COMPARATIVE REVIEW

VBSPAM COMPARATIVE REVIEW

Martijn Grooten

Thanks to many programs that are freely available on the Internet, building a spam filter is not rocket science. However, building a *good* spam filter is not a trivial task. And, with the email security market still growing and new products appearing every month, many customers will wonder whether a hitherto unknown product, with a shiny website and an impressive sales story, is actually any good.

The purpose of VBSpam testing is to provide an easy-to-recognize certification that tells potential customers that a product does what a good spam filter should do: i.e. block the vast majority of spam, with very few false positives. As such, we are delighted that, for the first time, all of the products in this month's test achieved a VBSpam award. This does not mean that no bad products exist – after all we only test products that have been submitted by their developers – but it does demonstrate that there is plenty of choice for customers, as well as a healthy amount of competition for product developers.

But in spam filtering, the devil is in the details. With recent reports suggesting that up to 95% of email traffic is spam, email can only be a viable form of communication for businesses if the vast majority of that spam is blocked – and blocking one or two per cent more will have a huge impact on users' inboxes. Similarly, a very low false positive rate is essential, and even a couple fewer false positives every month will significantly improve user experience. For this reason we provide detailed performance measurements for all 16 of the products tested this month.

THE TEST SET-UP

No major modifications were made to the test set-up, and as usual the full methodology can be found at http://www.virusbtn.com/vbspam/methodology/. In this test developers were offered the option of receiving information on the original sender's IP address and HELO/EHLO domain during the SMTP transaction, thus emulating a real environment where many messages are blocked because of the IP addresses and/or the domains of the senders. However, none of the developers chose to make use of the option on this occasion.

As in previous tests, the products that needed to be installed on a server were installed on a *Dell PowerEdge R200*, with a 3.0GHz dual core processor and 4GB of RAM. The *Linux* products ran on *SuSE Linux Enterprise Server 11*; the *Windows Server* products ran on either the 2003 or the 2008 version, depending on which was recommended by the vendor. (It should be noted that most products run on several different operating systems.)

To compare the products, we calculate a 'final score', defined as the spam catch (SC) rate minus three times the false positive (FP) rate. Products earn VBSpam certification if this value is at least 96%:

$$SC - (3 \text{ x } FP) \geq 96\%$$

THE EMAIL CORPUS

The test ran from 6pm GMT on 9 February 2010 until 7am on 1 March 2010, with an unscheduled break between 17 and 22 February when problems beyond our control left us without reliable information to base the test results on. The corpus contained 254,407 emails: 2,458 ham messages and 251,949 spam messages, where the latter consisted of 237,783 messages provided by Project Honey Pot and 14,166 messages sent to legitimate @virusbtn.com addresses.

Some new email discussion lists were added to the ham set and, as in previous tests, emails that claimed to be sent from @virusbtn.com addresses were removed from the test set. (Note that this check was only applied on the MAIL FROM, not on the email headers, and in future tests, these emails will not be removed from the test set.)

For each product, no more than four false positives were counted per sender. The 'image spam' and 'large spam' categories referenced in the test results are, respectively, spam messages containing at least one inline image, and those with a body size of over 50,000 bytes.

BitDefender Security for Mail Servers 3.0.2

SC rate (total): 97.96%

SC rate (Project Honey Pot corpus): 98.68%

SC rate (VB spam corpus): 89.85%

SC rate (image spam): 96.33%

SC rate (large spam): 93.16%

FP rate: 0.04%

Final score: 97.84%

Most products in this month's test saw a slight reduction in their spam catch rate, and this included *BitDefender*. However, *BitDefender* more than made up for this by missing just a single legitimate email out of 2,400. The product easily earns its sixth VBSpam award in a row.

(Note: On careful investigation of the previous test results – see VB, January 2010, p.23 – it was discovered that BitDefender's false positive score should

	True negative	False positive	FP rate	False negative	True positive	SC rate	Final score
BitDefender	2457	1	0.04%	5144	246805	97.96%	97.84%
FortiMail	2453	5	0.20%	5401	246548	97.86%	97.26%
Kaspersky	2449	9	0.37%	5148	246801	97.96%	96.85%
M86 MailMarshal	2454	4	0.16%	1717	250232	99.32%	98.84%
McAfee Email Gateway	2438	20	0.81%	2208	249741	99.12%	96.69%
McAfee EWSA	2456	2	0.08%	4281	247668	98.30%	98.06%
MessageStream	2452	6	0.24%	2762	249187	98.90%	98.18%
MS Forefront	2454	4	0.16%	602	251347	99.76%	99.28%
MXTools	2458	0	0.00%	4922	247027	98.05%	98.05%
Sophos	2454	4	0.16%	1787	250162	99.29%	98.81%
SPAMfighter	2446	12	0.49%	5099	246850	97.98%	96.51%
SpamTitan	2452	6	0.24%	1858	250091	99.26%	98.54%
Sunbelt VIPRE	2444	14	0.57%	4004	247945	98.41%	96.70%
Symantec Brightmail	2456	2	0.08%	2263	249686	99.10%	98.86%
Webroot	2456	2	0.08%	3192	248757	98.73%	98.49%
Spamhaus	2458	0	0.00%	5529	246420	97.81%	97.81%

have been 15, rather than the reported 17. This gave the product a FP rate of 0.534% and a final score of 96.51%.)

Fortinet FortiMail

SC rate (total): 97.86%

SC rate (Project Honey Pot corpus): 98.14%

SC rate (VB spam corpus): 93.03%

SC rate (image spam): 95.66%

SC rate (large spam): 93.50%

FP rate: 0.20%

Final score: 97.26%

Fortinet's *FortiMail* appliance has been filtering *VB* email without any problems for five tests in a row. A lower false positive rate on this occasion saw the product's final score improve a little to fully merit its fifth consecutive VBSpam award.

Kaspersky Anti-Spam 3.0

SC rate (total): 97.96%

SC rate (Project Honey Pot corpus): 98.47%

SC rate (VB spam corpus): 89.37%

SC rate (image spam): 97.56%

SC rate (large spam): 94.23%

FP rate: 0.37%

Final score: 96.85%

Kaspersky Anti-Spam did not miss a single legitimate email in the previous test but, thanks to a rather low spam catch rate, the product failed to win a VBSpam award. The product's developers used the feedback from the last test to improve its heuristics-based botnet traffic detection. Indeed, the spam catch rate saw a significant increase and although there were some false positives this time, the product easily reclaimed its VBSpam award.

M86 MailMarshal SMTP

SC rate (total): 99.32%

SC rate (Project Honey Pot corpus): 99.46%

SC rate (VB spam corpus): 96.95%

SC rate (image spam): 99.83%

SC rate (large spam): 98.86%

FP rate: 0.16%

Final score: 98.84%

M86 MailMarshal SMTP – which was tested on *Windows Server 2003*, but which also runs on *Windows Server 2008* – was the highest ranking product in the last test. On this occasion the product saw both its SC rate and its FP rate worsen a little, but not in a significant way, and with the third highest final score, the product is once again ranked highly in this test.

McAfee Email Gateway (formerly IronMail)

SC rate (total): 99.12%

SC rate (Project Honey Pot corpus): 99.37%

SC rate (VB spam corpus): 95.02%

SC rate (image spam): 99.19%

SC rate (large spam): 98.16%

FP rate: 0.81%

Final score: 96.69%

McAfee's *Email Gateway* appliance caught well over 99% of all spam for the fourth time in a row and the product wins its fourth consecutive VBSpam award. However, *Email Gateway* false positived on more legitimate emails than any other product – it had particular difficulties with emails from Eastern European and Asian countries – and there is definitely room for improvement in this area.

McAfee Email and Web Security Appliance

SC rate (total): 98.30%

SC rate (Project Honey Pot corpus): 98.75%

SC rate (VB spam corpus): 90.82%

SC rate (image spam): 92.89%

SC rate (large spam): 94.01%

FP rate: 0.08%

Final score: 98.06%

McAfee's *Email and Web Security Appliance* performed a little disappointingly in the last test, displaying a higher false positive rate than in earlier tests. Further investigation determined that this had been caused by the product sending some temporary failure responses over the Christmas period. The rules of the test stipulate that email that has not reached the back-end MTA one hour after its original delivery will be considered to have been marked as spam, but it is fair to say that in a real situation – with most legitimate senders resending over longer periods of time – this would have led to short delays in email delivery and probably not to false positives. Happily, the product has been working steadily since, missing just two legitimate emails on this occasion, and with an impressive spam catch rate.

MessageStream

SC rate (total): 98.90%

SC rate (Project Honey Pot corpus): 99.19%

SC rate (VB spam corpus): 94.11%

SC rate (image spam): 98.02%

SC rate (large spam): 96.88%

FP rate: 0.24%

Final score: 98.18%

The *MessageStream* hosted solution was one of the first products to join the VBSpam tests and the developers' confidence in their product has proven to be justified time and time again. With another good spam catch rate and missing just a handful of legitimate emails, the product fully deserves a VBSpam award.

Microsoft Forefront Protection 2010 for Exchange Server

SC rate (total): 99.76%

SC rate (Project Honey Pot corpus): 99.86%

SC rate (VB spam corpus): 98.06%

SC rate (image spam): 99.86%

SC rate (large spam): 99.04%

FP rate: 0.16%

Final score: 99.28%

One of the top performers in the previous test, *Microsoft*'s *Forefront Protection 2010 for Exchange Server* saw its performance improve even further and the product outperformed its competitors in all spam categories. Thanks to just four false positives, *Forefront* was the only product to achieve a final score of over 99%.

MXTools Reputation Suite

SC rate (total): 98.05%

SC rate (Project Honey Pot corpus): 98.66%

SC rate (VB spam corpus): 87.70%

SC rate (image spam): 96.79%

	Project Honey Pot spam		VB spam corpus		Image spam*		Large spam*	
	False negative	SC rate	False negative	SC rate	False negative	SC rate	False negative	SC rate
BitDefender	3142	98.68%	1438	89.85%	282	96.33%	186	93.16%
FortiMail	4413	98.14%	988	93.03%	334	95.66%	177	93.50%
Kaspersky	3642	98.47%	1506	89.37%	188	97.56%	157	94.23%
M86 MailMarshal	1285	99.46%	432	96.95%	13	99.83%	31	98.86%
McAfee Email Gateway	1503	99.37%	705	95.02%	62	99.19%	50	98.16%
McAfee EWSA	2980	98.75%	1301	90.82%	547	92.89%	163	94.01%
MessageStream	1927	99.19%	835	94.11%	152	98.02%	85	96.88%
MS Forefront	327	99.86%	275	98.06%	11	99.86%	26	99.04%
MXTools	3179	98.66%	1743	87.70%	247	96.79%	183	93.27%
Sophos	1120	99.53%	667	95.29%	79	98.97%	111	95.92%
SPAMfighter	3956	98.34%	1143	91.93%	151	98.04%	97	96.44%
SpamTitan	1280	99.46%	578	95.92%	40	99.48%	41	98.49%
Sunbelt VIPRE	3368	98.58%	636	95.51%	357	95.36%	132	95.15%
Symantec Brightmail	1397	99.41%	866	93.89%	89	98.84%	101	96.29%
Webroot	2709	98.86%	483	96.59%	41	99.47%	60	97.79%
Spamhaus	3530	98.52%	1999	85.89%	252	96.72%	193	92.91%

*There were 7,691 spam messages containing images and 2,721 considered large; the two are not mutually exclusive.

SC rate (large spam): 93.27%

FP rate: 0.00%

Final score: 98.05%

MXTools Reputation Suite combines *Spamhaus ZEN + RBL* (see below) with the *SURBL* URI blacklist and the *Server Authority* domain reputation service. I stated in the last review that the performance of the latter two depends on the way URIs are detected in emails. We have since found a bug in the script that detects URIs which caused the product to miss several domains, in particular most .cn domains.

Fixing this bug (while also realizing that a lot of spammers have recently moved from .cn to .ru domains) saw the suite's performance improve to a spam catch rate of well over 98%. Equally impressively, there were no false positives this time. The relatively low spam catch rate on the VBSpam corpus suggests that those employing the suite in a real situation would do well to run a filter on the full content of the email too, but nevertheless the suite is the deserving winner of another VBSpam award.

Sophos Email Appliance

SC rate (total): 99.29%

SC rate (Project Honey Pot corpus): 99.53%

SC rate (VB spam corpus): 95.29%

SC rate (image spam): 98.97%

SC rate (large spam): 95.92%

FP rate: 0.16%

Final score: 98.81%

Sophos has been active in the anti-virus industry for a quarter of a century and, like most of its competitors, has been offering anti-spam solutions for quite some time too. *Sophos Email Appliance* is a hardware solution that filters inbound and, optionally, outbound email for spam and malware, as well as offering data protection and email encryption. These and other policies can be configured using a simple web interface, which I found easy to work with; the various reports and trends on email traffic will no doubt help experienced administrators to fine-tune the settings so that they work even better in their organizations.

But even without an administrator's intervention the product worked very well, achieving the third-highest spam catch rate with only a handful of false positives and one of the better final scores. If there was anything that needed to be improved, it would be the product's performance on large emails, but even here it caught the vast majority of spam.

SPAMfighter Mail Gateway

SC rate (total): 97.98%

SC rate (Project Honey Pot corpus): 98.34%

SC rate (VB spam corpus): 91.93%

SC rate (image spam): 98.04%

SC rate (large spam): 96.44%

FP rate: 0.49%

Final score: 96.51%

SPAMfighter Mail Gateway saw a slight improvement in its spam catch rate compared to the previous test, while its false positive rate was about the same; with such a performance the product easily wins another VBSpam award. It was good to see *SPAMfighter*'s performance on both large and image spam improve significantly, and hopefully next time the false positive rate will improve too: while this mostly concerned newsletters (arguably emails that are less likely to be missed by end-users), there is still room for some improvement here.

SpamTitan

SC rate (total): 99.26%

SC rate (Project Honey Pot corpus): 99.46%

SC rate (VB spam corpus): 95.92%

SC rate (image spam): 99.48%

SC rate (large spam): 98.49%

FP rate: 0.24%

Final score: 98.54%

Like many products this month, *SpamTitan* had a lower spam catch rate than in the previous test – when it caught more spam than any other product – but it also saw its false positive rate reduced. This resulted in another impressive final score, putting the product firmly in position as one of this month's top five performers.

Sunbelt VIPRE Email Security

SC rate (total): 98.41%

SC rate (Project Honey Pot corpus): 98.58%

SC rate (VB spam corpus): 95.51%

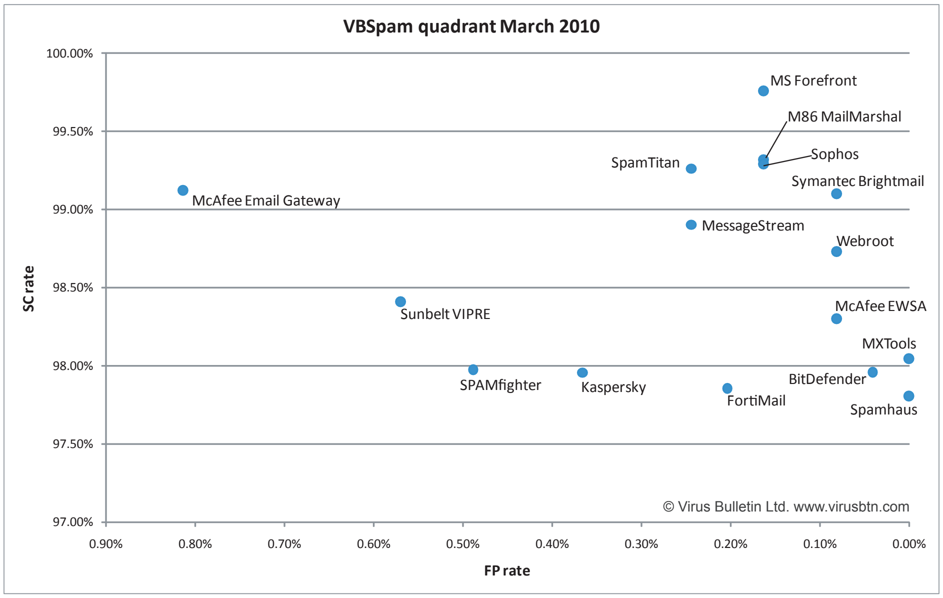
SC rate (image spam): 95.36%

SC rate (large spam): 95.15%

FP rate: 0.57%

Final score: 96.70%

Like many products in this test, *Sunbelt*'s *VIPRE* combines a slightly lower spam catch rate with a slightly lower false positive rate. The latter in particular still leaves some room for improvement, but it should also be noted that the product had a consistently high spam catch rate, even during periods when most other products saw their performance temporarily drop. This suggests that new spam campaigns are no problem for *VIPRE*.



Symantec Brightmail Gateway 9.0

SC rate (total): 99.10%

SC rate (Project Honey Pot corpus): 99.41%

SC rate (VB spam corpus): 93.89%

SC rate (image spam): 98.84%

SC rate (large spam): 96.29%

FP rate: 0.08%

Final score: 98.86%

Symantec Brightmail Gateway debuted in the previous test with an impressive performance and the third best final score. On this occasion we tested a new version of the product (a virtual appliance) which performed even better: like most products, its spam catch rate was slightly lower on this occasion, but this was more than made up for by the fact that it missed just two legitimate emails, resulting in the second best final score overall.

Webroot E-Mail Security SaaS

SC rate (total): 98.73%

SC rate (Project Honey Pot corpus): 98.86%

SC rate (VB spam corpus): 96.59%

SC rate (image spam): 99.47%

SC rate (large spam): 97.79%

FP rate: 0.08%

Final score: 98.49%

Webroot's hosted anti-spam solution has had a consistently high spam catch rate ever since joining the very first VBSpam test. In the past, the product has suffered from more false positives than average, but the developers must have worked hard on this and the product missed only two legitimate emails this time. If this is the reason fewer spam messages were caught, then I would say it's been worth it, as the product saw its final score improve.

Spamhaus ZEN plus DBL

SC rate (total): 97.81%

SC rate (Project Honey Pot corpus): 98.52%

SC rate (VB spam corpus): 85.89%

SC rate (image spam): 96.72%

SC rate (large spam): 92.91%

FP rate: 0.00%

Final score: 97.81%

As in the previous test, the IP address of every incoming email was checked against the *Spamhaus ZEN* DNS blacklist, while domain checks were performed against the new *Spamhaus DBL* blacklist. Once again, this resulted in a very good spam catch rate and again there were no false positives. While it is probably not a good idea to use a DNS blacklist as a standalone spam filter, with *Spamhaus* one can at least be sure that the vast majority of spam is blocked at an early stage.

Products ranked by final score	Final score
MS Forefront	99.28%
Symantec Brightmail	98.86%
M86 MailMarshal	98.84%
Sophos	98.81%
SpamTitan	98.54%
Webroot	98.49%
MessageStream	98.18%
McAfee EWSA	98.06%
MXTools	98.05%
BitDefender	97.84%
Spamhaus	97.81%
FortiMail	97.26%
Kaspersky	96.85%
Sunbelt VIPRE	96.70%
McAfee Email Gateway	96.69%
SPAMfighter	96.51%

CONCLUSION

For a few tests in a row we have been adding to the ham corpus the traffic of several email discussion lists. In the next test, we plan to take this one step further: we will use the emails sent to the lists, but rewrite the headers as well as the IP address and HELO/EHLO domain in such a way that, to the products in the test, it will look as if the emails have been sent directly to us rather than via the list server. This is not a trivial thing to do and certainly doesn't work for all mailing lists, but tests run over the past weeks show that it works well and that it creates a varied ham corpus.

We also plan to remove the *VB* corpora from the test. Over the past year our own email has given us a very realistic email stream to test against, but the fact that we have been unable to share full details of incorrectly classified emails with developers has become increasingly frustrating for all involved. Although developers have rarely questioned our decisions, in order for them to be able to improve their products – one of the most important aspects of the anti-spam tests – they need to have access to the full emails.

The products' performance on the *VB* spam and ham corpora will be included in the next report. However, these results will not count towards their final score.

The next test is set to run throughout April with the deadline for product submission being 26 March 2010; any developers interested in submitting a product should email martijn.grooten@virusbtn.com.