

# REVIEW AND ANALYSIS OF SPAM BLOCKING APPLICATIONS

Rami Khasawneh, Acting Dean, College of Business, Lewis University,  
[khasawra@lewisu.edu](mailto:khasawra@lewisu.edu)

Shamsuddin Ahmed, College of Business and Economics, United Arab Emirates University,  
[sahmed@uaeu.ac.ae](mailto:sahmed@uaeu.ac.ae)

## ABSTRACT

*The number of users using the Internet is increasing; as a result, the number of data and information that is being transferred is increasing. In a single day, you might receive over one hundred emails. A large number of these emails are junk emails. Unsolicited and generally unwanted email messages, commonly called Spam, have increased to such a level that users demand for blocking these unwanted messages is increasing. Using the Internet to send unsolicited messages is not new; however, the rapidly increasing proliferation of these unsolicited messages has many calling for new laws to severely restrict or block these messages. Legal changes are slow in coming and the Internet's borderless technology may render new laws ineffective. Spam blocking software is emerging as the solution. There are a number of Spam blocking technologies that are used in the market. This paper introduces Spam, presents a number of Spam blocking applications, reviews Spam blocking technologies: challenge response and pattern-match filtering, analyzes each of these methods and provide recommendations.*

**Keywords:** Spam Application, Spam Technology, Challenge Response, Pattern Matching, Pattern Filtering.

## INTRODUCTION

The Internet is still experiencing growth, which is continuing for several years. Also the number of users accessing the Internet is increasing; as a result, the number of emails sent over the Internet is increasing. A large number of these emails are junk emails, which is commonly called Spam. Unsolicited and generally unwanted email messages, commonly called Spam, have increased to such a level that users demand for blocking these unwanted messages is increasing. America Online reported that they block an average of 2.4 billion Spam messages every day [1]. Due to vast public pressure, the United States Senate overwhelmingly passed an anti-spam bill 97-0 on October 23, 2003 [1]. This demand for spam relief and the tremendous network cost to almost every organization is creating rapid growth in the spam blocking industry.

Spam Blocking software is responsible of allowing only wanted email messages to be in accessed by your email client.

## **History of Spam**

Electronic junk mail, also known as spam, is a problem that is no longer restrained by the physical limitations of paper and postage. Internet spammers can send out thousands of messages relatively easily and cheaply. Although spamming may be a relatively new problem, it was not unanticipated by IT developers more than 25 years ago.

In 1975, Internet pioneer Jon Postel realized that there was a fundamental flaw in electronic mail. As long as an e-mail message was being sent to a valid address, there was no way for a mail server (known as Host) to refuse a message from the network [3]. Postel published an article stating that it would be useful for a host to be able to decline messages from sources it believed were forwarding improper messages. At the time Postel's article was published, technological limitations rendered development of Postel's ideas impractical. During the next eighteen years, the inability of a network to refuse certain email messages caused occasional problems for users and network administrators. Some of these malicious events caused large disruptions.

During the mid 1980s, chain letters and a virus labeled "The Christmas Virus" were among the more notable disruptive events. Throughout the 1980s, a number of computer facilities were embattled by Internet chain letters [3]. Postal chain letters are self-limiting: it takes paper, postage, and time to send them out. But with a computer, sending five or ten copies of a chain letters is easy and consumes little human time and effort. It did not take long for system administrations to be effected by the problems associated with chain letters.

The Christmas virus also created a number of problems for computer networks. The Christmas virus arrived as a file named CHRISTMAS in the user's directory. The program then forwarded itself to individuals with whom the victim corresponded frequently. For some organizations, this caused significant disruption with mail servers and their supporting storage media.

As a result of chain letters and the Christmas virus, many became aware of the problems that could be created by automated programs forwarding junk mail to thousands or millions of addresses.

## **SPAM BLOCKING APPLICATIONS**

Spam blocking applications prevent unwanted emails from arriving at the user's inbox folder. Spam blocking software use one of two primary methodologies: Pattern-filtering and challenge-response. Pattern-match filtering technology uses defined methods to evaluate email and classify it as spam. The degree of filtering sophistication varies from one software manufacturer to another and it is the level of this sophistication that determines the effectiveness of the spam filtering. One concern with filtering is the frequency of false positives; that is, valid email messages labeled as spam and not delivered.

Challenge-response technology blocks messages from unknown sources. This method causes a problem for the email addresses that are received by an email client for the first-time. Their email

message will be rejected. This type of technology has potential to create havoc for those wanting to forward legitimate email [4].

Spam Blocking software that incorporates filtering technology use various pattern-matching methods to target spam and stop it from being delivered to an end user. The filtering software solutions are typically more expensive to buy and install, but also require much less human intervention.

## **FILTERING SPAM BLOCKERS**

There are several software applications in the in the industry that use this method. These methods are different mainly in how restrict they are in filtering email messages. Some of these applications are so restricted, such that, they filter a high rate of wanted emails. Selecting one of these applications over another depends on false positive rates that are acceptable in your organizations.

### **BrightMail**

BrightMail is a spam blocking software. It is an enterprise level software solution designed to support larger organizations.

For a larger organization, the added cost of a more sophisticated solution, like BrightMail, may well be offset by the added value of the more sophisticated features and functionality of this software. BrightMail uses filtering technology to protect a client's email. This filtering software has multi-layered spam defenses that utilizes at least five different methods or technologies to detect and remove spam from the legitimate email environment. These filtering methodologies are defined in more detail below.

BrightMail performance metrics are shown below [2]:

- It blocks over 95% of spam.
- A BrightMail installation will include a direct connection from BrightMail where updated filtering rules are provided to the client's mail server every 10 minutes.
- The accuracy rate is 99.9999%, or one out of every one million legitimate emails is erroneously label spam and blocked.

BrightMail uses what is called a probe network to gather and analyze 400 million pieces of spam per month in order to enhance their vast rule set. This probe network utilizes decoy email accounts to attract spam for analysis. The "captured" spam messages are forwarded to a BLOC server that is used to generate and enhance the rules for dealing with spam [2].

Spam is identified and removed from legitimate email by BrightMail though utilization of at least five methods or techniques for filtering spam from the legitimate email traffic. The techniques include:

- Source filtering
- Heuristic filtering

- Header filtering
- Spam signature identification
- URL filtering

Source filtering refers to identifying and blocking spam based on the source of the originating sender. This includes email addresses, domains, and range of addresses. Heuristic filtering analyzes the header and body of an email message. It envelops information from incoming messages and checks for the presence of distinct spam characteristics. Header filtering traps messages using targeted header-based filters. Spam signature identification analyzes string of bits or “an email signature” in order to label incoming mail as spam. Finally, URL filtering utilizes URL matching in messages to label targeted messages as spam.

BrightMail also allows individual users to identify messages as spam and forward it to a spam folder agent. This folder is used to manually capture and forward suspect messages back to BrightMail for analysis. The BrightMail analysts will look to identify unique characteristics that can be used to update and or improve the BrightMail spam-filtering rule set.

### **SpamWeed**

SpamWeed is an alternative to BrightMail. This product offers many of the same capabilities as BrightMail; yet at a much lower cost.

SpamWeed supports POP3 and works as a proxy between the server and the email clients. It supports all POP3 email clients such as Outlook Express, Outlook, Eudora, Incredimail, Mozilla Mail, and more. It does not add tags to spam and leave the filtering task to the email program; it filters spam out of the inbox. Once spam is detected, SpamWeed will send it to a quarantine folder. Typically, after seven days, SpamWeed will delete any spam from the quarantine folder. Any email falsely labeled as spam in the quarantine folder can be easily retrieve by the intended recipient during the seven-day quarantine period.

SpamWeed utilizes two methods for on-demand filtering. The first of these methods is labeled checker and the other is proxy. The checker filter periodically polls the POP server looking for prospective spam messages. Suspected messages are downloaded to the quarantine folder and removed from the POP3 server. The downside to utilizing the checker method is that the mail program must be synchronized with the spam checker. This requires user intervention. The user needs to turn off the automatic email checker in their mail client to allow this.

Unlike the checker method, the proxy method is very easy to use. It imposes no added overhead to a user’s daily email communication. Once installed, an anti-spam proxy user simply continues to use their email program. With this method there is no need to launch a specific spam remover. All operations are transparent to the email program and user.

SpamWeed may also be setup to function as a POP3 proxy. When this functionality is utilized, SpamWeed intercepts communication between the email client and the POP3 server and automatically handles any spam messages based upon a user’s specified action. For an email user

that does not leave messages on the email server, the email client will issue DELE command for all good messages that are known to exist. When SpamWeed detects that the mail client has issued the DELE command, it assumes the user does not leave message on the server and therefore any messages for this user label as spam are also deleted.

SpamWeed also incorporates features that maintain tables named blacklist and whitelist. The blacklist feature in SpamWeed is used as a convenient method to block a known sender or receiving address. The whitelist feature is used to prevent known addresses from being mark as spam. SpamWeed automatically maintain an enterprise-wide whitelist called the system whitelist. SpamWeed will capture communication information from its users, such as the people communicated with on a regular basis. This information is automatically populated in the system whitelist and is used to prevent false positives.

## **CHALLENGE-RESPONSE SPAM BLOCKING**

Another type of spam blocking technology utilizes a challenge-response mechanism. These challenge-response spam blockers are typically installed to support one email user at a time. These spam blockers reject messages from all new senders and require a manual response to authorize a sender to email the protected email address.

When an email site is protected by challenge-response technology, the spam blocking software creates a list of valid senders for the email account that is being protected. When a protected email account receives an email message, the challenge-response software intercepts the message and validates the sender against a list of approved senders. When a message arrives from a new sender, a reply is automatically generated to the sender indicating they have sent a message being protected by challenge-response software. The new sender must manually reply to the challenge [4].

### **SpamArrest**

SpamArrest uses the challenge-response method for blocking spam. It creates a message to the new sender with a word imbedded in an image that humans can read but a program cannot decipher [5]. The new sender must read the word layered in the image and reply to the challenge to gain email access to the protected email address. For the protected site, the original message is typically placed in a holding file for a prescribed period of time (e.g. seven days) and deleted if the challenge is not addressed within this time period. Different challenge-response software may work slightly different, but they all follow essentially the same mechanism of requiring a new sender to manually validate their desire to email the protected recipient. These manual mechanisms are intended to prohibit spammers from programmatically replying to the challenges.

## **COMPARISON**

We are going to provide a comparative analysis between the two technologies. The first thing we need to do is to select criteria for comparison that ensures an objective and fair comparison of

both approaches. The following is a list of our criteria elements: 1. Cost of implementation. 2. Rate of false positives. 3. Rate of false negatives. 4. User involvement.

The cost of implementing the pattern-match filtering (PMF) technology is higher than the cost of implementing a challenge-response (CR) technology. PMF applications are more sophisticated and more powerful than CR applications. They use artificial intelligence technology to increase the degree of accuracy in spam blocking. Their implementation requires more rules and more coding. CR applications are much simpler; they use a smaller number of rules. It only requires a small database to store the wanted and unwanted e-mail addresses.

The rate of false positives (FP) and false negatives (FN) is zero using CR applications. It is the email user job to determine if that email address is wanted or unwanted. The application will add those email addresses to either the blocked or allowed email-addresses list. The rate of FP and FN is not zero using PMF applications. This rate tend to get lower as more rules are built into the applications, but will never be zero.

Using CR technologies, the user is very involved in the process of wanted and unwanted emails. This involvement does not decrease by time. On the contrary, it might increase depending on the number of new email addresses received by the user. Using PMF technologies, the user is only involved in the initial setup of the program and rules. The user involvement in the process will be minimal after the setup is done.

Prior to installing or using any spam blocking software, one should be aware of the consequences of the various types of spam blockers. For the individual, challenge-response spam blocking offers an affordable, easy to implement solution. There are, however, problems with this type of technology. Challenge-response technology essentially causes an Internet email user to push their spam-blocking problem into the hands of anybody else who may wish to legitimately email the user of challenge-response software. Any individual looking to install this software needs to give careful consideration to how they may impact others, specifically if they are a part of a mailing list. One example of the problems with this technology involved Dave Farber at the University of Pennsylvania. Mr. Farber recently advised individuals on his mailing list "If I start getting a flood of challenges from EarthLink IPers that require my response I will most likely declare them spam and you will stop receiving IP mail" [4].

There are other issues that need to be considered in the rush to block spam. Internet mail servers are encountering significant added expenses due to the large quantity of spam dumped on them every day. In attempt to deal with this issue some are beginning to lash out at those they believe are causing or facilitating these abuses [4].

## **CONCLUSIONS**

Challenge-response spam blocking offers an affordable, easy to implement solution, and leaves the control in the hands of the user. It is recommended for individuals and small organizations. Pattern-match filtering is very powerful, expensive but has the drawback of filtering many wanted email messages. It is recommended for middle-size and large-size organizations

Although challenge-response technology is able to eliminate spamming, however, neither technology is able identify if the sender's email address is authentic or not. This fact creates another vulnerability that will eventually increase the rat of false negatives and consequently spamming.

## **REFERENCES**

1. Associated Press, (2003). Senate Votes for Tough Spam Limits. Retrieved October 24, 2003 from [HTTP://www.FoxNews.com](http://www.FoxNews.com).
2. BrightMail Product Information. (2003). Retrieved November 4, 2003 from <http://www.brightmail.com/>.
3. Killers. (2003). How to Control Spam, Use Spam Blockers and Spam New York Times, Retrieved October 30, 2003 from <http://www.nytimes.com/2003/10/28/technology/28SPAM.html?hp>
4. McCullagh, D. (2003). Spam blockers may wreak e-mail havoc. Retrieved November 4, 2003 from <http://www.news.com>
5. Schwartz, A., & Garfinkel, S., (1998). Stopping Spam. O'Reilly & Associates, Inc.
6. SpamWeed Product Information. (2003). Retrieved October 25, 2003 from <http://www.spamweed.com>.