The Radicati Group, Inc.
1900 Embarcadero Rd., Suite 206
Palo Alto, CA 94303
Phone 650-322-8059
Fax 650-322-8061
http://www.radicati.com

The Radicati Group, Inc.

An Analyst Review of Hotmail Anti-Spam Technology

A White Paper

www.radicati.com

TABLE OF CONTENTS

This white paper was sponsored in full by Microsoft Corporation.

1.0 INTRODUCTION

Spam has rapidly turned from a mere nuisance into a major security threat and financial drain for all email providers, as they attempt to stem the flood of unsolicited email while ensuring that legitimate messages are delivered correctly to the end user. To make matters worse, spam is often a mechanism used to carry viruses, malware, and numerous other security threats which can compromise a user's computer, gain access to sensitive personal information, and cause significant cost in terms of repairs to infected systems. Finally, there is the challenge of successfully blocking spam while at the same time avoiding the accidental deletion of legitimate user email.

While at first spam was relatively easy to block through the use of blacklists, or basic content filtering techniques, spamming methods have advanced to the point that these technologies are no longer sufficient or cost-effective. Anti-spam methods, therefore, have also had to evolve into a complex set of technologies that can be effective in this adversarial environment.

It is important to note, however, that not all unsolicited email necessarily represents spam. In particular in the case of commercial advertising, there is often a fine line between legitimate advertising and actual spam. This is referred to as graymail. Graymail poses a particular challenge for email providers as they need to provide easy ways for users to report un-wanted graymail messages, while at the same time containing the number of false positives (i.e. legitimate emails erroneously categorized as spam).

This white paper looks at the overall spam problem today: how much spam is received by email providers and successfully filtered, what percent spam still gets delivered to the end user and what how technologies have evolved over time to fight spam.

In particular, the paper looks at the efforts Hotmail is making to combat spam. As one of the largest email provider networks today, Hotmail presents a particularly large target for spammers. Yet through the use of the latest anti-spam technologies and a sophisticated layered approach to combating spam, Hotmail is able to deliver one of the cleanest inbox experiences to the end user. Hotmail's ability to fight spam, today, is very much on par or better than accepted industry averages.

2.0 UNDERSTANDING THE SPAM PROBLEM

Unsolicited bulk emails, or spam, first emerged in the mid 1990s as the Internet grew in popularity. Early spam consisted of simple emails attempting to sell something or conveying unsolicited information (e.g. often porn or chain letters). Over time, however, spam has evolved into much more complex forms of unsolicited messages. Often these emails are of a malicious nature and are sent for profit or to obtain sensitive information from the user as a way of then using this information for fraudulent purposes.

2.1 Types of Spam

Today, we differentiate among the following types of spam:

- **Commercial advertising** is the most common form of spam. These messages are a key nuisance to both businesses and consumers.

- **Illicit advertising** includes explicit advertisements, advertisements for illegal or counterfeit goods, pump-and-dump stock schemes and more.

- **Phishing** messages are designed to extract personal information from end users, including credit card numbers, user ids/passwords, social security numbers, and more. This is a quickly growing type of spam with significant profit potential for the spammers.

- **E-mail Fraud** is a quickly growing type of spam that can be particularly dangerous for end users. These messages include money transfer schemes, advertisements for non-existent products/services, advertisements for fake charities, and more, that are designed to directly bilk users out of money.

- **Chain Letters** quickly spread throughout communities of e-mail users. Some of these messages request money to be sent to other individuals, while other types of chain letters simply encourage users to forward the message to others.

Table 1, shows the different types of spam received in 2009.

Types of Spam	
Commercial Advertising	67%
Illicit Advertising	12%
Phishing	11%
Email Fraud	7%
Chain Letters	1%
Other	2%
Total	**100%**

Table 1: Types of Spam Received, 2009

2.2 Spam vs. Graymail

It is important to note, however, that not all unsolicited email is necessarily spam. In particular in the case of commercial advertising there is often a fine line between legitimate advertising and actual spam. We refer to this as:

Graymail - Graymail may include newsletters, social networking emails, and various types of alerts. Often the user has signed up for these and does not recall doing so, or may have inadvertently triggered the receipt of such email from websites they may have visited. Sometimes it's simply too much information at the wrong time from what would normally be a welcome sender (e.g. airline frequent flyer groups).

So in fighting spam it is important to understand that we are actually dealing with three possible types of email:

a. **Legitimate email** – which should not be accidentally categorized as spam and deleted.

b. **Spam** – which is totally unsolicited email, which should be eliminated before it reaches the user inbox.

c. **Graymail** – which may be legitimate with respect to spam filters (after all the user may at one time have signed up for it), however from the user perspective at any given point in time it may represent "spam" if too much is received, or if it is no longer of any interest.

Needless to say, Graymail is by far the most difficult one to control as it is highly subjective and cannot be done without the active participation of the user who needs to have an easy, user-friendly mechanism they can use to signal that receipt of a certain type of email is no longer desirable.

2.2 Spam Volumes

Spam is a very big problem for all email providers, with literally billions of spam emails that threaten the security of user inboxes on a daily basis. All email providers are deluged with billions of spam messages on a daily basis and all must fight to stop as much spam as they can before it reaches their subscriber's inboxes.

Table 2, below, shows the extent to which daily email traffic in the Internet is saturated with spam, and how the volume of spam is projected to grow from 2009 to 2014. In 2009, the Radicati Group estimated spam comprised 81% of the total worldwide email traffic, and this number is expected to increase to 85% by 2014. This is the total spam traveling on the Internet, however, the majority of this spam never reaches its intended destination as it is successfully blocked by anti-spam filters.

	2009	**2010**	**2011**	**2012**	**2013**	**2014**
Worldwide Message/Day (B)	247	294	349	419	507	613
Worldwide Spam Traffic/Day (B)	**199**	**238**	**286**	**347**	**424**	**517**
Total Spam %	*81%*	*81%*	*82%*	*83%*	*84%*	*85%*

Table 2: Worldwide Spam Traffic, 2009 - 2014[1]

Figure 1, below, shows what percentage of spam on a daily basis gets through spam filters and is delivered to the user's inbox (*note: this includes both actual spam and graymail*).

[1] Email Statistics Report, 2009-2013 – May 2009, The Radicati Group, Inc.

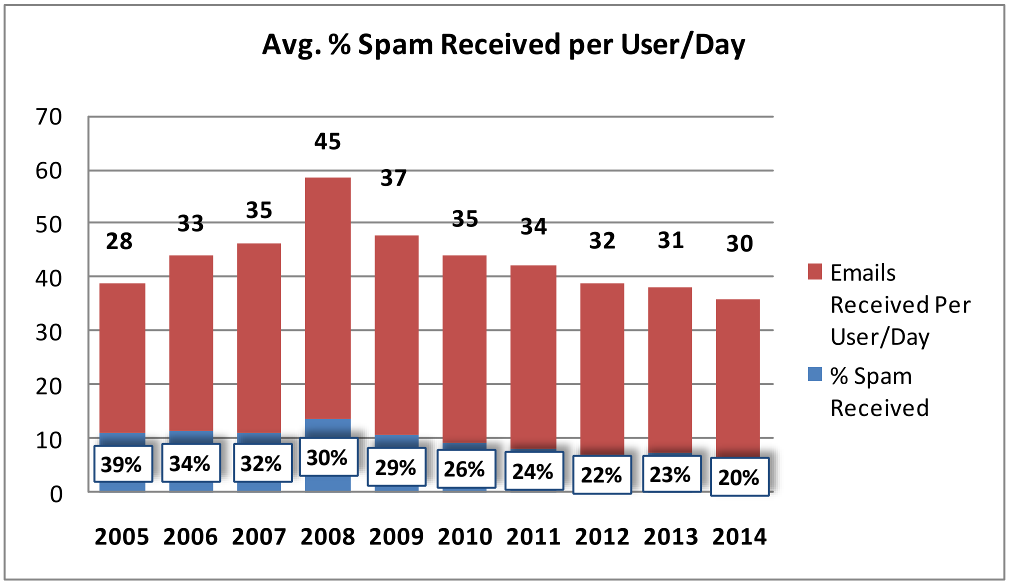

Figure 1 – Average % Spam Received Per User/Day, 2005-2014[2]

This data is based on input from users across all types of email networks and therefore serves as a baseline for understanding the typical level of spam penetration at any given time. In 2009, users reported that they considered an average 29% of emails they received on a daily basis to be spam (i.e. spam and graymail). Note that figure 1 also shows that despite the number of emails and spam received has trended upward the % of spam received has been decreasing. This is largely due to improved spam filtering techniques as well as greater synergies across the industry to fight spam.

3.0 THE EVOLUTION OF ANTI-SPAM TECHNOLOGIES

Naturally as spammers continue to evolve their techniques to bypass filters, vendors respond with more innovative approaches to defeat these spamming techniques. The following section provides an overview of how spamming techniques and anti-spam technologies have progressed over time.

[2] These figures are based on aggregate data from a number of user surveys we conduct on an annual basis where users report what percentage of their email received on a daily basis is spam.

Whitelists and Blacklists - are a straightforward method of blocking spam using lists of email addresses, IP addresses, and/or domains that are considered safe (whitelists) or unsafe (blacklists) to determine whether to accept or block messages from a sender. There are several public blacklists that are utilized by email security solutions today, such as Spamhaus, SORBS, DSBL, and many others.

In the early days of email, email providers could rely more heavily on blacklists and whitelists, as email was not used as extensively for communication as it is today. It became more impractical to rely on these techniques, however, as email traffic continued to increase. In addition, these lists depend largely on recipient submissions. However, recipients often mistakenly or inappropriately place IP or sender addresses on blacklists.

The growing usage of zombies and botnets has also made blacklists much less effective in blocking email. Zombies are computers that, unbeknownst to their owners, have been infected with malware that forces them to send spam emails. Botnets are a collection of multiple zombie computers that are controlled by a single spammer source, giving the spammer the ability to send mass emails with multiple machines. With this technique, otherwise legitimate email senders may be placed on blacklists because of their infected computer. Because blacklists are static, it can be difficult for users of infected computers to get their names removed from blacklists even if they have cleaned their computers and are once again only sending legitimate emails.

Content Filtering – was created as a complementary anti-spam technology to blacklists. Blacklists attempt to block spam from known spam senders while content filtering was introduced to prevent spam delivery from unknown spammers by inspecting the email body. This simple technique works by scanning the message body and subject line for keywords that flag a message as spam, such as words related to pornography, gambling, etc. This was one of the early anti-spam methods and is relatively easy to implement. It was initially effective because spammers were fairly straightforward in their messages, in that they did not try to bypass anti-spam filters. Content filters were able to easily detect spam keywords in the bodies of messages. This technique, however, was susceptible to false positives, because many of the spam keywords also had legitimate uses. Furthermore, spammers began using methods that could easily bypass these filters through the use of techniques that would obscure content, such as misspelling, spacing, symbols to replace letters, or more recently, images in messages.

Context Filtering – is similar to content filtering (and often labeled as such), and relies on scanning messages for specific word groups within a defined context, making it somewhat more effective than simple content filtering. This method is particularly helpful when differentiating the use of terms in a specific industry context versus spam. For example, the term Viagra can have legitimate uses in the healthcare industry but is also a strong indicator of spam. Context filtering can help distinguish among these different circumstances.

Just as with content filtering, context filtering is relatively straightforward to implement. Unfortunately, it shares the same drawbacks as content filtering in that it can be easily bypassed by more advanced types of spam.

Signature Filtering – uses a known spam email to block all identical spam messages. Initially spammers would send out the same spam email en masse. They could simply create one spam email and send it out to an extensive email list. Signature filtering was designed to combat this spam approach by capturing "signatures" of known spam emails. If an incoming email matches the signature of a known spam message the email is blocked immediately. This anti-spam technology is a good complementary technique in that it will block obvious spam messages with very low false-positive rates. However, spammers now bypass this approach by using templates to randomize spam emails. Small variations in each spam message make each email unique. Signature-based filtering will not work if spam emails are slightly different.

Heuristic Filtering – is a rule-based approach that looks for spam indicators in emails. Unlike simple content and context filters, heuristics filters can be coded to be quite sophisticated and complex. This approach was developed to detect the tricks spammers began using to try to fool anti-spam filters, with rules to detect several devious spamming techniques such as "transparent" font colors, tiny font, deceptive URLs, and more.

Heuristic filters are only effective when their rules are well-written and kept up to date. They require constant maintenance to remain effective, as the rules are static and spammers are continually finding ways to bypass these rules. Poorly written heuristic filters also have the tendency to have a high false-positive rate, leading to legitimate emails being misidentified as spam.

Statistical Filtering – is an advanced anti-spam technique that uses statistics to determine whether an email is spam. One approach to statistical filtering is to create a "score" for the email based on the number and weight of the spam indicators identified in the email. The email is determined to be either a spam or good email based on the score threshold. Often users can modify the anti-spam sensitivity by adjusting this threshold.

Other statistical filtering methods can "train" the filter using samples of spam and non-spam emails to determine the statistical probability that incoming emails are spam messages. This technique goes beyond the static rules of heuristics and uses a learning-based approach. The more email that is processed, the more effective the filter is.

The major downside of statistical filters, however, is that they must be well-tuned or well-trained in order to keep up with newer spamming techniques. The spam indicators must represent current spam trends and appropriate weights and score thresholds must be assigned. If using a training method, the email training sets must be updated frequently with large, representative email samples or the statistical filtering will be ineffective.

Even if statistical filters are maintained correctly, many spammers have developed methods which make it difficult to analyze the content of emails. For example, image spam displays the spam message in an image and not in text in the email body, making it more difficult to identify the spam indicators. Image spam has been successful in circumventing most advanced filters, and requires more innovative solutions than just the ones listed above.

In addition, most of the anti-spam methods discussed above require emails to enter the network to be scanned. With the current inundation of spam, scanning these emails within the network can require costly resources and may overload the system. In today's spam environment, anti-spam solutions must be able to stop the bulk of email threats before they even reach the network.

Reputation Services - In a nutshell, reputation services maintain an accurate knowledge about the nature of the email sender or embedded URL. If the sender is known to send spam or other malicious email or the link in the email connects to a malicious Web site, a bad reputation is assigned, ensuring that only emails from good sources are accepted.

Reputation services work by leveraging a large databases of sender and receivers to monitor messaging traffic. Unlike simple blacklists, reputation services continuously analyze the sending behavior of IP addresses and domains to determine if they are sending legitimate or illegitimate email. By identifying the sources of spam and other email threats, reputation services can block email based on the sender without having to scan the content of the email, increasing effectiveness, lowering false positives, and reducing the burden on the email network.

In addition, reputation services can be applied to the URLs embedded in emails. If a link connects the user to a spam site or other malicious Web site, the URL is given a “bad” reputation. An email containing a URL with a bad reputation is automatically blocked. This keeps that email out of the inbox and prevents users from following the link and falling victim to Web threats, such as malware downloads, phishing sites, and more.

Effective reputation services collect an email history and email samples from sending IP addresses or data on Web sites. Reputation services are able to support why they have given an IP address or URL a “bad” reputation. Reputation services continually update their lists, to weed out zombies and botnets which are responsible for most spam on the Internet. As victimized computers remove the bot code and once again send good email, they are no longer blocked as having a bad reputation.

The main benefit of reputation services is that they keep most spam and other email threats completely off the email network. They do not scan the body of the email, so they are not subject to the content tricks used by spammers. By keeping threats from even entering the network, reputation services ensure the email network stays secure and the available bandwidth is maintained for legitimate email.

4.0 FACTORS AFFECTING THE AMOUNT OF SPAM RECEIVED

As we look at spam rates across different email providers, we must keep in mind several factors that add complexity to email provider’s anti-spam efforts:

- **The size of the email provider’s network** – The larger the email provider’s user base the more massive the amount of spam it will receive and will need to deal with on a daily basis. So an email network such as Hotmail, which counts 365 million users will get about two times more spam directed at it than say Google Gmail, which only counts about 175 million users.[3] (i.e. two times the size of the user base means two times the amount of spam directed at the email provider).

- **The age of accounts on the email provider's network** – Typically the longer a user has had and used an email account on the Internet the wider the "trail" that account has left behind in terms of signups to various Internet sites, access logins and all forms of online registration which can be picked up by spammers and in time will result in more spam being directed at that email account. As Hotmail was one of earliest large-scale email providers on the Internet the average age of its email accounts is typically significantly greater than that of many newer email providers. Again this translates into Hotmail having to fend off much more spam on a daily basis than more recent email providers. This becomes particularly an issue when dealing with graymail as users with older accounts will typically have signed up for more newsletters, reminders, alerts or notifications than they remember and often do not wish to receive any longer. Distinguishing between these annoying graymail messages and true spam that needs to be blocked is a real challenge for the email provider.

- **Nature of usage of the account** – This is closely related to the age of the account, since earlier Internet users tended to be less careful about supplying their email addresses and somewhat more naïve and trusting about their overall online experience. As the spam problem and related phishing, malware and fraud problems have increased and gained more publicity, users habits have evolved and now create less exposure to attacks than in the past. Nevertheless, for many older accounts the damage has been done and they tend to continue to be bombarded by more unwanted emails. Figure 2, below, summarizes the results of surveys carried out in 2005 and 2009 that shows how user habits have evolved over time[4]. It shows that when asked if they ever clicked on a link contained in an unsolicited email (except unsubscribe), 42% of respondents said that "yes they had" in 2005, while only 13% indicated they had in 2009. Obviously, clicking on a link in an unsolicited email will typically result in more spam for that email account. So clearly user awareness and education has improved tremendously from 2005 to today.

[3] Comscore – January 2010.
[4] User Surveys carried out by The Radicati Group from 2005 to 2009.

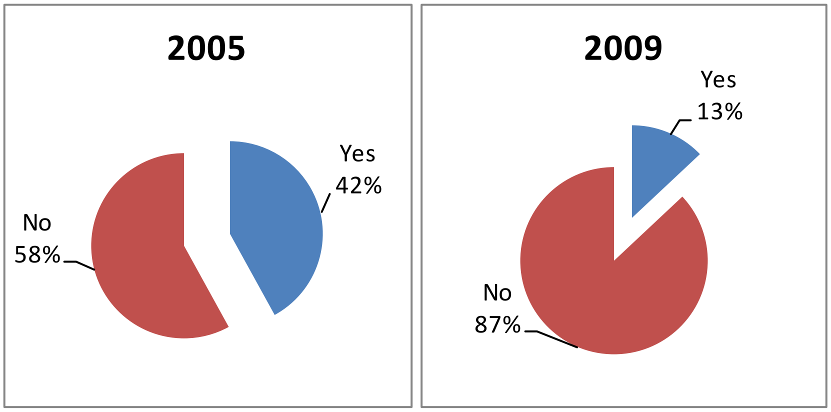

Figure 2 – User Behavior Evolution, 2005 & 2009

4.1 How well does Hotmail deal with Spam?

In evaluating Hotmail's anti-spam capabilities it is important to keep in mind a few key statistics[5]:

- Hotmail receives an estimated 8 billion messages daily. The majority of these are spam (5.5 billion) that gets filtered out before the messages are delivered.

- After all the filters, the final number of spam messages delivered to users represents 1.4% (i.e. less than 2%) of the original number of messages received by Hotmail.

- The number of legitimate messages delivered daily is around 2.4 billion.

These numbers show that Hotmail deals very effectively with the deluge of spam that targets its users on a daily basis. Through the use of highly sophisticated filtering technologies, Hotmail is able to reduce the amount of spam that gets delivered to the inbox to a very small fraction of what it receives. Hotmail's ability to fight spam is very much on par or better than accepted industry averages.

[5] Note: all the statistics presented in this section have been provided by Microsoft.

Figure 3, below, shows the type of message throughput that is processed by Hotmail on a daily basis.

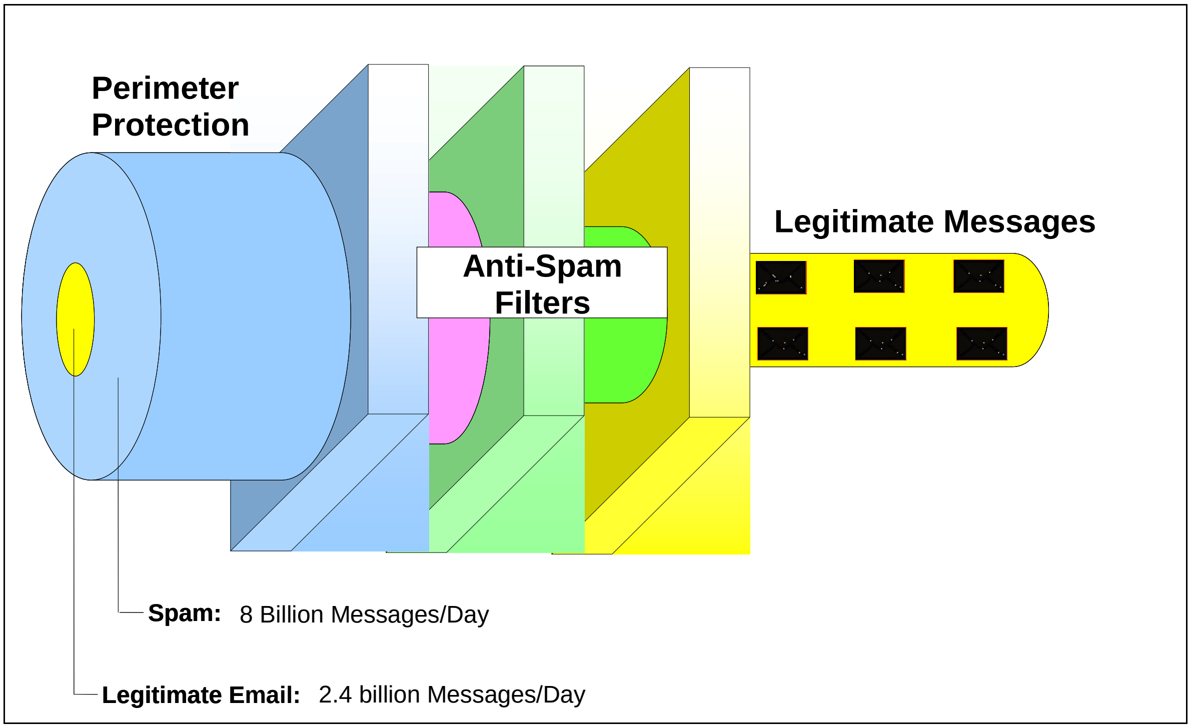
Perimeter Protection
Legitimate Messages
Anti-Spam Filters
Spam: 8 Billion Messages/Day
Legitimate Email: 2.4 billion Messages/Day

Figure 3 –Hotmail Spam Processing Throughput

5.0 Hotmail Anti- Spam Technology

As one of the early major network email providers, Hotmail has had a great deal of experience with fighting spam. Today, its anti-spam approach may be thought of as comprising three distinct layers as shown in figure 4 below, which support and reinforce each other to provide a safe user experience:

a. **Perimeter Technology** – meant to identify and delete most spam before it even reaches and enters the network.

b. **Spam Filters** – to weed out unwanted email at a more granular level.

c. **User Controls** – to allow users to easily and effectively signal when a sender or type of message are no longer wanted.

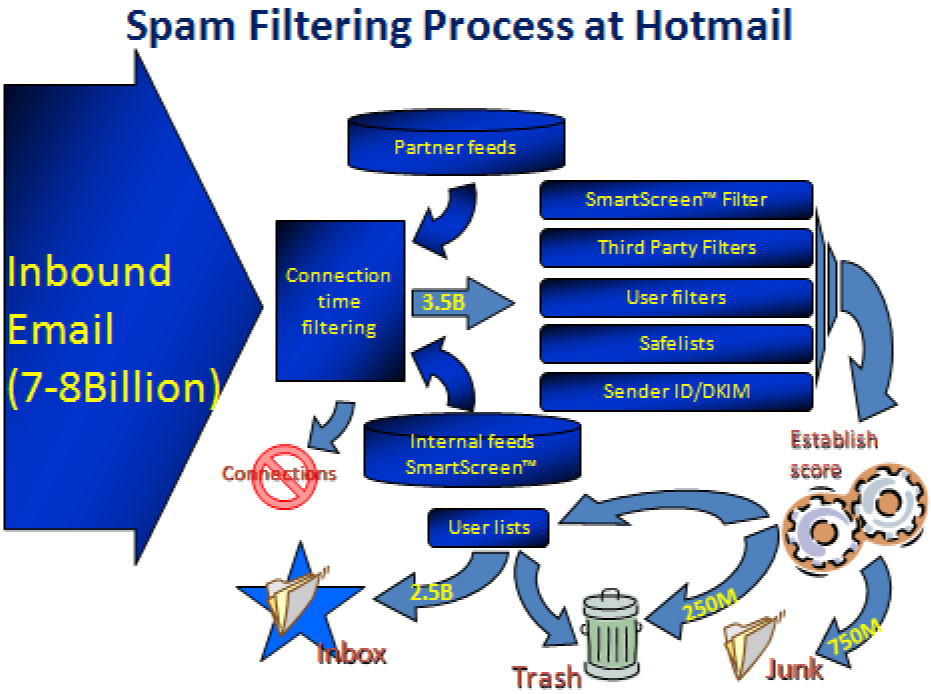

Figure 4 –Hotmail's Spam Filtering Process.

We will now look at each of these layers in more detail:

i. **Connection-time Filtering** – This is Hotmail's first line of defense. It is basically a reputation-based approach, where at any point in time, the Hotmail anti-spam system has knowledge of the world of e-mail senders based on various sources of reputation as well as recent trends in e-mail content. Sender reputation is generally tied to IP addresses or ranges of addresses. This information is then used to set limits on how many messages any given sender can deliver to Hotmail. Setting the limit to zero effectively blocks all e-mail from that sender. For good senders the limit is set to allow normal email delivery while minimizing the potential for abuse should the senders' own computers get hacked. Hotmail uses several sources to generate sender reputation:

- *IPs of Bots* - Hotmail tracks individual machines that are being used to send spam. Often, these machines are PCs that have been infected with malware and taken over by spammers.

- *Dynamic IPs* – it is a known fact that computers with dynamically assigned IP addresses should not be sending e-mail, so Hotmail tracks those and rejects email coming from those computers.

- *Known spam entities* – like all reputation-based services, Hotmail also makes use of external information, like IP address registration and Autonomous System Numbers (ASN) to identify ranges of IP address that are known to be used for spam and other malicious purposes.

- *3rd-party sources* – Hotmail also contracts with several 3rd-party reputation specialists to leverage the very best the industry has to offer.

ii. **Content Filters** – the emails that pass the reputation-based Connection-time Filtering and are accepted into the Hotmail network are then processed through a set of content filters. Microsoft's SmartScreen technology is applied to analyze the content of incoming emails. SmartScreen is a patented technology based on a machine learning approach where decisions regarding what emails are spam are made by the email user and incorporated into a feedback loop to train the filter to know what to look for. SmartScreen was originally developed by Microsoft Research in 2003, early versions were included in Outlook 2003, and applied across Microsoft's email platforms such as Exchange Server 2003 as part of Microsoft Exchange Intelligent Message Filter. Microsoft has continually improved SmartScreen to also look at urls contained in mails and tracks the urls down to the IP host. If anything related to the hosted IP is bad then the message is not delivered. In addition to Microsoft's own technology, the service also relies on a set of third party filters to increase its effectiveness. Once a message is reliably identified as spam it is deleted. If it is only suspected of being spam it is put in the user's junk folder.

iii. **User preferences** – this is the most effective tool by far in dealing with graymail. The users are able to set up their own blocklist and safelist rules to deal with incoming mail. In addition, Hotmail offers the user a set of powerful spam tools within Hotmail's user interface. The service displays a bar when a user is reading emails that warn them of potential danger based on the sender's reputation. Links and images are turned off by default in the case of unknown or untrustworthy senders in order to protect the user from malicious links and web beacons (i.e. malicious sites). As users contribute to the spam fighting effort by marking bad messages as junk or moving messages to their junk folders the system learns and gets smarter about what types of messages the user wants to receive. Likewise, every time a user moves a message from a junk folder back to the inbox the system learns and gets better at avoiding false positives.

iv. **Time-travelling filters** –Hotmail may not always be able to identify a new spammer the moment they start sending spam, however, once a spammer has been identified the system is able to go back and clean out spam that was delivered to the user inbox before the user even sees it. The tools are referred to as Time Travelling Filters, because in a sense they go back in time and remove spam after delivery..

v. **Malware detection** – Using SmartScreen, Hotmail scans attachments and blocks those that contain known malware and viruses.

6. 0 CONCLUSIONS AND RECOMMENDATIONS

There is no doubt that spam represents a significant, global threat to all email users. Users are increasingly frustrated by the influx of spam and the associated threats it brings in terms of malware, phishing and potential loss of personal information.

Nevertheless, things are getting better! New, improved anti-spam technologies are succeeding in keeping the amounts of spam delivered to the user inbox in check while ensuring fairly low false-positive rates.

Hotmail's anti-spam technology is at the forefront of combating spam through a layered approach that helps reduce the amount of spam delivered to the inbox while also providing easy-to-use mechanisms whereby users can reduce the amount of graymail they receive. Hotmail is continuing to invest heavily in new anti-spam technology in an effort to offer its users one of the cleanest possible online experiences. Through the effective use of various layers of technology, Hotmail is able to keep the amount of spam that gets delivered to the user’s inbox to a minimum. Hotmail’s ability to fight spam, today, is very much on par or better than accepted industry averages.

While spam will never completely go away, the ability to contain it within these levels is essential in restoring user confidence in the use of email and the overall Internet online experience as a whole. Microsoft’s solutions provide users with truly secure email and lower the potential for malware attacks without the risks of high false positive rates, ensuring that legitimate user emails are delivered safely to the inbox.

APPENDIX – NEW HOTMAIL ANTI-SPAM INVESTMENTS

Microsoft is constantly investing in new and innovative technologies to fight spam in Hotmail. This section provides a brief overview of some of the key new future capabilities now available in Hotmail.

1. **The ability to detect spammer infrastructure behind urls** – This is very important and fundamental to attacking spam where it originates. Through this capability Hotmail is able to identify different pockets of infrastructure that belong to spammers and increase their costs by forcing them to constantly get new IPs, new urls, new hosting domains, etc. The spammers whose infrastructure has been detected need to either close shop or find easier targets which are less expensive to attack.

2. **Personalization** – New investments will focus on developing a more personalized and intuitive spam management system. This will consist of:

   i. **Improved user filter block** – The user filter block will be refined using information obtained directly from the user on how they reacted to senders and what decisions were made on various senders and past email received. This is applied globally based on personalization data and works at the Mail Transfer Agent (MTA) level. The result will be a decreased incidence of false positives.

   ii. **Better understanding of user behavior** – The list management system is becoming more intuitive by tracking treatment of mail on a per individual basis. For instance, by noting that an email was categorized by Hotmail's anti-spam filters as junk and therefore consigned to the junk folder - yet the user later moved it back to the inbox as a legitimate email, the system will get smarter at classifying that type of email in the future.6 The key here is that such learning is done at the individual level since a message that might be one person's junk could very well be another person's legitimate message that they want in their inbox.

[6] Through the use of more lists and improved MTA-based learning techniques there is now more granularity in terms of how emails are handled, rather than a simple binary decision of what is to be delivered to the inbox and what isn't. So for instance, if a user classifies an email from a sender as Junk but the system sees that the user has had past conversations with that sender, the sender will not automatically be classified as blocked but will instead be put in an intermediary Junk list and the user's future behavior will be tracked to make a further recommendation.

This type of machine learning can be based on a variety of factors, such as: first contact, geography, language, delete history, and more.

3. **Trusted Sender** – Hotmail identifies the top phishing targets and uses identification protocols to demarcate messages from these senders that are legitimate – not phishing scams. This demarcation is done through placement of a visible icon next to the legitimate messages.

4. **Enhancements to Time Traveling Signatures** – Hotmail has made additional investments in its ability to retroactively (within milliseconds) remove spam based on signature discovery.

5. **Tagging** – Hotmail is further appending mail that's in a person's junk folder with educational messaging such that a person can understand why the message was placed there and, if legitimate, can avoid similar mail being placed there in the future, minimizing false positives.