St. Bernard Managed Protection Services

White Paper

Spam Filtering

Building a More Accurate Filter

By Lawrence Didsbury, MCSE, MASE

Executive Summary

Spam issues and volume have been escalating in severity for many years. It is one of the key productivity, security and user policy enforcement issues facing small and large businesses today. Large businesses can afford to set up hardware specific, or server side software solutions that while effective, are resource and cost intensive. Small businesses are either underprotected, or rely upon client side, software solutions that don't meet all their needs. Managed Security Services, or "on demand" solutions have emerged as a clear leader in SMB solutions, offering high end performance, management and maintenance, and minimal entry cost.

On average, small business employees manually remove three to seven spam mails per day. If you have 100 employees, this10 minutes per employee adds up to 16 man/hours per day of lost time. Unfortunately, some spam headings are too interesting for readers to ignore.=How many of these emails lead your employees to waste further time responding to or checking out the uninvited promotions? How many of these email promotions are unethical or improper in a business environment? How much of your email server capacity is being used to process this unwanted email? How much of your bandwidth is being wasted on this injurious content? **This unsolicited bulk email is commonly called "spam".** Spam has become such a problem that it has grabbed the attention of lawmakers around the globe who are now seeking to address the problem with legislation. Business leaders, particularly those with small to medium enterprises, should understand the seriousness of the spam problem, and become familiar with the methods for combating it. Even the open source community has tackled the issue, although the resulting solutions have been hard to deploy.

So, why not just install some software or device that will keep the spam out of your business network? It sounds easy enough, but upon closer inspection, you may realize that one single technology or software cannot necessarily serve as a panacea for this problem. Some commonly used spam filtering methods are Bayesian filtering, blacklisting, fingerprinting, and heuristics. **Bayesian filtering** is a method that essentially uses algorithms for text classification and uses that classification to determine the probability that a certain email is in fact spam. **Blacklisting** takes the sender's account, ISP, IP address or other known item and adds it to a "black list". All emails from the blacklisted account are then forbidden to enter the network using a software filter. **Fingerprinting** takes into consideration the entire email, the header, content and body, and uses this information to create a unique fingerprint. The fingerprint is then compared to other incoming email to determine if it is spam. **Heuristics** takes the header, the content, and the routing metrics of the message to create a score, and compares that score against a score that would indicate that the email is spam. Other techniques can be combined in the effort to improve filtering like using "bait" email accounts to attract and capture spam in order to create databases against which to score or compare live email. All of these technologies will be examined in more detail in this paper.

In addition to the specific spam filtering technologies, the application of various philosophies behind spam filtering will yield different results when applied along with a particular technology. The philosophy typically depends upon the audience being served with the filtering technology and the entity that is providing the filtering, as we will see in the discussion. Related to this aspect is the method of technology deployment; another factor that business leaders will be interested in when seeking to minimize the effect of spam on their business resources. Providing spam filtering as a network service is one approach, and using a server-based software application within the network is another. Each approach has its benefits. This part of the discussion will also help Small and Medium Business leaders determine which is right for their business.

The following begins with a discussion of the overall problem of unsolicited bulk email or spam, and the challenges presented when attempting to keep unwanted mail from reaching the user's inbox or your mail servers. Next, we will look at some of the philosophies behind the filtering of email as a starting place for understanding how spam filtering is accomplished. The discussion will also involve a brief look at deploying spam filtering as a network service or as a software

application. With the groundwork laid, the paper will then delve into specific spam filtering technologies as indicated above, namely Bayesian or content-based filtering, blacklisting, fingerprinting, and heuristics. Additional techniques employed by the market leaders in spam filtering will be illuminated where possible. Finally, specific input is provided as to the St. Bernard Managed Protection Services approach to filtering, and ultimately the St. Bernard On-Demand value proposition.

St. Bernard is a leading provider of email, and other security protection services for global clients using either a Global Gateway Service or a Local Messaging Router technology. As we will see in the following discussion, St. Bernard Managed Protection Services uses a combination of technologies and deployment methods to filter spam from enterprise networks, removing the injurious costs of this unwanted email from email systems. Interested business leaders are encouraged to contact St. Bernard for more information toll free at **(800) 782-3762**or via email at info@stbernard.com. A free 30-day trial of St. Bernard Managed Protection Services is available by visiting www.stbernard.com/ondemand

SPAM: A Serious Problem?

As previously mentioned, "spam" is the common term for unsolicited bulk commercial email. There have been other uses for the term spam, but it is widely accepted today as meaning unsolicited email in general. Spam might include emails with offers to purchase products, emails with offensive content, "business opportunity" emails and the like. Why does spam pose such a threat? Why does spam have so many business leaders upset? The bottom line is that **spam wastes your valuable business resources**. Some examples of this waste are

- Employee time spent removing spam emails
- Employee time spent with the distractions of interacting with spam content
- Physical server resources consumed by spam
- Network bandwidth consumed by spam
- Network administrator time spent combating the effects of spam (like viruses)
- Legal costs of pursuing spammers

To understand the cost of wasted employee productivity alone, an example calculation is provided below. The basis for the calculation was research from Gartner and eMarketer, which estimated that each employee spends 49 minutes each day dealing with email, and that 38% of that email, is spam. If these figures are true, then each employee spends 18.6 minutes of their productive time each day on spam!

If your company has 100 employees, with an average hourly salary of $25, and each employee spends 18.6 minutes a day on spam, then the costs to your organization are the following:

- 31 hours are spent on spam each day at your company!
- 7874 hours are spent on spam each year at your company!
- $196,850 is wasted each year on spam at your company!

If viewed in a wider framework, the costs of spam spread beyond the corporation to society as a whole. The costs of sending large quantities of spam are typically borne by the recipients, not the senders. When the overall costs are calculated for millions of citizens, even though they may be distributed among the recipients, the results are staggering. Some of the societal costs of spam include:

- Money lost by people vulnerable to email scams
- Avoidance of useful or productive online tools and resources due to the fear or effects of spam
- General annoyance

In an article titled *"Unsolicited Bulk Email: Definitions and Problems"* the Internet Mail Consortium (IMC) stated, "There is no other common form of unsolicited communication that shifts so much of the cost of each message onto the recipients[1]. Clearly, spam wastes resources that could be better applied towards productivity. This is particularly evident within the Small and Medium Business enterprise market.

The SPAM Filtering Challenge

The business problems caused by spam may seem obvious but the problems are not confined to those just outlined. When attempting to solve the problem of spam using filtering technology, a new set of technical challenges becomes apparent, namely, how to design a filter that stops spam while still allowing legitimate email to pass through to the recipient. A simple example can be used to illustrate this filtering challenge. Let us say that a spam filter only searches each email header for the presence of the word "Free" and deletes each email occurrence identified. How can the spam filter distinguish between two messages, one having the header "Order your Free Adult DVDs" and the other with a header of "Are you free for lunch today?" In this example, both emails would be deleted by the system, preventing one spam message, yet also preventing one recipient from receiving a lunch invitation from a client. The legitimate message recipient would not even be aware of the blocked lunch message and would not respond to the client, possibly resulting in lost business. This is an example of what is called a "false positive", or a situation where an email is falsely identified as spam and blocked. A "false negative" is the opposite, where an email is falsely identified as legitimate when it is in fact spam. Of course, new filtering technologies may offer much more accurate results, as we will see later in the discussion, but the example above illustrates the general challenge of spam filtering. The challenge of spam filtering is how to stop the highest percentage of unsolicited or unwanted email, while still allowing legitimate email to pass to the intended recipient.

Philosophies Behind SPAM Filtering

A service provider is offering the type of spam filtering that or software vendor typically depends upon their philosophy regarding spam filtering. Their spam filtering philosophy will be developed based on the needs of the customers they are serving. In general there are two philosophies of spam filtering, one that provides a centralized filter and control over what is filtered, and another that puts the control in the hands of the individual email recipient. The two approaches might also be coined "consumer" and "Small and Medium Business" spam filtering:

Consumer philosophy: When the individual email recipient requires control over what is blocked or allowed to pass through a spam filter, the filter would be considered consumer-focused. A solution such as this might be more interactive, allowing an individual user to allow or reject particular emails, or have access to the filter settings. The easiest method for deploying this type of solution would of course be individual client software or a profile-based web service with accessible filter settings. In the consumer environment each individual customer will have a different definition of spam, and will each have different requirements. One user may want to allow "business opportunity" emails, even if unsolicited, while another may want to block this content. Consumers with young children will likely want to ensure that all adult content is blocked from their systems and may want to further define the kind of material to block. Again, the main point is that the consumer philosophy would put the control over filtering in the hands of the recipient.

Small and Medium Business philosophy:

As opposed to the consumer philosophy, the Small and Medium Business approach would put the control of over the spam filtering rules in the hands of a centralized administration. Corporations understand their business needs and further understand the objectives of various

departments that need to be served by the filtering technology. Having centralized control over email filtering allows a corporation to uniformly apply filters to serve these business needs. For example, sexual or adult email content presents legal and human resources nightmares for businesses. The very presence of such content within a business environment can lead to sexual harassment lawsuits even if the corporation at large intends no fault. Centralized spam filtering allows the corporation to prevent this and demonstrates a business position and specific action against such material.

In addition to the centralized control over the administration of policies allowed with the Small and Medium Business philosophy is the control over network and computing resources that is needed within the organization. The decentralized approach of the consumer model does not allow for this control. Such an implementation in a Small and Medium Business environment would increase the costs associated with spam filtering. If each client had individual software to deploy or profiles to maintain, the administrative and helpdesk costs alone could be staggering to the organization.

Service or Software?

As we began to see in the previous section, spam filtering can be deployed with different philosophical objectives. In addition to the philosophy behind the deployment is the deployment method itself. The most common methods are as a network service offered by a service provider or as software deployed within the organization. The software method can also be further broken down into a centralized or decentralized approach.

Spam Filtering Services:
A service-based approach to spam filtering can be quite beneficial to corporations as it takes the challenges involved with operating a network service and transfers them to the service provider. The service provider maintains the equipment and software that performs the spam filtering and provides a Service Level Agreement (SLA) to the corporation, which guarantees the availability and performance of the system. This removes considerable headaches from the IT staff at the corporation while still providing the centralized control needed for a business environment. Additionally, a corporation can likely treat the costs of such a service as a business expense, with all or part of the cost being tax deductible during the same period it is incurred.

Spam Filtering Software
Some corporations are very sensitive about having any of their IT infrastructure in the control of an entity outside of their corporation. Additionally, some are cautious about allowing their intellectual property (like email content) to pass through systems outside of their corporation. While most of these concerns are well handled by spam filtering service providers, some corporations remain adamant about this point. For the corporations that require this control, understand the challenges, and have the necessary IT resources to accomplish the filtering in-house, there are spam-filtering software products to serve their needs. There are two deployment methods for the software approach, namely server/network based, and client based deployment. In the Small and Medium Business environment, the client approach is not recommended, as implementation would not likely provide for policies in a uniform manner. The client approach would also create additional costs in deployment, support and maintenance of the individual client software. The server or network-based approach would be most desirable for a business environment, where the IT department installs and centrally manages the spam filtering software. This implementation would be more cost effective than a client deployment and would allow for the centralized administration of spam filtering policies to serve the requirements of the business.

SPAM Filtering Methods Explained

Spam filtering methods have improved greatly from the simplest text comparison methods into those available today. As the filtering innovations improve, so do the methods of the spammers,

creating a sort of cat and mouse game where spammers attempt to circumvent the filters. Even as legislation is passed to prevent spammers from plying their unethical email practices, filtering technology continues to be updated to aid in the battle. The following are common types of spam filtering technology in use today:

Blacklisting:
Some of the earliest spam filters were blacklists, static lists or databases containing a list of items that should be prevented from passing through the filter. The items contained in the email message could be a common word or phrase in the header, an IP address, domain name, ISP of the originator or email address of the originator, for example. Any email passing through the filter would be searched for the blacklisted item, and if found, the filter would block the message from passing. A number of problems are encountered if blacklisting is the only method of filtering implemented. For example, if a filter blocks the ISP of a spam sender; it is quite likely that legitimate email may be blocked from other customers of that same ISP. Likewise, if a blacklist approach is used to block certain key words or phrases called "spam markers", this approach cannot always distinguish between these and legitimate email messages containing the same words or phrases.

Blacklists were later improved to include the hashed contents of entire known spam messages, which could be compared to incoming messages to remove the undesired email. Larger databases could be assembled with these known offending pieces of spam, and since the hashed contents were quite small, lengthy comparisons could be carried out rapidly enough to provide a high quality of service. Well-known services such as ORBS, MAPS and SPEWS are examples of blacklisting services that use the improved techniques. A number of other services exist and many vendors employ proprietary blacklisting techniques as well. The primary problem with this method was that the spammers could either change the contents of the email message or use some automated, random character string to change the contents of each message sent, thereby bypassing the filter.

Another method related to blacklisting is **"whitelisting".** As one might imagine, a whitelist takes the opposite stance as a blacklist, namely assuming that all email is spam unless it passes the whitelist. In order to get on the whitelist, an IP, domain, email address or similar identifier must be confirmed as acceptable before email messages from that entity will pass the filter. This listing method can be effective, but restricts the rapid communication that is sometimes necessary in business. To illustrate the challenge of a whitelist, if a referral is made via word of mouth, and the referred individual sends an email to someone protected by the filter, the sender will have to go through a confirmation process before their messages can pass the filter. The confirmation process is usually as simple as clicking on a link in an auto-generated message sent by the filtering software, or responding to a particular email address for clearance, however this can be perceived as an unnecessary inconvenience by some.

There is certainly some value in the blacklist or whitelist approach to spam filtering. The fact, however, is that they can be only partially effective, and will likely not provide an effective filter when used alone.

Bayesian Filtering (Content Focus):
Bayesian filtering or content-focused filtering holds promise as a filtering technology, as it employs a more technically advanced method of analysis. Paul Graham, a well published technologist who has written much on the topic of spam, argued that content-based filtering is likely the way to succeed in filtering spam because it targets the actual message of the spammer.[2] His point is that the spammers are getting better and better at avoiding spam filter technology, but that regardless of their efforts, they will still have to get their message to the recipient in order to achieve their goal. By attacking the actual content of the message with Bayesian filtering, the spam filter is likely to be more effective.

Bayesian filtering is an extension of text classification technology, which searches the textual content of an email and employs algorithms to determine the probability that a certain email is spam. The algorithms are able to classify the occurrence of certain words and phrases in terms of how and where they appear in the email message, not by their existence alone. The filter can assign a probability to the occurrence and then determine if the email is spam based on predetermined rules. Paul Graham uses the example of classifying the word "free" in his article titled "Better Bayesian Filtering" to explain the concept.[3] The word "free" can be presented a number of ways to the filter, each with a different probability assigned. For example, "FREE" in the subject line of the email would be assigned a higher probability than the word "free" in the body of the email. The word "FREE!!!" in the subject line would be assigned an even higher value than the word "Free" in the subject line. Without taking the reader down a deeply technical path, an effective Bayesian spam filter would need to incorporate a well-conceived algorithm for the effective classification and weighting of content contained in email messages. Paul Graham states in that same article that he was able to write a Bayesian filter that was capable of catching the majority of the spam messages while generating very few false positives, demonstrating the power of this approach. While effective, Bayesian or content-focused approach is not the best panacea for spam filtering. One challenge with content filtering is that spam emails are increasingly designed with HTML, and often contain image links that download image-based content to the receiver. A content-based engine will likely not be able to determine between a valid link and one from a spam message. In spite of this challenge, a content-focus contains a powerful method for efficiently identifying unwanted content.

Fingerprinting:
Fingerprinting is a filtering technique that identifies a spam message and assigns a unique identifier to that particular email. The system then builds a database with all of the unique identifiers and compares each incoming message to that "fingerprint" to see if it matches. The filter then blocks all matching messages. Fingerprinting is typically effective only when identifying repeat messages, or messages arriving after the first one received is fingerprinted. Spam filtering services often use a **"bait method"** to catch and fingerprint spam messages in order to proactively build a comparison fingerprint database. A "bait" email address is used, which is not filtered, and even encourages the receipt of spam in order to capture and identify offending pieces.

The challenge with fingerprinting is the speed at which the fingerprint information is obtained and disseminated through the system. This method would be most effective for a service-based offering, as the database would be centralized and easily updated. For a spam filtering software, it would be difficult to update all of the software clients in a short enough time as to make the fingerprints effective. Again, as we are seeing throughout the discussion of spam filtering methods, no one method seems effective enough to defeat all spam. Likely, a combined approach is the best one to adopt.

Heuristics:
Heuristics is a method of filtering that "tests" each email based on a set of rules to determine if that email is in fact spam. The rules can be designed to test any part of the email (header, subject, content, etc.) or multiple parts of the email in order to be most effective. Any number of rules can be designed to complete the test, but a balance has to be achieved between the detail level of testing and the performance of the system. The result of each test conducted on the email message results in a "score" for the email, and the score then compared to a passing or failing score. Some of the tests applied using heuristics might include:

- Validating the senders return address using a hostname lookup
- Checking that hostname against the domain name for validity
- Validating the originating server hostnames found in the header
- Validating the message ID
- Validating SMTP relays

Heuristics will only be as effective as the rules or tests that are employed against the email. The benefit of the heuristics method is that the tests can be adjusted to meet the needs of the business organization for a definition of spam that fits their specific requirements. From a technical standpoint, the tests can be strengthened or reduced to optimize the use of the computing resources involved. Ultimately the rules will reach a balance of serving the business needs and optimizing the performance of the system.

The St. Bernard On Demand Process Approach (Value Proposition)

Through the discussion of the various techniques and technologies employed in the filtering of undesired email content, or spam, we have seen that no single technology provides a completely effective solution.

St. Bernard Managed Protection Services understands that spam filtering is a composite process, not just a technology.

This understanding led St. Bernard to design a more effective spam filtering process for small and medium business, education and government customers that includes a number of steps and technologies. In addition to understanding the need for a process-driven approach to spam filtering, St. Bernard understands that every organization is different and has different IT requirements and that an effective solution must provide for these requirements.

St. Bernard Managed Protection Services offers email protection as a managed Global Gateway Service or as a Local Message Router software product.

By including a number of the available technologies in their spam filtering process, St. Bernard **Managed Protection Services** is able to achieve a much higher level of success in blocking spam than their competitors. The St. Bernard On-Demand process includes the following methods:

1. **Bait email** - St. Bernard On-Demand maintains a number of bait email inboxes that encourage spam. These bait emails are used to capture spam which is logged into a database of known spam. This database can then be used in the next step of the process.
2. **Fingerprinting** - Emails are content filtered and the resulting spam is fingerprinted. Fingerprints are stored in the database of known spam for cross-referencing against future emails that are received.
3. **Heuristics and Scoring** - Emails clearing the first steps are tested against a set or sets of rules and scored accordingly. These rules can be customized for the Small and Medium Business customer to serve their specific business needs.
4. **Client Feedback** - St. Bernard On Demand is constantly in contact with their customers to see if any false positives or false negatives are occurring within the system. All of this feedback is used to update the process to improve its effectiveness where possible.
5. **Volume Aggregation** - All of the email that is identified as spam during the St. Bernard On Demand spam filtering process is logged into the database of known spam for filtering of subsequent emails.

Because there is a small chance that false negatives could occur (less than a 2% chance), administrators might want to review the actual email that is captured by the St. Bernard On Demand process. The captured email is handled in one of four ways:

1. The first, and most popular, option is to leave the junk e-mail on the St. Bernard network. As we catch the junk e-mail, we create a temporary junk e-mail box for each of your users. We then notify each user that they have junk e-mail on our server and give them a link to a web-based control panel where they can check it. As the administrator, you can

set the notification interval. We hold the junk e-mail on our network for 30 days before it is deleted.
2. The second option is to redirect all the junk mail to a single e-mail address. For example, spam@yourcompany.com.
3. The third option is to have St. Bernard On Demand insert an x-header. The junk e-mail is delivered to your mail server giving you the ability to reroute the e-mail via the x-header variable.
4. The fourth option is the ability to modify the subject. For example, you could insert the word JUNK before the subject so that the individual users could set up their own filters at the e-mail client level.

Through flexible, effective technology, service and understanding, St. Bernard On-Demand Email Filter has implemented a process that ultimately delivers spam-free email and email protection to the recipients within your organization. To learn more, call St. Bernard at (800)-782-3762 or visit www.stbernard.com.

Summary

Throughout the discussion, we have focused on the methods for solving the problems that spam creates. Understanding the philosophies, methods and technologies behind spam filtering will allow organizations to make the most informed choice when selecting a vendor to protect their organization and email recipients from the hazards email might contain if left unchecked. The St. Bernard On-Demand Email Filter process was then described and was shown to include a number of the current technologies explained in the discussion, and further value was related in the process approach that is used by St. Bernard On Demand. Because St. Bernard offers their email protection as a service, any organization can benefit from their highly effective process. Any organizations desiring more information are encouraged to contact St. Bernard at the phone number or website provided above.

About St. Bernard

St. Bernard Software is a global provider of security solutions, including Internet and email filtering appliances, patch management and data backup solutions. St. Bernard Software also provides the SME market with a complete line of hosted security solutions, including email, IM and URL filtering services. Deployed across millions of computers worldwide, the company's award-winning products deliver innovative security solutions that offer the best combination of ease-of-use, performance and value. Established in 1995 with headquarters in San Diego, CA and an international office in the United Kingdom, St. Bernard Software sells and supports its products directly and through solution partners worldwide. For more information, please visit http://www.stbernard.com./

About the Author

Lawrence Didsbury, MCSE, MASE, is an independent writer focused on messaging, data storage and other computing technologies. Lawrence was most recently a Marketing Manager for Auspex Systems after serving at Compaq Computer Corporation as an engineering team lead, and a systems/software engineer. Before that Lawrence was an independent network/Internet consultant and served with systems integrators as a network services manager and systems engineer. Lawrence is a Microsoft Certified Systems Engineer (MCSE) with a Microsoft Exchange specialization, and is a Compaq Master Accredited Systems Engineer (MASE) with a concentration in Enterprise Systems Management. Lawrence has a BS degree from the University of Houston and is currently pursuing an E-Commerce MBA from Jones International University.