COMPARATIVE REVIEW

ANTI-SPAM COMPARATIVE REVIEW

Martijn Grooten

This month's anti-spam comparative review saw yet another increase in the field of competitors with 14 products taking their place on the test bench; the same 12 products that participated in the September test were joined by two new ones. One of the new products is the first anti-spam solution to take part in our test that runs on a virtual machine – demonstrating yet another possibility for administrators searching for a decent anti-spam solution to run in their organization. The 12 VBSpam awards given out this month – another record – demonstrate that there is plenty of choice when it comes to very good solutions.

THE TEST SET-UP

No changes were made to the test set-up, apart from some modifications to the corpora used, as is explained below. As usual, the full methodology can be found at http://www.virusbtn.com/vbspam/methodology/.

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. Those running on *Linux* ran on *SuSE Linux Enterprise Server 11*; the *Windows Server* products ran either the *2003* or the *2008* version, depending on which was recommended by the vendor.

THE EMAIL CORPUS

The test ran from 1pm UK time on 16 October 2009 to 12pm UK time on 30 October 2009 – with the end of British Summer Time coming in the middle of the test, this meant the test ran for two weeks exactly. The corpus contained a total of 199,842 emails: 2,121 ham messages and 197,721 spam messages. The latter consisted of 176,667 messages provided by Project Honey Pot and 21,054 spam messages sent to @virusbtn.com addresses.

The ham emails consisted of all legitimate emails sent to @virusbtn.com addresses. This time, however, some senders were excluded from the test set: these were the senders of emails that regularly discuss spam- and malware-related topics (for example anti-spam discussion lists) and as such *regularly* contain links to malicious and/or spamvertised URLs. We believe that not only are such emails unlikely to occur in the legitimate email stream of an average organization, but also that the recipients of such emails generally have the level of knowledge and technical ability required to whitelist these particular senders. All emails from these senders were removed from the test set, regardless of the contents of the individual emails. (Of course, it is possible that other legitimate senders also included malicious and/or spamvertised URLs in their emails – however, these were not excluded from the test set.)

Unsurprisingly, this affected the products' false positive rates and only one product blocked more than one per cent of all legitimate emails in the test. Interestingly, no legitimate email was blocked by more than four products – so while developers might argue that certain emails are hard to recognize as legitimate, it can also be pointed out that for every email they incorrectly blocked, there were at least ten other products that correctly recognized it as ham.

To make up for the exclusion of some senders, we subscribed some of our addresses to a number of email discussion lists. We believe this has several advantages: firstly, it adds to the variety of topics discussed in the ham stream, as well as to the variety of sending domains and IP addresses, and thus makes the test results more representative for an average company. Secondly, these emails are generally very much wanted by their recipients and as such do not fall in the grey area of legitimate-yet-not-particularly-wanted emails. And thirdly, because we can (and will) vary the lists subscribed to over time, we can give the full contents of the emails to developers whose products blocked them – in doing so neither compromising our own confidentiality nor introducing the possibility for developers to whitelist these senders and thus gain unfair advantage over their competitors. Finally, it should be noted that spam is occasionally sent to discussion lists – for instance when a subscriber's email account has been compromised. This happened once during the running of the test and this email was classified as spam.

RESULTS

In previous reviews we have published both the overall false positive (FP) rate and the false positive rate as a ratio of the total *VB* mail stream – the latter number is of little practical use, but has been included in the past for reference. However, because of the modifications described above, the mail corpora used are not those of a real company and therefore we have decided to leave this FP ratio out of the report; interested readers will still be able to compute the ratio themselves.

BitDefender Security for Mail Servers 3.0.2

SC rate (total): 97.89%

SC rate (Project Honey Pot corpus): 98.90%

SC rate (VB spam corpus): 89.37%

FP rate: 0.707%

Two interesting papers presented at VB2009 demonstrated that *BitDefender* does more than simply use existing technologies to fight spam: the developers in the company's Bucharest-based anti-spam lab are working hard to find new ways to stay ahead of the spammers. The product has won a VBSpam award in each of the three previous anti-spam tests and while this month the spam catch rate is slightly lower than that of the previous test, it is still sufficient for the product – again, the *Linux* version – to win a VBSpam Gold award.

(Note: In the previous test report it was stated that *BitDefender* had 11 false positives. Careful investigation of these showed that a mistake was made and one reported false positive should not have been counted as such. This did not affect the level of the award earned by the product.)

Fortinet FortiMail

SC rate (total): 98.47%

SC rate (Project Honey Pot corpus): 98.98%

SC rate (VB spam corpus): 94.12%

FP rate: 0.047%

FortiMail, a hardware appliance from Canadian company *Fortinet*, won a VBSpam Silver award in the two previous tests and while not entirely unhappy with that, its developers believed the product was capable of doing better. For this test, the product's spam criteria were loosened in an attempt to reduce the false positive rate (which, so far, has prevented it from winning a higher level award), while an upgrade of the firmware was intended to help maintain a high spam catch rate. The latter worked very well, but even more impressive was the product's low false positive rate: out of well over 2,000 emails, only one newsletter was missed. A VBSpam Platinum award is well deserved and the developers' faith in their product fully justified.

Kaspersky Anti-Spam 3.0

SC rate (total): 97.52%

SC rate (Project Honey Pot corpus): 98.58%

SC rate (VB spam corpus): 88.65%

FP rate: 0.141%

In previous reports I have lauded *Kaspersky*'s anti-spam solution for the minimal maintenance it requires: it is installed on a *Linux* machine and works straight away. Of course ‘works’ doesn’t necessarily mean ‘works well’, but it does in the case of *Kaspersky*. Particularly impressive is the product’s consistently low false positive rate – only three emails were incorrectly blocked during the test. This combined with a good spam catch rate earns the product yet another VBSpam Gold award.

McAfee Email Gateway (formerly IronMail)

SC rate (total): 99.02%

SC rate (Project Honey Pot corpus): 99.85%

SC rate (VB spam corpus): 92.00%

FP rate: 0.707%

Like last time, *McAfee*’s *Email Gateway* appliance (also sold under its former name *IronMail*) was the only product that scanned and, in cases of suspected spam, blocked emails during the SMTP transaction, with only the harder-to-filter emails being scanned at a later stage. This solution worked well: the product once again had a very high spam catch rate. The false positive rate was significantly lower than on the last occasion and all but a few of these false positives were scanned at a later stage; in a real scenario these emails would probably have been stored in quarantine rather than being discarded altogether. With still a few too many false positives for a platinum award, the product won its second consecutive VBSpam Gold award.

McAfee Email and Web Security Appliance

SC rate (total): 98.75%

SC rate (Project Honey Pot corpus): 99.28%

SC rate (VB spam corpus): 94.36%

FP rate: 0.189%

‘Never change a winning formula’, they must have thought at *McAfee* and in a system administrator’s ideal scenario the appliance – the only product to win a VBSpam Platinum award in the last test – was run using exactly the same set-up. This scenario worked well for the product and combining a very low false positive rate with a very high spam catch rate, it won its second consecutive VBSpam Platinum award.

M86 MailMarshal SMTP

SC rate (total): 99.62%

SC rate (Project Honey Pot corpus): 99.94%

SC rate (VB spam corpus): 96.92%

FP rate: 0.519%

The brand *M86 Security* has been around in the world of computer security for barely two months; before that the company was known as *Marshal8e6*, which in turn was the merger of *Marshal* and *8e6*. The company offers a number of security solutions including its *MailMarshal SMTP* spam filter.

This product, which comes with its own MTA and was run on *Windows Server 2003*, uses a multi-layered approach where an email has to pass several tests before it is sent to the user’s inbox. Among these tests are SpamBotCensor, which uses knowledge about the engines used by various spam bots to detect spammers at the SMTP level, and SpamCensor, which uses heuristics to block spam based on the contents of the email. The product’s user interface gives the administrator plenty of opportunities to modify the rules for the various tests and can easily be fine-tuned to meet the needs of a particular organization.

Unfortunately, the SpamBotCensor could not be applied during our test, but *MailMarshal* still had the highest spam catch rate of all participating products. Combined with a low false positive rate, it just missed out on a platinum-level award; a VBSpam Gold award nevertheless marks an excellent debut for *MailMarshal*.

MessageStream

SC rate (total): 99.49%

SC rate (Project Honey Pot corpus): 99.82%

SC rate (VB spam corpus): 96.64%

FP rate: 0.471%

One reason why organizations may want to choose a hosted anti-spam solution is the little maintenance it requires. That is certainly the case with *MessageStream*, the hosted solution provided by *Giacom*. Without a lot of intervention from the developers it achieved yet another very high spam catch rate and missed out on a platinum award by just a few emails; it is the only product to have won four VBSpam Gold awards in a row.

Messaging Architects M+Guardian

SC rate (total): 98.75%

SC rate (Project Honey Pot corpus): 99.26%

SC rate (VB spam corpus): 94.47%

FP rate: 0.943%

It is always disappointing to see a product win a lower-level award in a test than in the previous one. In reality, the *M+Guardian* appliance performed better on this occasion than in the last test – however, since the thresholds have become stricter the product's fourth VBSpam award is a silver one. It will be interesting to see whether the product will be able to do better again next time around.

Microsoft Forefront Protection 2010 for Exchange Server

SC rate (total): 99.00%

SC rate (Project Honey Pot corpus): 99.46%

SC rate (VB spam corpus): 95.16%

FP rate: 0.471%

Few will have been awaiting this review more eagerly than the developers at *Microsoft*: their *Forefront* product won a VBSpam Silver award in its first test in September. At the time the product was still a release candidate, and in the weeks following that test they believed some issues had been solved – thus they were eager to see if the changes had made an improvement. They had: the product's false positive rate was reduced by almost four-fifths compared to the last test, while it maintained a high spam catch rate. A VBSpam Gold award will be an extra reason to celebrate the official release of the product in the second week of November.

Sanesecurity signatures for ClamAV

SC rate (total): 72.40%

SC rate (Project Honey Pot corpus): 73.24%

SC rate (VB spam corpus): 65.34%

FP rate: 0.33%

In previous reviews it has not been made clear enough that while the *Sanesecurity* signatures work together with *ClamAV*, they have little to do with that product (which is mainly an anti-malware product). Perhaps unsurprisingly for something that scans emails purely based on content, this product sees a greater fluctuation from day to day than other products; in this case it means that some 'bad days' in the first week of the test caused the product's final spam catch rate to be significantly lower than during the previous test. Still, for what is only a partial solution – which would be an effective part of a multi-layered solution – a spam catch rate of well over 70% is a rather good score, although a number of false positives caused by incorrectly blacklisted URLs demonstrate that the product isn't entirely without fault either.

SPAMfighter Mail Gateway

SC rate (total): 97.22%

SC rate (Project Honey Pot corpus): 97.36%

SC rate (VB spam corpus): 96.10%

FP rate: 0.66%

SPAMfighter's *Mail Gateway* debuted in the previous VBSpam test, but failed to win an award. The developers at the Danish company believed this may have been the result of the product being set up in a manner that was less than ideal for our test; they also believed their product might have been disproportionately disadvantaged by issues with the network. While these issues were solved, the product was set to filter less stringently to reduce the number of false positives, while at the same time the linger filter was turned on. This filter will hold on to emails that aren't immediately recognized as either ham or spam and rescan them after a certain amount of time, by which time the content might be recognized by the updated spam filter. Of course, this may cause delays for legitimate email, but the filter can be set to work only at certain times of day (such as outside office hours), when delays aren't generally noted; in this test it was turned on 24 hours a day.

The changes certainly had a very positive effect on the product's performance: the false positive rate was reduced greatly and the spam catch rate was still rather good; the product performed almost equally well on both spam corpora, showing that its performance wasn't just luck. A VBSpam Gold award is well deserved.

SpamTitan

SC rate (total): 99.48%

SC rate (Project Honey Pot corpus): 99.97%

SC rate (VB spam corpus): 95.41%

FP rate: 0.377%

Spam filters are essential for any organization, but for smaller companies buying separate hardware for spam filtering might not always be an option. Running the filter on a virtual machine could then be a solution and *SpamTitan*, a company based on the Irish west coast, offers such a solution. The product can easily be installed under *VMware* – for larger organizations, the same product is available as an ISO image that contains a complete operating system – and works almost immediately after installing. That is not to say the spam rules cannot be customized to suit a particular organization's needs: a web interface lets the user customize many rules of the blended approach the product uses to fight spam. I was particularly charmed by the simple, yet accurate explanations of the various anti-spam rules.

The fact that this approach worked well to block spam can be seen from the spam catch rate – which was among the highest in this test. At the same time, the product had a very low false positive rate, missing out on a platinum award by just a single email; a VBSpam Gold award is more than deserved.

Vircom modusGate

SC rate (total): 94.01%

SC rate (Project Honey Pot corpus): 94.37%

SC rate (VB spam corpus): 90.92%

FP rate: 3.772%

Vircom's *modusGate* product has failed to win an award in the last two VBSpam tests, but its developers are working hard to fix the issues that they believe are the cause of the poor performance in our tests. Still, with a false positive rate of more than three per cent and a spam catch rate significantly lower than that of most of its competitors, we cannot but deny *Vircom*'s *modusGate* a VBSpam award on this occasion.

Webroot E-Mail Security SaaS

SC rate (total): 99.31%

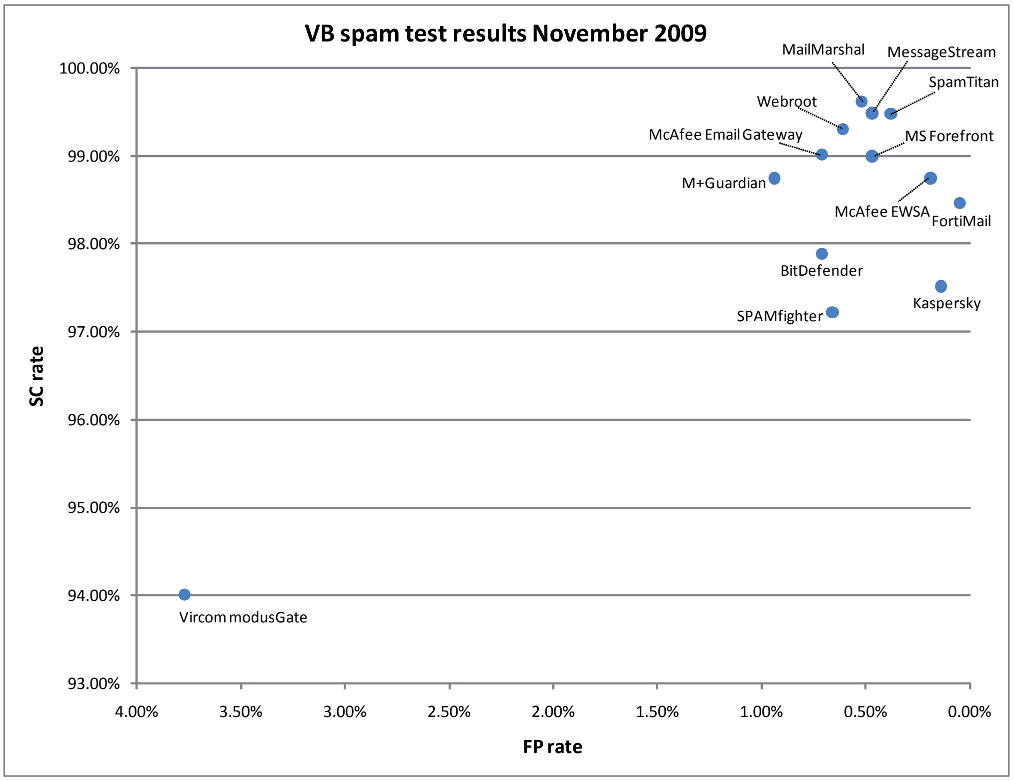
SC rate (Project Honey Pot corpus): 99.67%

SC rate (VB spam corpus): 96.31%

FP rate: 0.613%

Webroot only just missed out on a VBSpam Gold award in the last round of testing, winning its second VBSpam Silver award instead. Making small improvements is not a trivial task though, especially if competitors do the same thing and the thresholds thus become stricter. However, the developers of this hosted solution managed to improve their product enough to see the number of false positives reduced, while still having among the highest spam catch rates and thus this time *Webroot* earns a VBSpam Gold award.

				Total spam			Project Honey Pot spam			VB corpus		
	True negative	FP	FP rate	False negative	True positive	SC rate	False negative	True positive	SC rate	False negative	True positive	SC rate
BitDefender	2,106	15	0.71%	4,172	193,549	97.89%	1,935	174,732	98.90%	2,237	18,817	89.37%
FortiMail	2,120	1	0.05%	3,033	194,688	98.47%	1,794	174,873	98.98%	1,239	19,815	94.12%
Kaspersky	2,118	3	0.14%	4,904	192,817	97.52%	2,515	174,152	98.58%	2,389	18,665	88.65%
McAfee Email Gateway	2,106	15	0.71%	1,941	195,780	99.02%	257	176,410	99.85%	1,684	19,370	92.00%
McAfee EWSA	2,117	4	0.19%	2,466	195,255	98.75%	1,278	175,389	99.28%	1,188	19,866	94.36%
MailMarshal	2,110	11	0.52%	752	196,969	99.62%	103	176,564	99.94%	649	20,405	96.92%
MessageStream	2,111	10	0.47%	1,017	196,704	99.49%	310	176,357	99.82%	707	20,347	96.64%
M+Guardian	2,101	20	0.94%	2,472	195,249	98.75%	1,307	175,360	99.26%	1,165	19,889	94.47%
MS Forefront	2,111	10	0.47%	1,975	195,746	99.00%	955	175,712	99.46%	1,020	20,034	95.16%
Sanesecurity	2,114	7	0.33%	54,567	143,154	72.40%	47,269	129,398	73.24%	7,298	13,756	65.34%
SPAMfighter	2,107	14	0.66%	5,488	192,233	97.22%	4,667	172,000	97.36%	821	20,233	96.10%
SpamTitan	2,113	8	0.38%	1,025	196,696	99.48%	59	176,608	99.97%	966	20,088	95.41%
Vircom modusGate	2,041	80	3.77%	11,851	185,870	94.01%	9,940	166,727	94.37%	1,911	19,143	90.92%
Webroot	2,108	13	0.61%	1,358	196,363	99.31%	581	176,086	99.67%	777	20,277	96.31%



AWARDS

As in the previous test, the levels of the awards earned by products are defined as follows:

- VBSpam Platinum for products with a total spam catch rate twice as high and a false positive rate twice as low as the average in the test.
- VBSpam Gold for products with a total spam catch rate at least as high and a false positive rate at least as low as the average in the test.
- VBSpam Silver for products whose total spam catch rate and false positive rates are no more than 50% worse than the average in the test.

To avoid the averages being skewed by one or more malperforming products, the scores for any product with a false positive rate of more than 10% and/or a spam catch rate of less than 70% are removed from the computation of the averages; this did not apply to any of the products this month.

This month's benchmarks are then as follows:

- Platinum: SC 98.25%; FP 0.36%
- Gold: SC 95.60%; FP 0.71%
- Silver: SC 94.75%; FP 1.07%

The table shows the scores for all of the products on test. The highlighted columns show the scores used for the benchmark calculations. In the graph, *SaneSecurity* has been left out: this is only a partial solution and, as such, should not be compared directly with the other products.

CONCLUSION

The period between tests is used by developers to make improvements to their products. At the same time, we use this period to make improvements to the test set-up and to review our methodology. With the catch rates and (especially) the false positive rates of the various products edging closer to each other than ever, we believe that the way in which the product certifications are determined could do with some improvements. These changes will be announced in due course (well before the start of the next test) at http://www.virusbtn.com/vbspam.

The next test is set to run throughout December and the deadline for product submission will be 27 November 2009; any developers interested in submitting a product should email martijn.grooten@virusbtn.com. A number of new products have already committed to their participation and we are looking forward to an even bigger test.

December has traditionally been the month when spam levels rise to unprecedented heights, so it will be interesting to see which products are best at keeping their users' inboxes clean during the holiday period.