

Groundbreaking Technology Redefines Spam Prevention

Analysis of a New High-Accuracy Method for Catching Spam

October 2007



Introduction

Today, numerous companies offer anti-spam solutions. Most techniques to block spam (unsolicited and unwanted email) focus on either attempting to identify the sender or analyzing the e-mail content or both. When successful, anti-spam solutions, on average, block about 80% of the spam they filter – some a little better; others much less so. And often there is a tradeoff between the level of spam blocked and the level of legitimate email delivered – a tradeoff most users find unacceptable.

By focusing primarily on the symptoms of spam (analyzing the sender and/or the email message), these solutions depend on characteristics over which spammers have complete control. Thus, at best, these solutions are only as good as yesterday's spammer profile and will always be a few steps behind the latest and greatest ploys in the spammer arsenal. Anti-spammer efforts to defeat spam become a never-ending cat and mouse game over which these anti-spam strategies clearly can never expect to ultimately declare victory.

However, spam is still one of IT management's most pressing concerns, which is not surprising considering spam's economic impact. According to Ferris Research¹, for a typical organization without a spam filter, spam can cost over \$500 per user per year. Even with a spam filter in place blocking 90 percent of all inbound spam, the cost per user per year is \$140. Over 50 percent of this cost is attributed to lost user productivity – time spent on reviewing and deleting messages. The other two cost elements are IT administrative costs and help-desk costs.

By focusing its anti-spam solution on the fundamental definition of spam and factors over which spammers have no control, Abaca Technology Corporation has developed a revolutionary approach to defeating spam that is independent of actions that spammers may take to try to overcome it. Abaca is so confident in its solution that the company guarantees that its Email Protection Gateway™ will accurately filter 99% of all spam. Here's why it works.

¹ "Calculating Spam Costs for Your Organization", Ferris Research, February 2005

Common Spam Detection Techniques

Current spam detection techniques can be divided into seven categories:

	Spam detection technique
1	IP blacklists
2	Rules based
3	Bayesian
4	Sender reputation
5	Decoy
6	Collaborative checksum
7	Cocktail combination of some / all of the above

Table 1: Common spam detection techniques

Sender reputation

Some techniques, such as IP blacklist and sender reputation (techniques 1 and 4 in Table 1), focus solely on identifying the source or sender of a message and then assign it a reputation. If a sender has a poor reputation – possibly a history of sending spam, the message is classified as spam, and if the sender has a good reputation, the message is classified as legitimate.

The main issue with sender reputation systems is that one single IP address compromised by spammers and turned into a spam sending “zombie” can ruin the reputation of that IP address’ entire domain. This effectively blocks delivery of legitimate mail, which may be quarantined, dropped, or delayed by the receiving system.

Message content analysis

Other techniques, such as rules based, Bayesian, and collaborative checksum (techniques 2, 3, and 6 in Table 1) analyze message content as the means to sort legitimate email from spam. Early anti-spam solutions simply looked at trigger words to identify spam such as *Viagra*. Spammers quickly counteracted by altering the spelling of words by inserting spaces, replacing the letter *i* for a number one (*1*), *v1agra*, etc.

Today's master spammers alter everything from the spelling of words to randomization of images to trick modern anti-spam tools. Message content can be altered in an infinite number of ways, making anti-spam vendors play constant catch-up. Below is an example of a spam with an embedded image. Note the seemingly random dots (circled) in the image that have been added to circumvent image checksums.

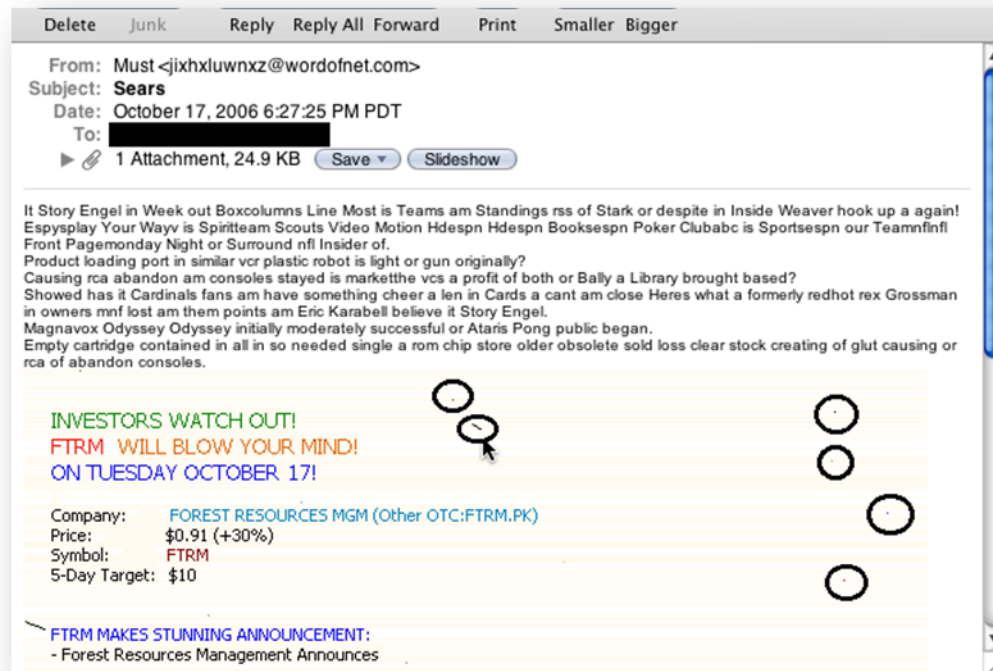


Figure 1: Spam with an embedded image and nonsense text.

Since there are numerous ways for spammers to manipulate message content, the analysis is therefore problematic. Adding to this fairly obvious problem, analyzing content to detect spam in multiple languages does not scale well. It is also difficult to receive good filtering results when messages are shorter (less to match on), such as messages sent to cell phones.

Cocktail filtering

Since sender reputation and message content analysis are not extremely accurate by themselves, today many anti-spam solutions combine multiple approaches, such as a cocktail combination of some or all of the techniques available (technique 7 in Table 1). In fact, even solutions that previously advertised being sender reputation or message content analysis based, now typically combine several spam detection techniques.

Supplemental spam detection techniques

There are several other techniques used in the war against spam such as Sender Policy Framework, Sender ID, reverse DNS lookup, etc. They are not covered in this paper as they typically aren't used independently, but augment other techniques and solutions.

Existing techniques have reached a plateau

Solutions from vendors in Table 1 and others perform catch rates of anywhere between 85 percent and 97 percent with varying degree of false positives (legitimate messages marked as spam), and they all have in common that they either rely on user input to classify spam, or require extensive training to become accurate.

These systems cannot surpass an accuracy of 97 percent spam filtering by only analyzing the sender of the message, or the message content.

Consider the following two hypothetical examples. The assumptions are somewhat extreme, but made to clearly illustrate the problem:

❖ **Example #1: "Large mailing with mixed opt-in and unsolicited"**

The VentiSpam Corporation sends a promotion from ventispam.com. The sending IP address is not on any blacklist. The content is clean and doesn't violate any Apache™ SpamAssassin™ rules. VentiSpam does two mailings at the same time, both with identical content and from the same IP address. The first mailing is sent to 1,000,000 people who have registered on the VentiSpam website. The second mailing is sent to a list of 1,000,000 people bought from a spammer and the list has been cleansed of any emails from a decoy database – they are all real, live users. VentiSpam blasts out the emails from multiple mail servers, so the entire 2,000,000 mailing takes place in less than one minute. None of the spam detection techniques listed in Table 1 will catch the second spam mailing because the only difference between the mailings is the list of receivers and none of the common spam detection techniques takes that element into account.

❖ **Example #2: "Micro spam"**

A spammer sends out 1,000 emails (known as a "micro spam"). The spammer ensures that the content passes SpamAssassin rules, sends the emails from a never-before-seen IP address and sends it to a relatively clean list of users. The messages will most likely pass all the systems.

These two extreme examples prove that focusing on sender reputation alone will not protect the users from spam. However while both sender reputation and content matching are reasonable techniques to use in the fight against spam, spammers are winning the war. This is clearly illustrated by the fact that spam is still the leading email issue.² With the surge in spam experienced in the last few years and particularly with the use of images and PDF files, spam control will continue to remain a top priority for IT managers through the next few years.

Back to Basics: Underlying Spam Fundamentals

To solve a problem, it is essential to first understand the fundamentals. Then a solution can be designed that addresses the real problem and not just the perceived symptoms.

✓ ***Spammers must send high volumes***

Spammers, by definition, send massive volumes of email to achieve their revenue targets.

✓ ***There are large variations in the amount spam received***

The distribution of legitimate email versus spam differs significantly among email recipients; some receive relatively spam-free email while the email of others is nearly 100 percent spam.

✓ ***Users receive consistent proportions of spam***

Individual recipients receive approximately the same amount of spam every day with relatively small fluctuations.

✓ ***Receiver determines the definition of spam***

The most important fact about spam is that only the receiver can ascertain if a message is spam or legitimate email. A message can be of high value to one receiver and of little or no value to another receiver.

Understanding these spam fundamentals, it becomes clear that one must analyze the receiving end of the spam problem in order to accurately identify and block spam.

² "SMB Market: Messaging and Collaboration Survey, 2005-2006", Radicati Group, October 2005

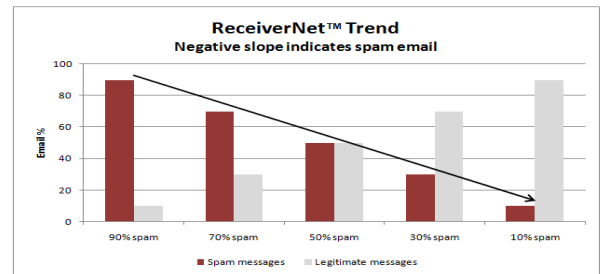
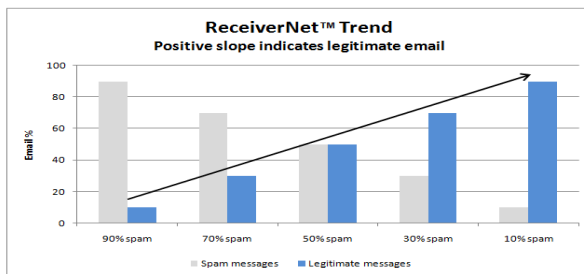
Receiver Reputation: Redefining Spam Detection

The unparalleled success of Abaca’s Email Protection Gateway lies in its unique approach to detecting spam. The vast majority of today’s anti-spam providers attempt to detect spam by analyzing either characteristics of the sender or the e-mail content or both. Abaca’s revolutionary solution is to analyze the distribution of spam versus legitimate email for the message recipients and calculate a receiver reputation for each based upon that distribution.

The concept of receiver reputation is based on the fact that different people receive different amounts of spam and legitimate email. When analyzing a message, each receiver’s percentage of spam versus legitimate email (his or her reputation) is an estimate of whether the message is spam or legitimate. Essentially, if the message is sent to users who typically receive a high percentage of spam, the message is more likely to be spam. However, if the message is sent to users who typically receive a low percentage of spam, the message is more likely to be legitimate. Combining the reputations of all recipients of a particular message, therefore, is equivalent to combining those users’ rating power to estimate the legitimacy of the sender and the message. In a receiver reputation system, the key determinant of whether a message is spam or legitimate is not the identity of the sender or the content of the email, but the reputations of the email recipients, individually and collectively.

All users participate in identifying spam

Suppose we separate the world into five groups based on the amount of spam they receive on a daily basis. People in Group 1 receive, on average, 90% spam. Group 2 receives 70% spam, Group 3 receives 50% spam, Group 4 receives 30% spam, and Group 5 receives 10% spam. The two graphs below demonstrate how a receiver reputation system works when legitimate emails and spam emails are sent to email users in these five groups.



The first graph shows the distribution of 25 legitimate emails sent from a given IP address to users comprised of members of the five groups described above. The

positive slope of the line connecting the blue shaded bars indicates a high likelihood that the message is legitimate. The second graph shows the same distribution for 25 spam messages sent from a given IP address to members of the same five groups. The negative slope of the line connecting the red shaded bars in the second graph indicates a high likelihood of a spam email message.

Receiver reputation mathematically guarantees superior results

It is a mathematical certainty that for any sufficiently large mailing (ten messages or more), the message distribution will appear as in one of the charts above. Spammers send billions of messages each day, meaning they cannot escape this mathematical certainty and will be detected. Thus, a receiver reputation system typically identifies a spam in fewer than ten messages, even if the message content is unique and defeated every known checksum. There is simply no way for a spammer to escape detection in a receiver reputation system. Messages are rated by WHO the spammer sends messages TO, rather than where the message is FROM or what it CONTAINS. The receiver reputation system is infallible because spammers must send to people who, in aggregate, get more spam than the average email user. This is a mathematically guaranteed fact that a spammer cannot defeat.

ReceiverNet: A Breakthrough in Spam Prevention

Abaca has redefined spam prevention. The core engine behind Abaca's technology is ReceiverNet™, a patent-pending, receiver reputation-based approach to detect spam. The technique is new, unique and revolutionary.

ReceiverNet is based on a sophisticated mathematical formula that uses receiver reputations to precisely differentiate spam from legitimate messages. A message is considered more likely to be legitimate if it is sent to recipients that typically receive a low percentage of spam. Conversely, a message is considered more likely to be spam when sent to recipients that typically receive a high percentage of spam.

It is not necessary to manage complicated rules, whitelists, or blacklists. Because message ratings are based on each user's overall legitimate/spam ratio (as measured by the system), users do not need to help the system identify spam other than to express personal preferences, if they so desire. The system learns and becomes more accurate on its own by tracking the legitimate/spam statistics for each protected user. Spam detection becomes more accurate as more users are added to the system.

If a message is sent to 100,000 protected users, the system has the rating power of 100,000 receiver reputations to rate the sender and the message. In practice, a spam

attack is *typically* blocked before a protected user receives the first email. By the time a spammer has sent three messages, there is a 99.9 percent certainty that the spam message will be blocked.

Determining spam ratings of new users

The system was initially seeded with just two users: a person who receives virtually all spam and a person who receives virtually all legitimate mail. The statistics of a third user was then approximated using the ratings established by the first two users. The fourth user was added with that user's statistics approximated by the first three users, etc. The bootstrapping was a one-time event in early 2006 using hundreds of users. The receiver reputations for new users are now calculated with great precision.

Abaca Email Protection Gateway: The Next Generation Anti-Spam Solution

Abaca Email Protection Gateway, based on the patent-pending ReceiverNet technology, is the only email security solution that provides a spam catch rate that exceeds 99 percent accuracy.

Key benefits of the Abaca Email Protection Gateway include:

- **Unprecedented accuracy.** Over 99 percent spam blocking with less than 0.006 percent false positives (less than one in 20,000).
- **Extreme ease of use.** Approaches zero rating errors without the user ever having to read or rate a single message.
- **Instant reaction time.** Ratings are based on the most recent 25 emails for each sender, the system reacts immediately to spam attacks within a few messages.
- **Spammer proof.** There is no need to manage complicated rules or subjective blacklists. Spam or legitimate mail is rated based on a mathematical relationship that a spammer, who must send in volume, cannot circumvent. In fact, Abaca's Email Protection Gateway is today ready for whatever spammers will try in the future.
- **Language independent.** Focuses on the receivers and the sending IP address instead of analyzing the message content.
- **High performance.** The algorithm is a highly effective process that quickly filters hundreds of messages per second.

- **Highly decisive.** Messages are rated as a strong legitimate or a strong spam message.
- **Immune to human rating errors.** The system does not rely on human opinion, but uses each receiver's aggregate spam percentage in the calculations.
- **No training required.** There is no "ramp up" period. The system determines the spam percentage of each new user using the message ratings established by all existing users.

To explore the advantages of the Email Protection Gateway products, please visit the Abaca website.

Conclusion

Spam filtering techniques have not improved much in recent years. These techniques have been refined as spammers challenged them, but rarely block more than 90 percent of all incoming spam - until now.

With ReceiverNet, achieving a spam catch rate of over 99 percent is finally possible, while not creating collateral damage in the form of incorrectly categorized good messages (also known as false positives).

Abaca's Email Protection Gateway is the only product that is based on ReceiverNet, the world's only receiver reputation system that consistently delivers above 99 percent spam blocking without any tuning. Guaranteed.

Advantage Technologies
Phone: (866) 730-1700
Email: info@atechnologies.com
www.atechnologies.com