Anti-Virus Comparative

On-demand Detection of Malicious Software

includes false alarm test

Language: English
March 2012

Last Revision: 10th April 2012

www.av-comparatives.org

Table of Contents

Tested Products	3
Conditions for participation and test methodology	4
Tested product versions	4
Comments	5
Graph of missed samples	6
Results	8
False positive/alarm test	9
Award levels reached in this test	10
Copyright and Disclaimer	11

Tested Products

- AhnLab V3 Internet Security 8.0
- avast! Free Antivirus 7.0
- AVG Anti-Virus 2012
- AVIRA Antivirus Premium 2012
- BitDefender Antivirus Plus 2012
- BullGuard Antivirus 12
- eScan Anti-Virus 11.0
- ESET NOD32 Antivirus 5.0
- F-Secure Anti-Virus 2012
- Fortinet FortiClient Lite 4.3
- G DATA AntiVirus 2012
- GFI Vipre Antivirus 2012
- Kaspersky Anti-Virus 2012
- McAfee AntiVirus Plus 2012
- Microsoft Security Essentials 2.1
- Panda Cloud Free Antivirus 1.5.2
- PC Tools Spyware Doctor with AV 9.0
- Sophos Anti-Virus 10.0
- Trend Micro Titanium AntiVirus+ 2012
- Webroot SecureAnywhere AV 8.0

Conditions for participation and test methodology

The conditions for participation in our tests are listed in the methodology document at http://www.av-comparatives.org/seiten/ergebnisse/methodology.pdf. Before proceeding with this report, readers are advised to first read the above-mentioned document.

The participation is limited to not more than 20 international well-known Anti-Virus products, which vendors agreed to get tested and included in the public test-series of 2012.

Tested Product Versions

The Malware sets have been frozen the 16th February 2012 and consisted of 291388 sample variants. The products were updated on the 1st March 2012. The following twenty up-to-date products were included in this public test:

- AhnLab V3 Internet Security 8.0.5.19
- avast! Free Antivirus 7.0.1407
- AVG Anti-Virus 2012.0.1913
- AVIRA Antivirus Premium 12.0.0.915
- Bitdefender Anti-Virus+ 15.0.36.1530
- BullGuard Antivirus 12.0.215
- eScan Anti-Virus 11.0.1139.1146
- ESET NOD32 Antivirus 5.0.95.0
- F-Secure Anti-Virus 12.49.104
- Fortinet FortiClient Lite 4.3.3.0436
- G DATA AntiVirus 22.1.0.2
- GFI Vipre Antivirus 5.0.5134
- Kaspersky Anti-Virus 12.0.0.374
- McAfee AntiVirus Plus 11.0.654
- Microsoft Security Essentials 2.1.1116.0
- Panda Cloud Free Antivirus 1.5.2
- PC Tools Spyware Doctor with Antivirus 9.0.0.909
- Sophos Anti-Virus 10.0
- Trend Micro Titanium AntiVirus Plus 5.0.1280
- Webroot SecureAnywhere 8.0.1.95

Please try the products[1] on your own system before making a purchase decision based on these tests. There are also some other program features and important factors (e.g. price, ease of use/management, compatibility, graphical user interface, language, support, etc.) to consider. Although very important, the file detection rate of a product is only one aspect of a complete Anti-Virus product. AV-Comparatives provides also a whole product dynamic "real-world" protection test, as well as other test reports which cover different aspects/features of the products. Please visit our website to find the other tests and reports.

[1] Information about used additional third-party engines/signatures inside the products: **Bullguard**, **eScan** and **F-Secure** are based on the BitDefender engine. **G DATA** is based on the Avast and Bitdefender engines. **PC Tools** is using the signatures of Symantec.

Most products run with highest settings by default. Certain products switch to highest settings automatically when malware is found. This makes it impossible to test against various malware with real "default" settings. In order to get comparable results we set the few remaining products to highest settings or leave them to lower settings - in accordance with the respective vendors. We kindly ask vendors to provide stronger settings by default, i.e. set their default settings to highest levels of detection, esp. for scheduled scans or scans initiated by the user. This is usually already the case for on-access scans and/or on-execution scans. We allow the remaining vendors (which do not use highest settings in on-demand scans) to choose to be tested with higher setting as they e.g. use in on-access/on-execution higher settings by default. So the results of the file detection test are closer to the usage in the field. We ask vendors to remove paranoid settings inside the user interface which are too high to be ever of any benefit for common users. Below are some notes about the settings used (scan all files, scan archives, etc. is being enabled), e.g.:

AVG, AVIRA: asked to do not enable/consider the informational warnings of packers as detections. So, we did not count them as detections (neither on the malware set, nor on the clean set).

Avast, AVIRA, Kaspersky: tested with heuristic set to high/advanced.

F-Secure, Sophos: asked to get tested and awarded based on their default settings (i.e. without using their advanced heuristics / suspicious detections setting).

Several products make use of cloud technologies, which require an active internet connection. Our tests are performed using an active internet connection. We do not longer show the baseline detection rates without cloud and show instead only the results with active cloud. Users should be aware that detection rates may be in some cases drastically lower if the scan is performed while offline (or when the cloud is unreachable for various reasons). The cloud should be considered as an additional benefit/feature to increase detection rates (as well as response times and false alarm suppression) and not as a full replacement for local offline detections. Vendors should make sure that users are warned in case that the connectivity to the cloud[2] gets lost e.g. during a scan, which may affect considerably the provided protection and make e.g. an initiated scan useless.

This report does no longer contain the splitted detection categories. As announced last year we also do not longer provide the on-demand scanning speed test. We invite users also to look at our other tests and not only this type of test. This is to make users aware of other types of tests that we are providing since some years, like e.g. our Whole-Product "Real-World" Protection Test.

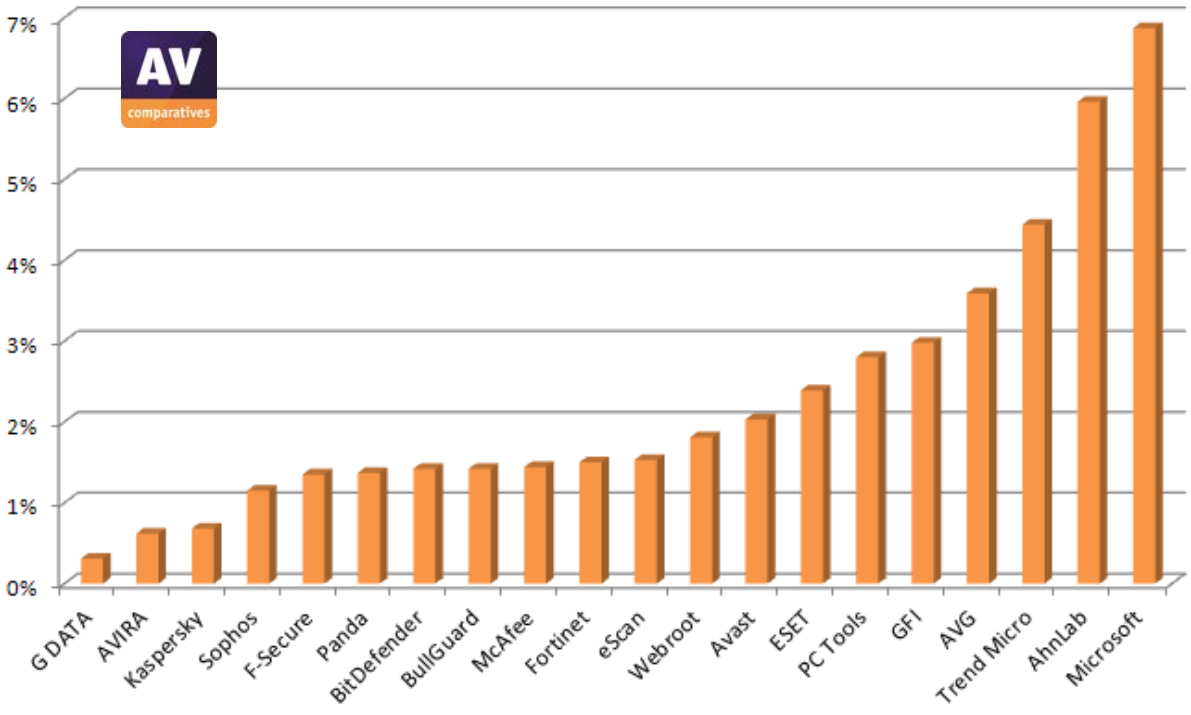
The used test-set has been built consulting telemetry data in attempt to include prevalent samples from the last months which are/were hitting users in the field. Due to the increasing amount of polymorphic malware, we applied a clustering method to classify similar files. This allows us to evaluate prevalent polymorphic samples and reducing the size of the set without introducing bias. Therefore, each miss represents one missed group/variant of files. We decided to apply this method as it helps reducing influence of e.g. inappropriate samples of same family in the set. We compared the results/rankings with and without clustering the files. Using fewer samples (one per variant) will reduce the workload for all in the next tests, as a smaller set can be analyzed better.

[2] A good paper about the pro and contra of clouds can be found here: http://www.av-test.org/fileadmin/pdf/publications/vb_2009_avtest_paper_why_in-the-cloud_scanning_is_not_a_solution.pdf

The malware detection rates are grouped by the testers after looking at the clusters build with the hierarchal clustering method. By using clusters, there are no fixed thresholds to reach, as the thresholds change based on the results. The testers may group the clusters rationally and not rely solely on the clusters, to avoid that if e.g. all products would in future score badly, they do not get high rankings anyway.

	Detection Rate Clusters/Groups (given by the testers after consulting statistical methods)			
	4	**3**	**2**	**1**
Very few *(0-2 FP's)* **Few** (3-15 FP's)	**TESTED**	**STANDARD**	**ADVANCED**	**ADVANCED+**
Many (16-100 FP's)	**TESTED**	**TESTED**	**STANDARD**	**ADVANCED**
Very many (101-500 FP's)	**TESTED**	**TESTED**	**STANDARD**	**STANDARD**
Crazy many (over 500 FP's)	**TESTED**	**TESTED**	**TESTED**	**TESTED**

Graph of missed samples (lower is better)



Files have been clustered. Each miss represents a missed variant. All tested products scored quite good and missed less than 10% of the test-set, so we decided to cluster the detection rates into three groups.

The results of our file detection rate tests are not applicable for on-execution protection technologies (like HIPS, behaviour blockers, etc.). Such technologies are evaluated in our new heuristic/behavioral tests which we will provide during this year. Please do not miss the second part of the report (it will be published in a few months) containing the new heuristic/behavioral test. It evaluates how well products are at detecting and blocking completely new/unknown malware (by on-demand/on-access scanner with local heuristic and generic signatures without cloud, as well as by blocking malicious behavior when files get executed).

A good file detection rate is still one of the most important, deterministic and reliable features of an Anti-Virus product. Additionally, most products provide at least some kind of HIPS, behaviour-based, reputation-based or other functionalities to block (or at least warn about the possibility of) malicious actions e.g. during the execution of malware, when all other on-access and on-demand detection/protection mechanism failed.

Even if we deliver various tests and show different aspects of Anti-Virus software, users are advised to evaluate the software by themselves and build their own opinion about them. Test data or reviews just provide guidance to some aspects that users cannot evaluate by themselves. We suggest and encourage readers to research also other independent test results provided by various well-known and established independent testing organizations, in order to get a better overview about the detection and protection capabilities of the various products over different test scenarios and various test-sets.

Results

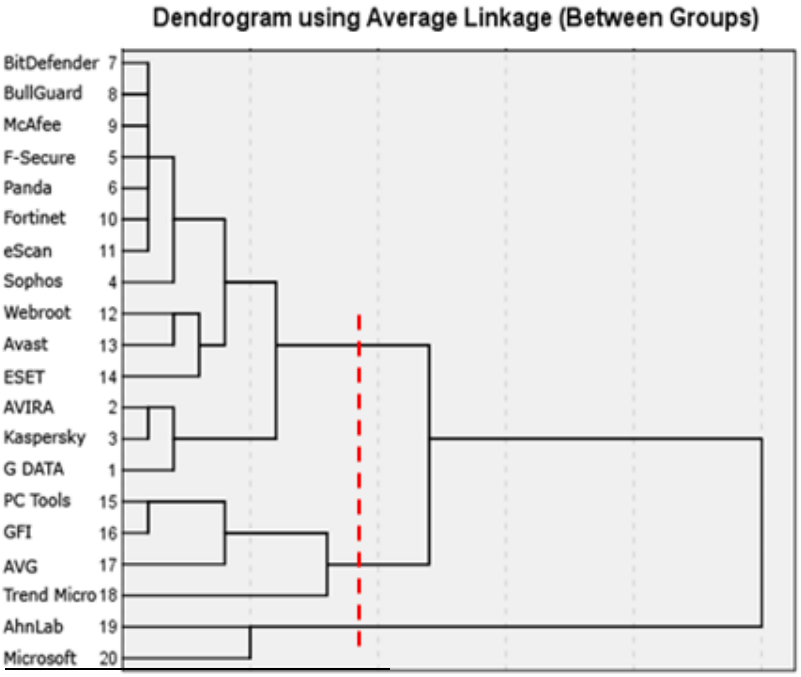
Please consider also the false alarm rates when looking at the below file detection rates[3].

Total detection rates (clustered in groups):

1.	G DATA	99.7%
2.	AVIRA	99.4%
3.	Kaspersky	99.3%
4.	Sophos	98.9%
5.	F-Secure, Panda, Bitdefender, BullGuard, McAfee	98.6%
6.	Fortinet, eScan	98.5%
7.	Webroot	98.2%
8.	Avast	98.0%
9.	ESET	97.6%
10.	PC Tools	97.2%
11.	GFI	97.0%
12.	AVG	96.4%
13.	Trend Micro	95.6%
14.	AhnLab	94.0%
15.	Microsoft	93.1%

The used test-set contained almost 300-thousands recent/prevalent samples from last months.

Hierarchical Cluster Analysis



This dendogram shows the results of the cluster analysis[4]. It indicates at what level of similarity the clusters are joined. The red drafted line defines the level of similarity. Each intersection indicates a group (in this case 3 groups).

[3] We estimate the remaining error margin on the final set (after applying clusters) to be under 0.2%
[4] For more information about cluster analysis, see e.g. this easy to understand tutorial: http://strata.uga.edu/software/pdf/clusterTutorial.pdf

False positive/alarm test

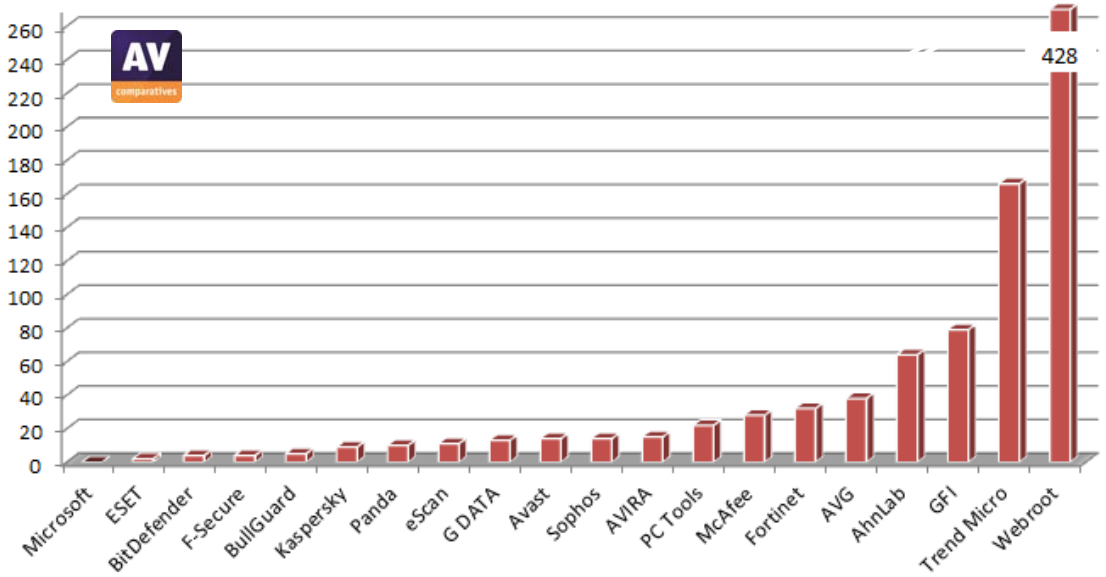
In order to better evaluate the quality of the file detection capabilities (distinguish good files from malicious files) of anti-virus products, we provide a false alarm test. False alarms can sometimes cause as much troubles as a real infection. Please consider the false alarm rate when looking at the detection rates, as a product which is prone to cause false alarms achieves higher scores easier. All discovered false alarms were reported/sent to the respective Anti-Virus vendors and should by now have been fixed.

False Positive Results

Number of false alarms found in our set of clean files (lower is better):

1.	Microsoft	0	
2.	ESET	2	very few FPs
3.	BitDefender, F-Secure	4	
4.	BullGuard	5	
5.	Kaspersky	9	
6.	Panda	10	few FP's
7.	eScan	11	
8.	G DATA	13	
9.	Avast, Sophos	14	
10.	AVIRA	15	
11.	PC Tools	22	
12.	McAfee	28	
13.	Fortinet	32	
14.	AVG	38	many FP's
15.	AhnLab	64	
16.	GFI	79	
17.	Trend Micro	166	
18.	Webroot	428	very many FP's

Details about the discovered false alarms (including their assumed prevalence) can be seen in a separate report available at: http://www.av-comparatives.org/images/stories/test/fp/avc_fp_mar2012.pdf



Award levels reached in this test

AV-Comparatives provides a ranking award (STANDARD, ADVANCED and ADVANCED+). As this report contains also the raw detection rates and not only the awards, expert users that e.g. do not care about false alarms can rely on that score alone if they want to.

AWARDS (based on detection rates and false alarms)	PRODUCTS
	✓ G DATA ✓ AVIRA ✓ Kaspersky ✓ Sophos ✓ F-Secure ✓ Panda ✓ Bitdefender ✓ BullGuard ✓ eScan ✓ Avast ✓ ESET
	✓ McAfee* ✓ Fortinet*
	✓ PC Tools* ✓ GFI* ✓ AVG* ✓ Microsoft ✓ Trend Micro* ✓ Webroot*
	✓ AhnLab*

*: those products got lower awards due to false alarms

The Awards are not only based on detection rates - also False Positives found in our set of clean files are considered. On page 6 of this report you can see how awards are being given.

A product that is successful at detecting a high percentage of malicious files but suffers from false alarms may not be necessarily better than a product which detects less malicious files but which generates less false alarms.

Copyright and Disclaimer

For more information about AV-Comparatives and the testing methodologies, please visit our website.

AV-Comparatives e.V. (April 2012)