

HOW TO GET RID OF SPAMS: MFILTRO Technologies explained

**“NETASQ MFILTRO has one of the highest spam catch rates on the market today,
with zero false positives during our tests”**

- DataNews 7 Sept 2007

PURPOSE OF THIS DOCUMENT

The biggest demand of the market for an anti-spam solution is **efficiency** and **simplicity**. This whitepaper explains in more detail how the different technologies are used in NETASQ MFILTRO.

ANTI SPAM Techniques

The anti-spam policies used in NETASQ MFILTRO consist of two main parts:

- ✓ E-mail context, works at the protocol level
- ✓ E-mail content, works on data

The first part includes all rules that act at the protocol level during the SMTP connection. Among these techniques NETASQ MFILTRO uses IP Reputation – dynamic RBL, dynamic URLBL, local IP white and black lists. Early talker detection is used to block spam scripts.

Other techniques used include valid domain check based on MX and/or A records, sender domain and sender email address white and black lists, tarpitting dynamic IP addresses etc.

Valid recipient email addresses can be checked using a local list, SMTP RCPT check or LDAP/Active Directory queries.

All these techniques work on the connection level, which implies that the potential spam email never comes over your Internet connection. In an average situation, **80% of all SPAM is already stopped at the protocol level.** NETASQ MFILTRO does not only free resources on your internal mail servers, but also frees bandwidth by blocking 80% of all spam before it even crosses your Internet connection.

The second part of the NETASQ MFILTRO policy, is looking at the content of the emails. NETASQ MFILTRO anti-spam engine will look at the entire email, headers and body included. Attachments can be stripped to enforce company policy. Content inspection based on words or sentences can be configured.

This heuristic filtering consists of several thousands of rules of different nature, each being applied to a message in order to deliver an elementary score.

Heuristic rules are empirical, non predictable rules derived from advanced analysis of all parts of a message:

1. header fields, with a special mention for the subject text,
2. plain text part if any,
3. HTML part if any,
4. attachment names and contents where applicable.

Heuristic rules are established by human experts searching for unexpected common features of various messages (especially when partly or entirely generated by robots), which will of necessity be shared by future spam, no matter the topic.

Deriving heuristic rules requires perfect command of all protocols involved in email over the Internet, as well as deep knowledge of spamming nuts and bolts in general. Human search for new heuristic rules is computer aided, with dedicated tools that have come of age along with the filter itself. New hypotheses are swiftly checked against new SPAM as well as against established SPAM and ham corpora. Though less frequent, negative score rules are of paramount importance in limiting "false positive" risk.

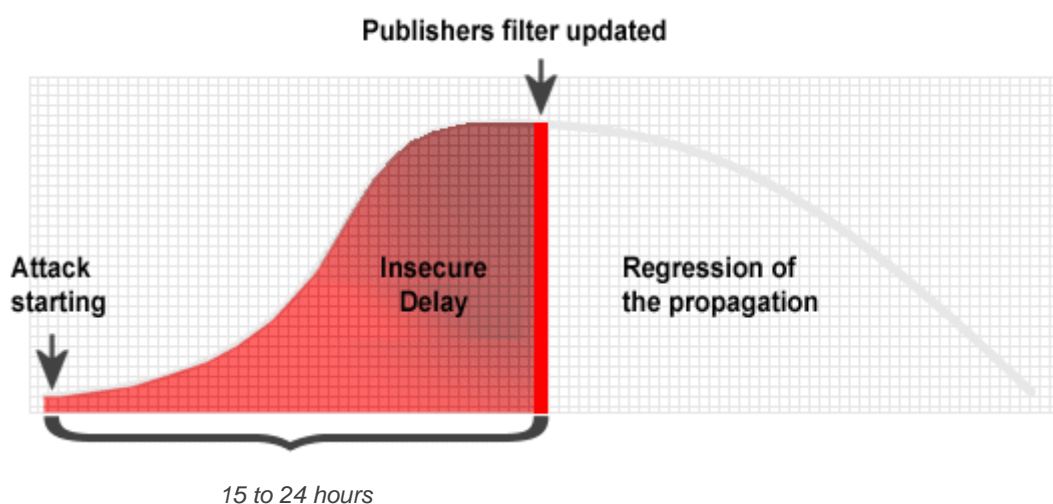
Thanks to the Predictive Heuristic Filter, NETASQ MFILTRO technology is able to anticipate some unwanted emails and viruses before they spread. This innovative feature was designed to ensure protection against a new attack whereas other software publishers need time to update their anti-spam or antivirus engine.

Predictive Heuristic Filter

Heuristic prediction is an original concept developed by NETASQ MFILTRO engineers and part of the anti-spam engine. It enables to answer to the problem of critical and necessary intervention period that publishers encounter when they face outbreak.

Unlawful viral attacks are more and more frequent and devastating. A few years ago they needed days or weeks to go around the world. Nowadays a few hours are enough to infect millions of workstations. Trojans and worms have created a network which enables SPAM and viruses to spread all around the world very quickly (called botnets).

Attacks are more and more devastating so antivirus publishers have to react more quickly. However there is always a critical insecurity period during which viruses and unwanted emails can spread through thousands of computers.



This schema represents insecurity period during outbreaks.

Thanks to Predictive Heuristic Filter, NETASQ MFILTRO protects users from the start of the attack.

Heuristic prediction enables to filter spam and virus attacks without requiring any intervention or any update. Thanks to this, users are protected even during the insecurity period.

Looming systems which come by concealed doors to spread massively have a weakness: they focus on viral innovation or the visible content of the message. Creating infected emails requires techniques which evolve less quickly. Smart analysis and action on the message's structure are enough to block a virus-carrier message. To do this it is not necessary to analyse the infected file.

Developing the heuristic prediction system consists in keeping only the filtering rules which are based on unlawful emails sequences with a high reappearance probability. A rule become and may be considered as predictive when it has been developed to answer specifically to an attack and it keeps on react to new ones.

Other Features of the anti-spam Engine

Counter-reaction

Counter-reaction filtering is one of the most advanced and effective parts of the NETASQ MFILTRO engine. It basically consists of detecting the nuts and bolts used by spammers to outsmart the anti-spam filters of the first generation. Below is a joint list of first generation anti-spam techniques, spamming counter-techniques and MFILTRO counter-counter-techniques:

Anti-spam basic technique: blacklists of websites pointed to by links in messages

Spammers answer: redundant encoding of links

MFILTRO efficient protection: redundant URL encoding detection

Anti-spam basic technique: message footprint -based filtering

Spammers answer: random string insertion to make no two message contents identical

MFILTRO efficient protection: random string detection

Anti-spam basic technique: word association statistical (Bayesian) analysis

Spammers answer: adding neutral words, either readable or low-contrast / small print

MFILTRO efficient protection: detection of unrelated word sequences, unreadable or barely readable text, excessive use of small print

Anti-spam basic technique: keyword -based filtering

Spammers answer: spelling changes within human readability and/or English phonetics

MFILTRO efficient protection: detection of typical alterations to sensitive words and HTML tags. NETASQ MFILTRO engine is able to detect hundreds of alterations for "viagra", from V1AGR4 to VVIAAGGRRRIA to VIGRA , none of which being memory -stored as such

HTML patterns

Whenever there is an HTML part within a submitted RFC-822 message (which nowadays occurs more often than not), MFILTRO computes an exclusive HTML code footprint (HTML pattern). It then compares it to a list of known patterns, demonstrably typical of generated spam. This technique, combined to statistics on the image sizes within, provide for a **particularly effective filtering of spam mainly or exclusively made of online images.**

Detection of forged SMTP timestamps and other header entries

Detection of forged Received: and other message header entries emerged in time as a major technique for detecting spam.

Anti-scam

Cyber-scam comes in various flavours, most often as carefully worded financial propositions, aimed at luring the addressee into supposedly overseas high-earners. NETASQ MFILTRO runs a dedicated scam detection module, as e-mail scam has little likeness to everyday, advertising spam.

ANTIVIRUS Techniques

NETASQ MFILTRO can use different third party antivirus solutions. Both ClamAV and Kaspersky are implemented. For most effective use of the Predictive Heuristic techniques described above, ClamAV is tightly bound to the Heuristic Engine. For this reason only one signature update is needed for both ClamAV and MFILTRO Heuristic signatures.

Next to the NETASQ MFILTRO Engine application, the Kaspersky application can be called during the body analyses of the email. This leaves great flexibility in the policy: for some domains only ClamAV can be used, for other domains Kaspersky and for other domains **even both as a double protection.**

Since NETASQ MFILTRO can invoke these applications from within the policy, this can not only be configured on the domain level, but also on email addresses, LDAP users or groups, local lists etc. This flexibility is especially useful in Managed Services environments or complex WAN Networks.

NETASQ MFILTRO MTA

The Mail Server or Mail Transfer Agent (MTA) is a concept unique to NETASQ MFILTRO.

Schematically, all messages coming to the MTA, get processed according to the specific policy applied to the message stream it belongs to (typically, but not necessary based on recipients domain), and is then destroyed, quarantined, stored, or routed to a mail server according to the result of the processing. This means great flexibility, for example mail routing could be based on an LDAP query.

The MTA is the structure which manages the rules to be applied to each message stream. Its configuration web-GUI allows the system administrator to easily define and implement processes performed for each message stream.

The MTA can store several Workflow Policies while only one is active at any given time. This makes it simple to switch from one policy to another in a matter of minutes, without stopping the MTA or losing messages. Each Workflow Policy is made of rules which are processed in the set order. The MTA is usually configured through its Web-GUI, but configuration files, in

XML format can also be uploaded, and edited. In any case, when enabled, the rule set is compiled for faster execution.

Beyond Mail Firewall applications, the uniqueness of the MTA lies in the incredible simplicity of integrating NETASQ MFILTRO in an existing message infrastructure. Many functions, such as domain masquerading, routing to several mail servers, or checking specific services subscribed by the recipient, which usually require coding or scripting, are just a matter of point and click with the GUI. Furthermore, the very rich abilities of the MTA allows NETASQ MFILTRO to replicate any given legacy mail service, including conditional recipient address transcription. The numerous functions which control the traffic from the MTA towards the mail servers will make sure that your servers are never overloaded, since the MTA can act as a temporary store while you maintain your servers.

Advanced Policy Rules can include modifications of addressees or senders as well as content. Information can be added or removed from messages, and message routing can be modified. Messages can be forwarded, copied, sent to distribution lists, rejected, delayed or queued.

All these actions can be combined in complex rules and tied to fulfilment of specific prerequisites.

Simplicity

The User Quarantine is very easy to use and has no administrative overhead. The User Quarantine can be accessed by one-time tokens provided in the daily quarantine reports. Thanks to this mechanism, there is no need for usernames or passwords that the administrator has to create and maintain, and no need for support on lost passwords. Users can create their own white lists based on domains or e-mail addresses directly from the daily report without the need to log into their quarantine first.

Also thanks to the heuristic engine there is no need to instruct MFILTRO about spam. This relieves the users from spending time on categorizing emails and reporting back to the appliance.

NETASQ MFILTRO can be fully configured with a 10 minute installation wizard, including registration, firmware upgrade, e-mail routing, exceptions, and user quarantine settings. This allows the administrator to plug and forget, with close to zero maintenance after installation.

Simplicity both for the administrator and users ensures the shortest time spent on configuration and maintenance, **which provides the best Return On Investment possible.**

Conclusion

The techniques explained in this document make NETASQ MFILTRO one of the most **efficient** spam combatants on the market today. The 10-minute installation wizard, together with the easy-to-use User Quarantine, guarantees **simplicity**.

NETASQ MFILTRO uses a variety of best-of-breed solutions to deal with spam and viruses. For anti-virus functions, NETASQ MFILTRO uses ClamAV and/or the award-winning gateway anti-virus solution Kaspersky. The Predictive Heuristic Engine is used to stop virus and spam outbreaks before signature updates are available.