail} = \text{'not a spam'}; \]
\[ \text{exit;} \]
\[ \text{else} \]
\[ \text{Go to step3;} \]

3) Analysis the mail content using 'spam keywords knowledge base (already declared by the user).

\[ \text{If (mail content matches with the spam keywords knowledge base)} \]
\[ \text{then mail} = \text{'spam'}; \]
\[ \text{exit;} \]
\[ \text{else} \]
\[ \text{Go to step4;} \]

4) Analysis the sender mail addresses using the contact list and previous received mails.

\[ \text{Extract mail header then Separate sender mail address.} \]
\[ \text{If (sender mail address is available in (contact list or previous communicated mails)} \]
\[ \text{then mail} = \text{'not a spam'}; \]
else
    Go to step 5;

5) Detect the spam mail using the Sender location step.

**Sender_Location()**

{

S_location=find_location();

/* Using some crawling operations over the internet*/
/* Find the location of Sender mail server using the mail header*/

*If(S_location not belongs to the receiver Location/nation)*
Then mail = ‘a spam’;
Else
then mail = ‘not a spam’;
/* (sender belongs to the receiver location) */
Go to step 6;
}

6) It is the complex step of artificial network; we are not able to map the step using the functions. The step is executed using some artificial tools and API.

7) Detect the spam mail using the cross validation approach.

**Cross validation()**

{
Send (simple equation / puzzle, sender mail address)
If (validation=true)
Then mail = ‘not a spam’;
Else
then mail = ‘a spam’;
/* (sender is a machine user) */
We have recorded the incoming mail activities and sender mail addresses over 4 months (Apr, 2010 to July, 2010) at mailbox of an organization. We have not implemented our proposed methodology for the detection of spam mails in Apr, 2010 but during May, 2010 and July, 2010, we have implemented it and recorded the activities of incoming mails and also analyzed the behavior of incoming for the artificial neural network step. The following table data represents the recorded activities over the various mailboxes.

Table 2. Represents the data of recorded activities over mailboxes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Inbox</td>
<td>15870</td>
<td>17961</td>
<td>18460</td>
<td>17123</td>
</tr>
<tr>
<td>Spam</td>
<td>4692</td>
<td>7234</td>
<td>7494</td>
<td>7031</td>
</tr>
<tr>
<td>False Match</td>
<td>83</td>
<td>43</td>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>Total mail</td>
<td>19562</td>
<td>25195</td>
<td>25954</td>
<td>23157</td>
</tr>
<tr>
<td>% Spam Caught</td>
<td>24.8%</td>
<td>28.7%</td>
<td>28.9%</td>
<td>30.4%</td>
</tr>
<tr>
<td>% False Match</td>
<td>0.42%</td>
<td>0.17%</td>
<td>0.089%</td>
<td>0.099%</td>
</tr>
</tbody>
</table>

We can get the performance information of the proposed methodology using the experimented results which are shown in Table 2. We can easily compare these results and performance with the previously described approaches of spam filtering.

Fig 12 and 13 represents all the complete scenario of experiments results.
6. CONCLUSIONS AND LIMITATION

Our work is inspired by a situation of large number of spam mails over the mailbox, those we have easily encountered. We have recorded the incoming mail activities of various mail boxes of an university server over 4 months and analyzed those mails to get the better results and better performance of spam filtering. From table data, we can all results of spam mails, inbox mails, false match easily for the given time period. The experiment results provide the complete scenario of the problem and accuracy of spam detection. Our system indicated that the spam
was filtered out with 98.17% with 0.12% false positive. Table 2 represents the recorded data over the 4 months time period.

Limitation of the proposed method is that it needs more hardware for the execution and higher memory space. So many times, it increases the workload of the mail server. So to implement the proposed methodology for large mail servers, we need intelligent mail servers which are can be reduced the time complexity and provide better performance of spam filtering. So that we can easily manage higher computation load. Due to more hardware specification and higher computation load, the cost of implementation of proposed methodology is much higher.

**ACKNOWLEDGEMENT**

The authors would like to thank ABV-Indian Institute of Information Technology and Management, Gwalior for the support provided for this work.

**REFERENCES**


Authors

Gaurav Kumar Tak is a student of 5th Year Integrated Post Graduate Course (B.Tech. + M.Tech. in Information and Communication Technology) in ABV-Indian Institute of Information Technology and Management Gwalior, India. His primary research areas of interest are Cyber Crime and Security, Wireless Ad-Hoc Network, Web Technologies.

S. Tapaswi is Professor in IT Dept., ABV-IIITM, Gwalior, India. She earned her Ph.D. (Computer Engineering) from Indian Institute of Technology, Roorkee, India in 2002, M.Tech (Computer Science) from University of Delhi, India in 1993 and B.E. from MITS, Gwalior, India in 1986. Her primary research areas of interest are AI, ANNs, Fuzzy Logic, Digital Image Processing, Computer Networks, Mobile Networks, Adhoc networks, Information Security etc.