Anti-Virus Comparative

File Detection Test of Malicious Software

includes false alarm test

Language: English
September 2012

Last Revision: 5th October 2012

www.av-comparatives.org

Table of Contents

Tested Products

- AhnLab V3 Internet Security 8.0
- avast! Free Antivirus 7.0
- AVG Anti-Virus 2012
- AVIRA Antivirus Premium 2012
- BitDefender Antivirus Plus 2013
- BullGuard Antivirus 12
- eScan Anti-Virus 11.0
- ESET NOD32 Antivirus 5.2
- F-Secure Anti-Virus 2012
- Fortinet FortiClient Lite 4.3
- G DATA AntiVirus 2013
- GFI Vipre Antivirus 2012
- Kaspersky Anti-Virus 2013
- McAfee AntiVirus Plus 2012
- Microsoft Security Essentials 4.0
- Panda Cloud Free Antivirus 2.0.0
- PC Tools Spyware Doctor with AV 9.0
- Sophos Anti-Virus 10.0
- Trend Micro Titanium AntiVirus+ 2013
- 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 21st August 2012 and consisted of 240859 sample variants. The products were updated on the 28th August 2012. The following twenty up-to-date products were included in this public test (most current ones available at time of testing):

- AhnLab V3 Internet Security 8.0.6.13
- avast! Free Antivirus 7.0.1466
- AVG Anti-Virus 2012.0.2197
- AVIRA Antivirus Premium 12.0.0.1183
- Bitdefender Anti-Virus+ 16.18.0.1406
- BullGuard Antivirus 12.0.231
- eScan Anti-Virus 11.0.1139.1225
- ESET NOD32 Antivirus 5.2.9.12
- F-Secure Anti-Virus 12.56.100
- Fortinet FortiClient Lite 4.3.5.472
- G DATA AntiVirus 23.0.3.2
- GFI Vipre Antivirus 5.0.5162
- Kaspersky Anti-Virus 13.0.1.4190
- McAfee AntiVirus Plus 11.6.385
- Microsoft Security Essentials 4.0.1526.0
- Panda Cloud Free Antivirus 2.0.0
- PC Tools Spyware Doctor with Antivirus 9.0.0.2308
- Sophos Anti-Virus 10.0.8
- Trend Micro Titanium AntiVirus Plus 6.0.1215
- Webroot SecureAnywhere AV 8.0.1.233

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.

We invite users to look at our other tests and not only at this type of test. This is to make users aware of other types of tests and reports that we are providing already since several years, like e.g. our Whole-Product “Real-World” Protection Test. 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 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 (these protection features are evaluated in other types of tests that we provide on our website).

[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).

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

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

Several products make use of cloud technologies, which require an active internet connection. Our tests are performed using an active internet connection. 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 gets lost e.g. during a scan, which may affect considerably the provided protection and make e.g. an initiated scan useless.

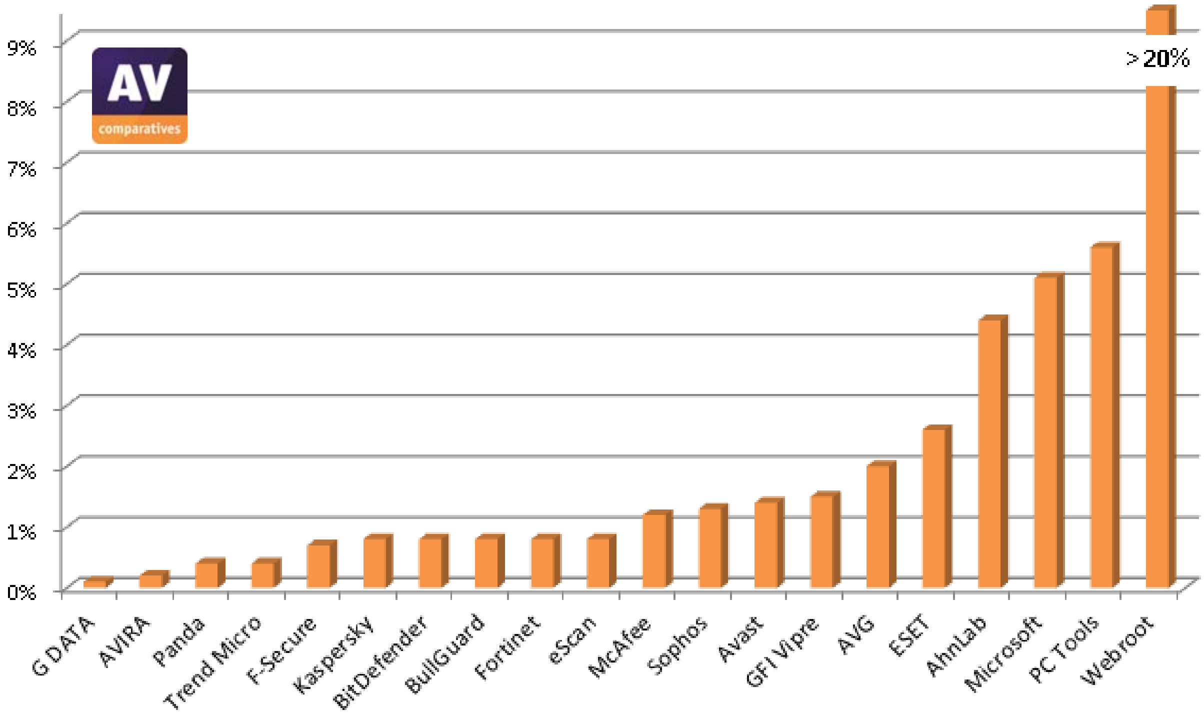
The used test-set has been built consulting telemetry data in attempt to include prevalent samples from the last weeks/months which are/were hitting users in the field. We applied a clustering method to classify similar files (one sample - the most prevalent one - per variant). This allows us to evaluate prevalent similar samples and reducing the size of the set without introducing bias. Therefore, each miss is intended to represent one missed group/variant of files. As we used even more recent malware samples and even more prevalent samples than previously, the size of test-set is even smaller (and should be easier to get discovered/detected by AV products).

As usual, all test results have been peer-verified before publishing. We also include crosschecks e.g. checking cloud connectivity from different IPs, service providers, countries, using various product licenses, changing test environment and related variables, testing over a period of time, etc. Furthermore, we sometimes provide random sets of samples to other trusted independent labs to let them check if the product behaves the same and if they get the same results as we do.

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. As we reduced also the clean set, we updated also the awarding system, as some products manage to get way too false alarms even in the reduced clean set. We are still evaluating whether to apply clusters also for the FP ranges or to set next year the threshold for "Many" at 11 instead 16.

	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 (16-50 FP's)	**TESTED**	**TESTED**	**STANDARD**	**ADVANCED**
Very many (51-100 FP's)	**TESTED**	**TESTED**	**TESTED**	**STANDARD**
Crazy many (over 100 FP's)	**TESTED**	**TESTED**	**TESTED**	**TESTED**

Graph of missed samples (lower is better)



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

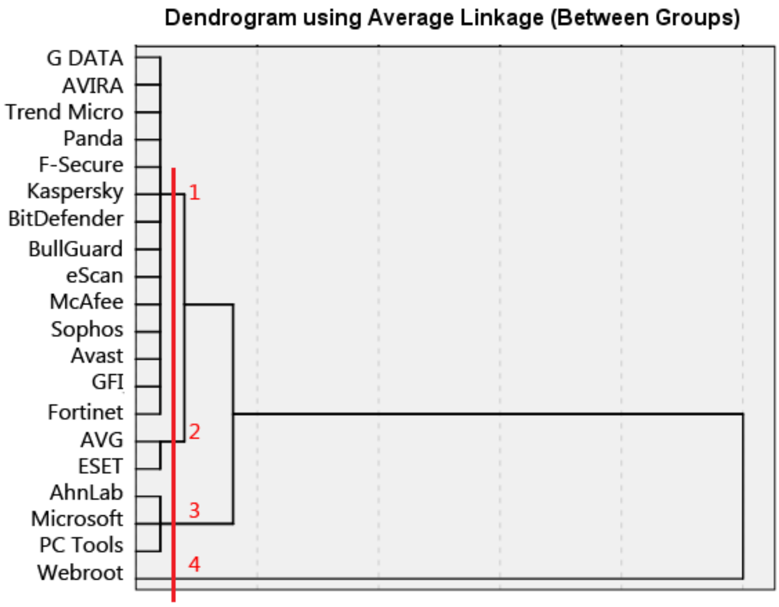
Please consider also the false alarm rates when looking at the below file detection rates[2].

Total detection rates (clustered in groups):

1.	G DATA	99.9%
2.	AVIRA	99.8%
3.	Panda, Trend Micro	99.6%
4.	F-Secure	99.3%
5.	Kaspersky, BitDefender, BullGuard, Fortinet, eScan	99.2%
6.	McAfee	98.8%
7.	Sophos	98.7%
8.	Avast	98.6%
9.	GFI Vipre	98.5%
10.	AVG	98.0%
11.	ESET	97.4%
12.	AhnLab	95.6%
13.	Microsoft	94.9%
14.	PC Tools	94.4%
15.	Webroot	below 80%[3]

The used test-set contained around 240000 recent/prevalent samples from last weeks/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 four groups).

[2] We estimate the remaining error margin on the final percentages to be below 0.2%

[3] Results and misses have been confirmed with several tests and also by the vendor. Even two weeks after the test-date, the detection rate was still lower than 80%. Also a representative set of misses was sent for peer-review to an independent party.

[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

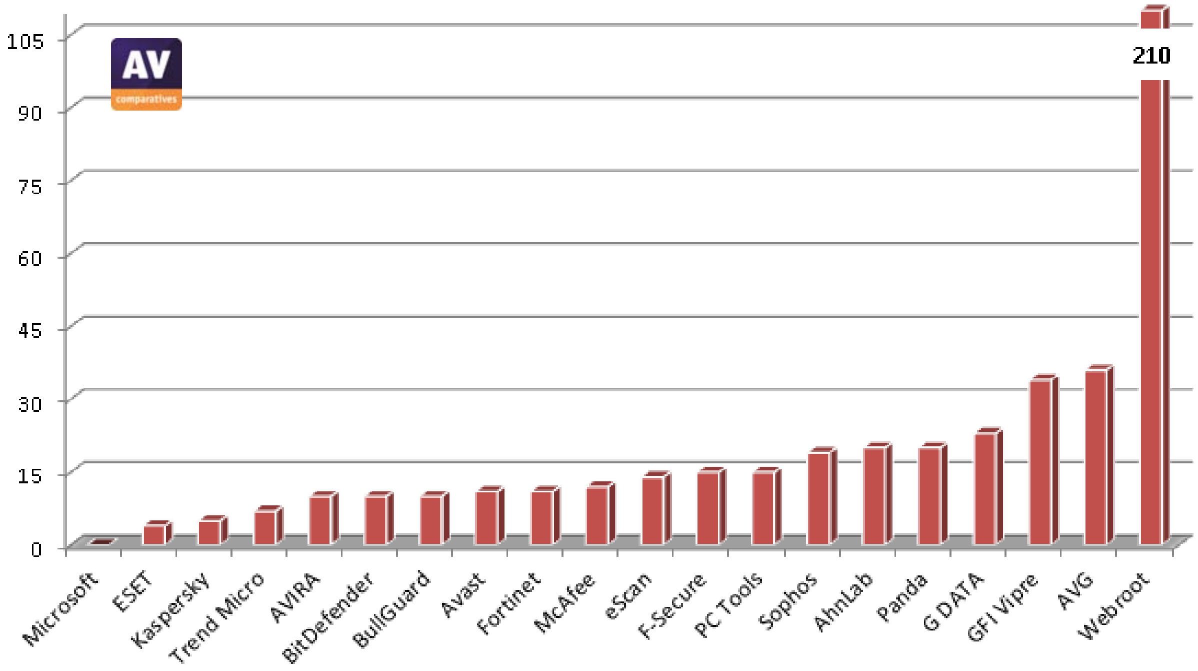
In order to better evaluate the quality of the file detection capabilities (distinguish good files from malicious files) of anti-virus products, we provide a false alarm test. False alarms can sometimes cause as much troubles as a real infection. Please consider the false alarm rate when looking at the 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	very few FPs
2.	ESET	4	
3.	Kaspersky	5	
4.	Trend Micro	7	
5.	AVIRA, BitDefender, BullGuard	10	few FP's
6.	Avast, Fortinet	11	
7.	McAfee	12	
8.	eScan	14	
9.	F-secure, PC Tools	15	
10.	Sophos	19	
11.	AhnLab, Panda	20	
12.	G DATA	23	many FP's
13.	GFI Vipre	34	
14.	AVG	36	
15.	Webroot	210	crazy 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/docs/avc_fps_201209_en.pdf



Award levels reached in this test

AV-Comparatives provides a ranking award. 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**
AV comparatives ADVANCED+ ★★★ ON DEMAND DETECTION TEST OCT 2012	✓ AVIRA ✓ Trend Micro ✓ F-Secure ✓ Kaspersky ✓ BitDefender ✓ BullGuard ✓ Fortinet ✓ eScan ✓ McAfee ✓ Avast
AV comparatives ADVANCED ★★ ON DEMAND DETECTION TEST OCT 2012	✓ G DATA* ✓ Panda* ✓ Sophos* ✓ GFI Vipre* ✓ ESET
AV comparatives STANDARD ★ ON DEMAND DETECTION TEST OCT 2012	✓ AVG* ✓ Microsoft ✓ PC Tools
AV comparatives TESTED ON DEMAND DETECTION TEST OCT 2012	✓ AhnLab* ✓ Webroot

*: 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

This publication is Copyright © 2012 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 a liability for the correctness of the test results cannot be taken by any representative of AV-Comparatives e.V. 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 a registered Austrian Non-Profit-Organization.

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

AV-Comparatives e.V. (October 2012)

validEDGE
MIS 1100