Anti-Virus Comparative

Retrospective test

(Static detection of new/unknown malicious software)

Language: English
August 2011
Last revision: 15th November 2011

www.av-comparatives.org

Content

1. Introduction

This test report is the second part of the August 2011 test[1]. The report is delivered in late November due the high-required work, deeper analysis and preparation of the retrospective test-set. Due to the time spent in analyzing and assuring the quality of those samples, they usually also get included in the next detection test to see if they will be covered by then.

Many new viruses and other types of malware appear every day, this is why it's important that Anti-Virus products not only provide new updates, as often and as fast as possible, but also that they are able to detect such threats in advance (also without executing them or while offline) with generic and/or heuristic techniques. Even if nowadays most Anti-Virus products provide daily, hourly or cloud updates, without heuristic/generic methods there is always a time-frame where the user is not reliably protected.

The products used the same updates and signatures they had the 12th **August** 2011, and the same detection settings as used in August (see page 6 of this report). This test shows the proactive detection capabilities that the products had at that time. We used new malware appeared between the 13th and 20th August 2011. The following products were tested[2]:

- avast! Free Antivirus 6.0
- AVIRA AntiVir Personal 10.2
- BitDefender Anti-Virus Plus 2012
- eScan Anti-Virus 11.0
- ESET NOD32 Antivirus 5.0
- F-Secure Anti-Virus 2011
- G DATA AntiVirus 2012
- Kaspersky Anti-Virus 2012
- Microsoft Security Essentials 2.1
- Panda Cloud Antivirus 1.5
- Qihoo 360 Antivirus 2.0
- Trustport Antivirus 2012

2. Description

Anti-Virus products often claim to have high proactive detection capabilities – far higher than those reached in this test. This is not just a self-promotional statement; it is possible that products reach the stated percentages, but this depends on the duration of the test-period, the size of the sample set and the used samples. The data shows how good the proactive (on-access/on-demand) detection capabilities of the scanners were in detecting the new threats (sometimes also named as 0-day threats by others) used in this test. Users should not be afraid if products have, in a retrospective test, low percentages. If the anti-virus software is always kept up-to-date, it may be able to detect more samples. For understanding how the detection rates of the Anti-Virus products look with updated signatures and programs, have a look at our regular on-demand detection tests. By design and scope of the test, only the heuristic/generic detection capability was tested (offline). Some products may be had the ability to detect some samples e.g. on-execution or by other monitoring tools, like behaviour-blocker, reputation/cloud heuristics, etc. Those kinds of additional protection technologies are considered by AV-Comparatives in e.g. whole-product dynamic tests, but are outside the scope of retrospective tests.

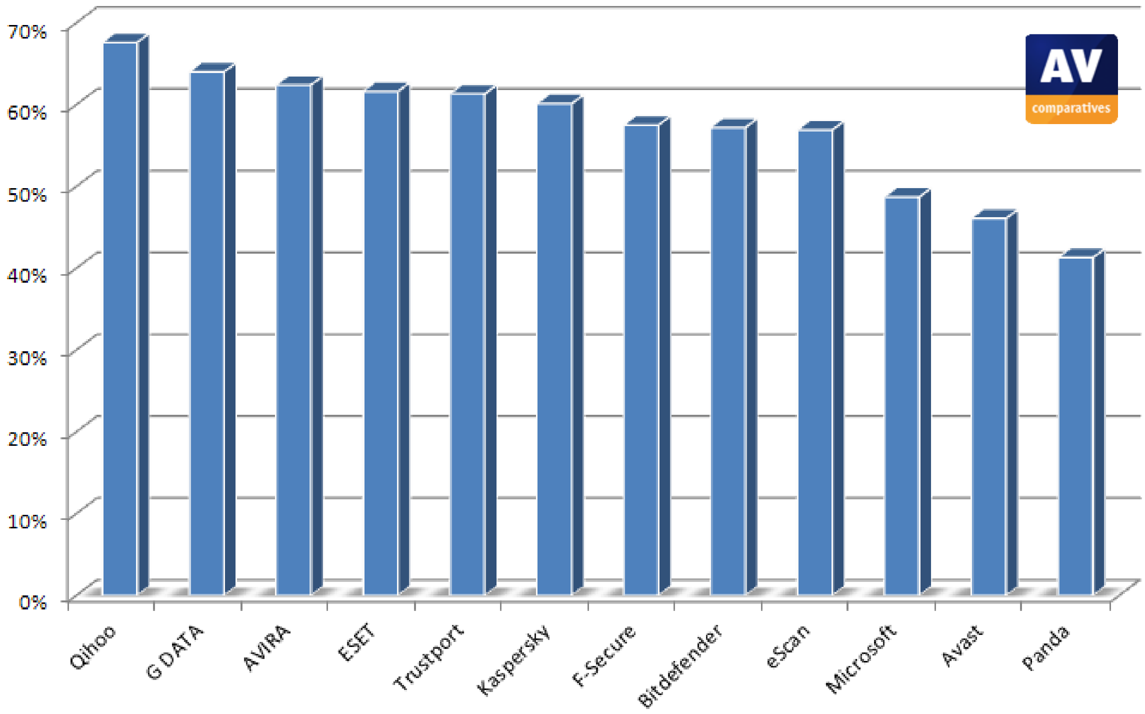
[1] http://www.av-comparatives.org/images/stories/test/ondret/avc_od_aug2011.pdf

[2] AVG, K7, McAfee, PC Tools, Sophos, Symantec, Trend Micro and Webroot decided to not get included in this test and to renounce to get awarded

3. Test Results

Note: If you are going to republish those results, it is compulsory to include a comment that products use also additional protection features (like behavior-blockers, etc.) to protect against completely new/unknown malware. As described on previous and next pages, this test evaluates only the offline heuristic/generic detection of the products against unknown/new malware, without the need to execute it or to submit anything online.

The below table shows the proactive on-demand detection capabilities of the various products, sorted by detection rate. The given awards (see page 8 of this report) are based not only on the detection rates over the new malware, but also considering the false alarm rates.



As it can be seen above, the tested products are able to detect a quantity of completely new/unknown malware proactively even without executing the malware, using passive heuristics, while other protective mechanisms like HIPS, behavior analysis and behavior-blockers, reputation/cloud heuristics, etc. add an extra layer of protection. The retrospective test is performed using passive scanning and demonstrates the ability of the products under test to detect new malware proactively, without being executed. In retrospective tests „in-the-cloud" features are not considered, as that is not the scope of the test.

This test does not include some vendor's products who decided to do not be included in this "proactive/retrospective" test, e.g. because in their opinion their product's real-life capabilities are not adequately represented in the retrospective test due to the absence of a live Internet connection or because URL blocking is not considered. The methodology and design of our "proactive/retrospective" testing indeed does not allow cloud-based products to connect to their remote knowledge bases and we also do not consider URL blocking, as this is not what we want to measure/compare in this type of test. Several other included products also have cloud-based technologies (and some don't), but at the same time they still provide good offline generic/heuristic detections, without having to rely on / sent data to their clouds, without having many false alarms and without being dependent of the malware vector (i.e. without relying on blacklists of URL filters). Cloud and other technologies should be seen as an additional protection enhancement, but never as a replacement of basic technologies.

Some further (unofficial) reasons given by some vendors for not taking part in retrospective tests were e.g. that they know that they score low in this type of tests and do not want that users see tests where their results are very low compared to others and not close to 100%. Although the given technical and marketing reasons may appear to make sense, users should have the right to know how products score in various aspects and various test scenarios; as long as the users are informed/educated about what the results are showing, they will be able to understand by themselves to what extend the data is useful for their needs, and if it is not of interest for them, the users will look at results provided in other types of tests provided by AV-Comparatives, like e.g. the Whole-Product-Dynamic test, which aims to simulate a real-world scenario which takes into account various protection features of the products.

Nowadays, hardly any Anti-Virus products rely purely on "simple" signatures anymore. They all use complex generic signatures, heuristics etc. in order to catch new malware, without needing to download signatures or initiate manual analysis of new threats. In addition, Anti-Virus vendors continue to deliver signatures and updates to fill the gaps where proactive mechanisms initially fail to detect some threats. Anti-Virus software uses various technologies to protect a PC. The combination of such multi-layered protection usually can provide good protection.

Almost all products run nowadays by default with highest protection settings (at least either at the entry points, during whole computer on-demand scans or scheduled scans) or switch automatically to highest settings in case of a detected infection. Due to that, in order to get comparable results, we tested all products with highest settings, if not explicitly advised otherwise by the vendors. To avoid some frequent questions, below are some notes about the used settings (scan of all files etc. is always enabled) of some products:

AVIRA, Kaspersky: asked to get tested with heuristic set to high/advanced. Due to that, we recommend users to consider also setting the heuristics to high/advanced.
F-Secure: asked to get tested and awarded based on their default settings (i.e. without using their advanced heuristics).
AVIRA: asked to do not enable/consider the informational warnings of packers as detections. Due to that, we did not count them as detections (neither on the malware set, nor on the clean set).

AV-Comparatives prefer to test with default settings. As most products run with highest settings by default (or switch to highest automatically when malware is found, making it impossible to test against various malware with "default" settings), in order to get comparable results we set also the few remaining products to highest settings (or leave them to default settings) in accordance with the respective vendors. We hope that all vendors will find the appropriate balance of detection/false alarms/system impact and will provide highest security already by default and remove paranoid settings inside the user interface which are too high to be ever of any benefit for normal users.

This time we included in the retrospective test-set only new malware which has been seen in-the-field and prevalent in the few days after the last update in August. Additionally, we took care to include malware samples which belong to different clusters (i.e. which differ from each other, in order to e.g. do not include too many samples which are practically the same malware). As malware which became prevalent may be spotted faster by reactive measures when many users got infected, initial proactive rates may be lower (because if they would have been spotted proactively, they may not become prevalent if they would be blocked/detected in advance).

4. Summary results

The results show the proactive (generic/heuristic) file detection[3] capabilities of the scan engines against new malware. The percentages are rounded to the nearest whole number. Do not take the results as an absolute assessment of quality - they just give an idea of who detected more, and who less, in this specific test. To know how these anti-virus products perform with updated signatures, please have a look at our detections tests of February and August. To know about protection rates provided by the various products, please have a look to our ongoing Whole-Product Dynamic tests. Readers should look at the results and build an opinion based on their needs.

Below you can see the proactive on-demand detection results over our set of new and prevalent malware appeared in-the-field within some few days of August (9003 different malware samples):

ProActive detection of new malware:

1.	**Qihoo**	**67.6%**
2.	**G DATA**	**64.0%**
3.	**AVIRA**	**62.4%**
4.	**ESET**	**61.6%**
5.	**Trustport**	**61.3%**
6.	**Kaspersky**	**60.1%**
7.	**F-Secure**	**57.5%**
8.	**Bitdefender**	**57.2%**
9.	**eScan**	**56.9%**
10.	**Microsoft**	**48.7%**
11.	**Avast**	**46.1%**
12.	**Panda**	**41.4%**

5. False positive/alarm test

To better evaluate the quality of the detection capabilities, the false alarm rate has to be taken into account too. A false alarm (or false positive)[4] is when an Anti-Virus product flags an innocent file to be infected when it is not. False alarms can sometimes cause as much troubles like a real infection.
The false alarm test results were already included in the test report of August. For details, please read the report available at http://www.av-comparatives.org/images/stories/test/fp/avc_fp_aug2011.pdf

Very few false alarms (0-3):	Kaspersky, Microsoft, Panda, ESET
Few false alarms (4-15):	F-Secure, Bitdefender, Avast, AVIRA, G DATA
Many false alarms (over 15):	Qihoo, eScan, Trustport

[3] This test is performed offline and on-demand – it is NOT an on-execution/behavioral/cloud test.
[4] All discovered false alarms were already reported to the vendors in August and are now already fixed.

6. Certification levels reached in this test

The following certification levels are for the results reached[5] in the retrospective test:

CERTIFICATION LEVELS	PRODUCTS
AV comparatives ADVANCED+ ★★★ RETROSPECTIVE / PROACTIVE TEST NOV 2011	G DATA AVIRA ESET Kaspersky F-Secure Bitdefender
AV comparatives ADVANCED ★★ RETROSPECTIVE / PROACTIVE TEST NOV 2011	Qihoo* TrustPort* eScan* Microsoft Avast Panda
AV comparatives STANDARD ★ RETROSPECTIVE / PROACTIVE TEST NOV 2011	-
NOT INCLUDED[6]	*AVG, K7, McAfee, PC Tools, Sophos, Symantec, Trend Micro, Webroot*

*: Products with "many" false alarms were rated according to the below award system[7]:

	Proactive Detection Rates			
	0-10%	**10-25%**	**25-50%**	**50-100%**
None - Few FP	**tested**	**STANDARD**	**ADVANCED**	**ADVANCED+**
Many FP	**tested**	**tested**	**STANDARD**	**ADVANCED**
Very many FP	**tested**	**tested**	**tested**	**STANDARD**
Crazy many FP	**tested**	**tested**	**tested**	**tested**

[5] All the products that were included in the test achieved good results, and received either the Advanced or Advanced+ award. Remarkably, Panda's Cloud AntiVirus product achieved respectable detection rates of unknown malicious programs, despite not allowed to use the cloud, and was rewarded with an Advanced award. Unfortunately, not all vendors chose to participate in this test. This may be because many of the non-participating programs would only achieve sub-optimal results in this type of test which does not make use of the cloud etc.

[6] As those products are included in our yearly public test-series, they are listed even if those vendors decided to do not get included (read more on page 4 and 5 of this report).

[7] Considering that certain vendors did not take part, we decided that it makes more sense in this case to keep our fixed thresholds instead of using the cluster method (as by the non-inclusion of the low-scoring products clusters may be built "unfairly").

7. Copyright and Disclaimer

This publication is Copyright © 2011 by AV-Comparatives e.V. ®. Any use of the results, etc. in whole or in part, is ONLY permitted after the explicit written agreement of the management board of AV-Comparatives e.V., prior to any publication. AV-Comparatives e.V. and its testers cannot be held liable for any damage or loss, which might occur as result of, or in connection with, the use of the information provided in this paper. We take every possible care to ensure the correctness of the basic data, but no representative of AV-Comparatives e.V. can he held liable for the accuracy of the test results. We do not give any guarantee of the correctness, completeness, or suitability for a specific purpose of any of the information/content provided at any given time. No one else involved in creating, producing or delivering test results shall be liable for any indirect, special or consequential damage, or loss of profits, arising out of, or related to, the use or inability to use, the services provided by the website, test documents or any related data. AV-Comparatives e.V. is an Austrian Non-Profit Organization.

AV-Comparatives e.V. (November 2011)

validEDGE
MIS 1100