White Paper X-Spam for Exchange 2000-2003 Server

X-Spam for Exchange 2000-2003 (X-Spam) is a highly adaptive Anti-Spam Software that protects the Microsoft Exchange 2000-2003 servers from Spam. X-Spam uses advanced regular expressions and meta tests to positively identify Spam and block it. This white paper discusses Spam and provides information on X-Spam.

What is Spam

Spam is unsolicited junk mail sent to you or your mail server. People who indulge in such activities are called Spammers. Commercial advertisers who may offer dubious products, get rich schemes, products that do not suit your life style, promote illegal activities, etc, send Spam. The intent here is to make you spend money. Almost 60% - 70% of Spam is related to Porn. It costs the Spammer almost nothing to send mails and invariably get an eyeball count.

There is another type of spammer who sends large number of e-mails that flood your mailbox or mail server. The intent here is to cripple your e-mail service to such an extent that you cannot receive genuine mails. This is termed as Denial-of-Service (DoS) attack.

Losses Caused by Spam

Experts estimate that almost 50% of all email exchanged across the world's networks is Spam and this assessment is expected to increase despite anti-Spam laws in countries such as the United States and United Kingdom. Apart from clogging important network and email resources such as servers and network bandwidth, Spam is a major nuisance for many email users in terms of the time wasted sifting through their mailboxes for legitimate mail as well as being exposed to obscene and at times offensive content. Though the majority of Spam targets individuals for the sale of products and services, some are also used as tools for committing frauds as well as for stealing account and personal information.

Earlier, Spam mails could be filtered using simple word/phrase based content filters. However, over the past couple of years, Spammers have started using techniques (such as message hiding using rogue html tags, including non-alphabetic characters in their messages, using invisible fonts, impersonation, etc.) that allow their messages to bypass simple filters. Since the message content and presentation also keeps changing, it becomes even more difficult for administrators and users to keep their custom filters updated.

Sometimes, Spam is indistinguishable from legitimate mail because of it's innocent content. Setting filters for such mails may lead to a high false positive (legitimate mails tagged as Spam) rate. What is required is a dynamic solution that identifies new Spam, keeps the false positive rates low and requires minimum intervention.

How Spam is delivered

Some of the methods used to spread Spam are listed below:

- Spam is received through e-mails and may have alluring subjects like: Free offer, Chance of a lifetime, etc. Invariably you try to open the mail and read it. That is what the Spammer wants you to do. Opening the mail, reading it and then deleting it, consumes your Internet access time and costs you money. The mail servers that have delivered the mail through a series of servers have spent money and used bandwidth to deliver junk you did not want. Probably the junk mail was ahead in the queue for mails to be delivered and was given precedence over an urgent mail.
- Some Spam mails have attachments and the mail asks you to open it. If you do so, you risk running a virus that may be hidden in the mail. The costs involved in removing a virus from your system are massive.
- Some Spam mails after enclosing an alluring description of products or services, ask you to click on a link for further information. These links may open porn or other sites that you had no business visiting. But details of the visit are recorded in your server and you may have a lot of explaining to do.
- Products advertised through Spam require that you provide your credit card number and other personal information. Besides getting your account billed for junk items, you also open yourself to more Spam.

How Spamming is Done

There are many ways in which Spamming is done. The commonly used ones are:

- **Rogue ISPs:** Spammers who have enough cash to fund their illegal activities run these. Rogue ISPs obtain their own network numbering and multiple domain names from the InterNIC. Spammers use multiple domain names and manage to get across Spam blocks. While it is possible to block a domain, it is not possible to block an ISP provider.
- **On-the-fly' Spammers**: Such type of Spammers, register as multiple genuine users for trial accounts with ISPs. Forged identity or stolen credit cards are used to establish identities. They then use these accounts to start their Spam hits. By the time the ISP realizes that they are hosting a Spam run, the Spammer uses another account.
- **Blind Relays**: Some innocent servers allow Blind Relaying – relaying messages without authentication. Spam is routed through such blind relays.

What is X-Spam for Exchange 2000-2003

X-Spam is a highly adaptive Anti-Spam product that keeps the mailboxes of users free of Spam at the Mail Server itself. X-Spam for Exchange 2000-2003 fights Spammers both before accepting mail at the mail protocol level and after the mail is accepted.

X-Spam is installed as a Windows Service on Windows 2000/2003 running Exchange Server 2000/2003. X-Spam is a scalable solution that uses various key Spam-fighting techniques to stop Spam at the mail server level itself. Just like anti-virus software, X-Spam keeps its Anti-Spam engine updated by automatically downloading the latest filters.

Brief Overview of Design

X-Spam consists of an SMTP proxy server that sits as a gateway between the Internet and Exchange Server. All the mail protocol level checks are carried out in the proxy. Along with the proxy server is a COM+ application that is closely integrated with Exchange Server 2000-2003. This application is responsible for all the content filtering on the email when it arrives in the user's inbox on Exchange Server and also the actions related to mails once they are identified as Spam or legitimate mail.

X-Spam Anti-Spam Technology

X-Spam works at the mail protocol level and also after a mail is accepted. Thus it provides two levels of security for your email infrastructure. X-Spam uses Heuristics Scoring to reduce the probability of catching legitimate emails as Spam. These scores are generated using a special algorithm from the area of Artificial Intelligence.

Apart from its own filters, X-Spam allows users to design their own special filters either for simple content filtering or for powerful regular expression filtering. Multiple tests can be combined using Boolean expressions to create filters that test multiple properties of mails. X-Spam fights Spam at two levels:

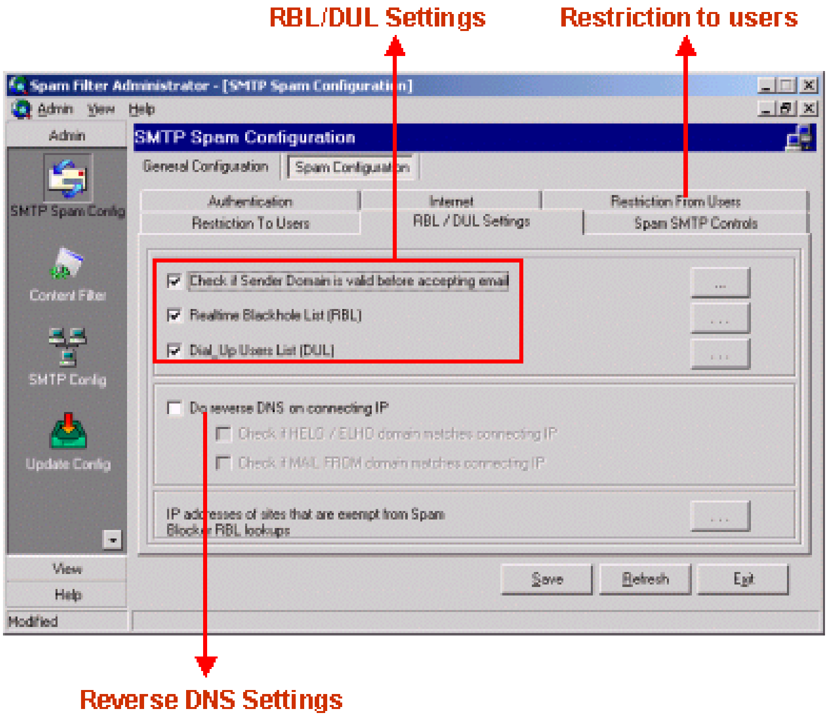
❿ Before accepting mail at the mail protocol layer: If a connecting mail server is listed in a Spam sending black list, mails are not accepted from it. Users can easily add preferred black lists to the default list. There are several other options that allow users to check the validity of the domain in the sender's email address.

❿ After mail has been accepted, it goes through a series of special filters that add to the total Spam score of the mail. If the mail crosses the Spam threshold, it is tagged as Spam.

Application Features

MAIL PROTOCOL LAYER CHECKS

X-Spam performs a series of checks at the protocol layer to check for Spam. Figure 1 gives a screen shot from X-Spam with these features.



RBL/DUL Settings

Restriction to users

Reverse DNS Settings

☑ Check if Sender Domain is valid before accepting email

☑ Realtime Blackhole List (RBL)

☑ Dial_Up Users List (DUL)

☐ Do reverse DNS on connecting IP

☐ Check if HELO / ELHO domain matches connecting IP

☐ Check if MAIL FROM domain matches connecting IP

IP addresses of sites that are exempt from Spam Blocker RBL lookups

Figure 1. Mail Protocol Layer Checks

The checks are explained below:

Real-time Black List (RBL)

An RBL is a DNS server that lists IPs of known Spam sending machines. There are many such lists available to fight Spam and X-Spam allows you to add your own choice of RBLs to the list that gets shipped with the product. RBL checking in X-Spam for Exchange 2000-2003 is carried out at the SMTP protocol level when the sending mail server connects to your mail server. If the IP is found to be in any one of the blacklists, the connection is terminated.

Dialup User's List (DUL)

DUL is a DNS server that lists IPs of machines that are known to be part of the IP-range assigned to dialup Internet accounts. Spammers prefer using dialup accounts to send mail since they are easy and cheaper to sign up for and dispose. Most legitimate dialup user's usually connect to their ISPs mail servers or an Internet email account to send mail. Also, since dialup accounts have their IPs dynamically assigned they can't have DNS MX (mail server) records pointing to them. This implies that one cannot possibly have a properly configured mail server on a dialup account IP. An IP found to be on a DUL is more often than not probability a spammer. A good practice is to require mobile users to use authentication when connecting to the company mail server. The list of DULs is extensible just like RBL. If a connecting IP is found to be in any one of the blacklists, the connection is terminated.

MX/A DNS record verification

The domain part of the email address given in the "MAIL FROM" SMTP command is checked to see if it has a DNS MX (mail server) and/or A (IP) record. Every email address domain should at the very least resolve to an A record and every properly configured mail server domain should also have an MX record. Spammers or misconfigured client programs are the only ones who use non-existent domains in their emails. However, in either case it is absolutely valid to refuse acceptance of mail.

Reverse DNS

A reverse DNS check is performed to see if the connecting IP resolves to a valid domain name. Most properly configured mail servers usually have a reverse lookup (PTR) record entry in DNS. Further verification for the legitimacy of the resolved domain is done by checking if there is a sub domain match with the domain in "HELO/ELHO" and/or the "MAIL FROM:" SMTP commands. Since a majority of mail servers do not have a reverse lookup, this feature is disabled by default.

IP/Host/User Restriction From List

⑩ **Whitelist**: If an IP/Host/User is on this list, all mails from that IP or from that user/domain are accepted.

⑩ **Blacklist**: If an IP/Host/User is on this list, all mails from that IP or from that user/domain are rejected.

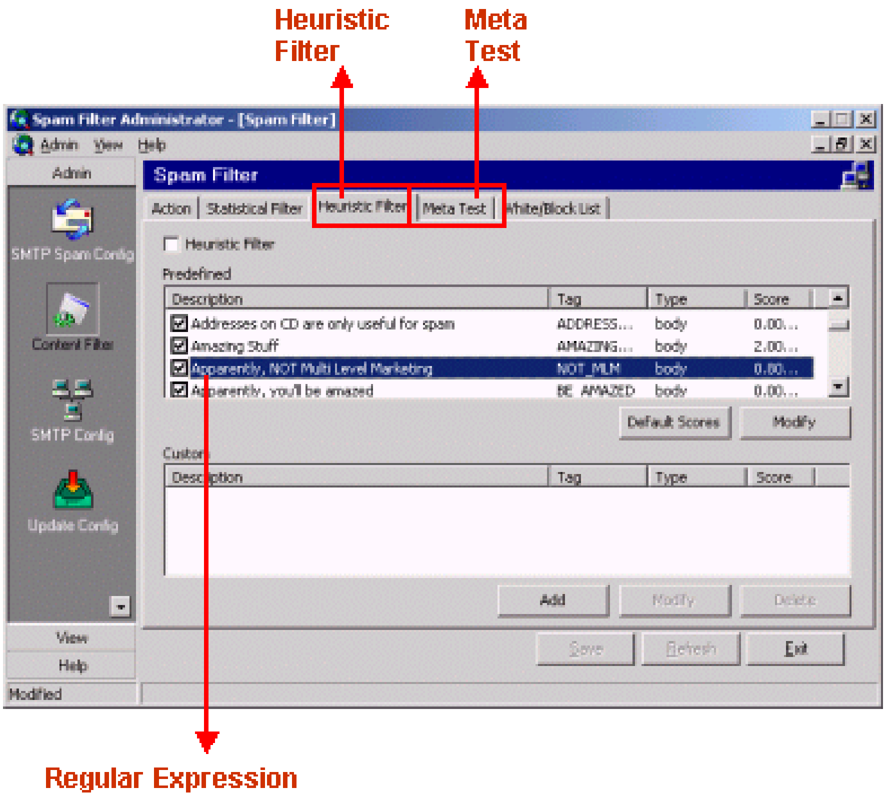
Host/User Restriction To List

⑩ **Whitelist**: If a Host/User is on this list, all mails to that user/domain are accepted.

⑩ **Blacklist**: If a Host/User is on this list, all mails to that user/domain are rejected.

EMAIL CONTENT CHECKS

X-Spam performs rigorous email content checks. Figure 2 gives a screen shot from X-Spam with these features.



□ Heuristic Filter

Predefined

Description	Tag	Type	Score
☑ Addresses on CD are only useful for spam	ADDRESS...	body	0.00...
☑ Amazing Stuff	AMAZING...	body	2.00...
☑ Apparently, NOT Multi Level Marketing	NOT_MLM	body	0.00...
☑ Apparently, you'll be amazed	BE_AMAZED	body	0.00...

Custom

Description	Tag	Type	Score

Figure 2. Email Content Checks

Domain/User Content Check White List If a Domain/User is on this list, all mails to that user/domain are accepted "without content checks".

Regular Expression and Content Filter After mail has been accepted, it is checked with a whole range of predefined and user defined regular expression and content tests. Unlike simple filtering based on simple words and phrases, the regular expression tests are used to check for various properties of restricted content within the header, body and text/html attachments of emails. A regular expression provides a powerful way to describe several different permutations of content as well as properties of restricted text in the email body. With regular updates from MicroWorld, these tests will be kept updated. A facility is provided so that a user can add his/her own customizable tests to the predefined tests. One can also disable certain tests to prevent a high false positive rate although we doubt if that will ever be needed. Regular expression tests can be specified for specific headers, the text body, urls and the unprocessed text/html body. Not only do these tests filter emails based on reserved content but also perform header and content analysis.

Meta Filter Meta Filter combines various non-meta tests to form even more powerful Spam fighting tests. Boolean expressions are used to combine various tests into more complex and concrete tests. X-Spam for Exchange 2000-2003 comes with predefined meta tests and these will be regularly updated through X-Spam for Exchange 2000-2003's update mechanism.

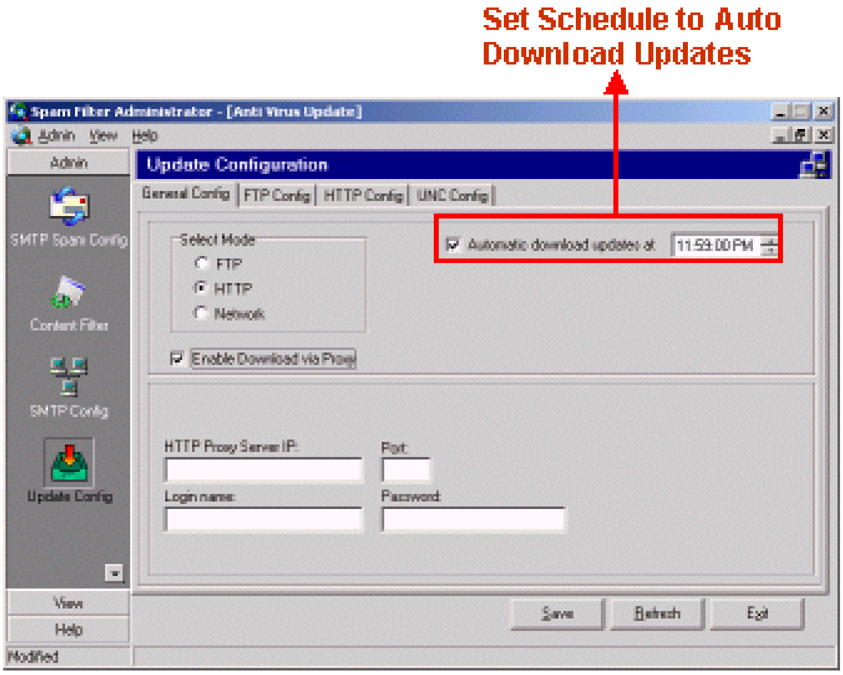
Heuristics When heuristics are enabled, each regular expression is associated with a score. When a test is positive, its score is added to the total Spam score of that email and tested to see if it has either crossed a user configurable threshold for Spam and non-Spam. If the test crosses either, no further tests are executed. New scores will be regularly generated using a special AI-based algorithm on MicroWorld's archive of recent Spam and legitimate mails so as to prevent false positives. These new scores are then downloaded using the update mechanism. Users can also specify their own scores using the User Interface.

OTHER FEATURES

Automatic Updates X-Spam for Exchange 2000-2003 has self-learning and automatic detection capabilities that recognize the Internet connectivity. When X-Spam for Exchange 2000-2003 detects an Internet connection, it automatically connects to MicroWorld's FTP server to check for and download filter and heuristics score Updates. With this feature (that works on dedicated as well as dial-up lines), all X-Spam for Exchange 2000-2003 users across the world receive updates on a regular basis.

This feature makes our product adaptive and resilient to new Spam mails that might emerge months after you have purchased our software.

Automatic updates can be via FTP or HTTP. Proxy and Firewall (Passive FTP) support is provided. Broken FTP Updates can be resumed from breakpoint.



Set Schedule to Auto Download Updates

Spam Filter Administrator - [Anti Virus Update]

Update Configuration

Select Mode: ○ FTP ● HTTP ○ Network

☑ Automatic download updates at 11:59:00 PM

☑ Enable Download via Proxy

HTTP Proxy Server IP:

Port:

Login name:

Password:

Figure 3. Auto download Updates

Customizable Actions All the settings are highly configurable and customizable. After a mail is accepted and identified as Spam, the following actions can be taken:

- ❿ It can be moved to a special Spam folder in a user's inbox for an IMAP setup.
- ❿ It can be tagged in the subject line.
- ❿ A special header can be inserted to indicate it is Spam.

❿ The mail can be deleted although we do not recommend this since false positives can then never be recovered.

Reports ❿ X-Spam supports the generation of daily, weekly and monthly reports with intuitive graphs and also sends daily summaries of mail classification to the administrator.

Relay-Control It provides Relay Control using IP-blocks (to ensure that unauthorized IP addresses are blocked from sending mails through the Mail Server). X-Spam for Exchange 2000-2003's SMTP proxy can handle multiple domains. It provides Hop Control to ensure that the mails do not loop and also provides routing facilities.

System Requirements

Available for Platforms: Windows 2000/2003 Server.

Exchange Server: 2000 (SP3+) / 2003.

P-II and above Processor

128-256 MB of Physical Ram

50 - 100 MB of free H/Disk space

Compatibility

X-Spam for Exchange 2000-2003 can be combined with the following MicroWorld products for a complete content security solution:

- ❿ MailScan 4.2a for SMTP Servers
- ❿ MailScan 4.2a for Exchange 2000-2003
- ❿ eScan

Contact us

For details, please visit	www.mwti.net
For sales enquiries	sales@mwti.net
For support	support@mwti.net