

Anti-Virus Comparative



Retrospective/Proactive test

Heuristic and behavioural protection
against new/unknown malicious software

Language: English

March 2014

Last revision: 4th July 2014

www.av-comparatives.org

Contents



1. Introduction	3
2. Description	3
3. False-alarm test	5
4. Test Results	5
5. Summary results	6
6. Awards reached in this test	7
7. Copyright and disclaimer	8

1. Introduction

This test report is the second part of the March 2014 test¹. The report is delivered several months later due to the large amount of work required, deeper analysis, preparation and dynamic execution of the retrospective test-set. This type of test is performed only once a year and includes a behavioural protection element, where any malware samples are executed, and the results observed. Although it is a lot of work, we usually receive good feedback from various vendors, as this type of test allows them to find bugs and areas for improvement in the behavioural routines (as this test evaluates specifically the proactive heuristic and behavioural protection components).

The products used the same updates and signatures that they had on the 7th March 2014. This test shows the proactive protection capabilities that the products had at that time. We used 1,200 new, unique and very prevalent malware samples that appeared for the first time shortly after the freezing date. The size of the test-set has also been reduced to a smaller set containing only one unique sample per variant, in order to enable vendors to peer-review our results in a timely manner. The following products were tested:

- AhnLab V3 Internet Security 8.0
- AVG Internet Security 2014
- Bitdefender Internet Security 2014
- BullGuard Internet Security 2014
- Emsisoft Anti-Malware 8.1
- eScan Internet Security 14.0
- ESET Smart Security 7.0
- F-Secure Internet Security 2014
- Fortinet FortiClient 5.0
- Kaspersky Internet Security 2014
- Lavasoft Ad-Aware Free Antivirus+ 11.1
- Microsoft Security Essentials 4.4
- Tencent QQ PC Manager 8.9

This test is currently an optional part of our public main test-series, that is to say, manufacturers can decide at the beginning of the year whether they want their respective products to be included in the test. Although we may continue to perform the test in the future, it will no longer be part of our public main test-series and be provided only on request as it is very resource-intensive.

Readers may be interested to see a summary and commentary of this test which was published by PC Mag last year: <http://securitywatch.pcmag.com/security-software/315053-can-your-antivirus-handle-a-zero-day-malware-attack>

2. Description

Many new malware samples appear every day, which is why it is important that antivirus products not only provide new updates, as frequently and as quickly as possible, but also that they are able to detect such threats in advance with generic/heuristic techniques; failing that, with behavioural protection measures. Even if nowadays most antivirus products provide daily, hourly or cloud updates, without proactive methods there is always a time-frame where the user is not reliably protected. The aim of this test is to evaluate the proactive detection and protection rates in this time-frame (without

¹ http://www.av-comparatives.org/wp-content/uploads/2014/04/avc_fdt_201403_en.pdf

cloud). The data shows how good the proactive heuristic/generic detection and behavioural protection capabilities of the scanners were in detecting new threats used in this test. The design and scope of the test mean that only the heuristic/generic detection capability and behavioural protection capabilities were tested (offline). Additional protection technologies (which are dependent on cloud-connectivity) and infection vectors are considered by AV-Comparatives in e.g. Whole-Product Dynamic (“Real-World”) Protection Tests and other tests, but are outside the scope of the Retrospective/Proactive Tests.

We included in the retrospective test-set only new malware that was very prevalent in-the-field shortly after the freezing date. Samples which were not detected by the heuristic/generic detection capabilities of the products were then executed in order to see if behaviour-blocking features would stop them. In several cases, we observed that behaviour blockers only warned about some dropped malware components or system changes, without protecting against all the malicious actions performed by the malware; such cases were not counted as a block. As behaviour blockers only come into play after the malware is executed, a certain risk of being compromised remains (even when the security product claims to have blocked/removed the threat). Therefore, it is preferable that malware be detected before it is executed, by e.g. the on-access scanner using heuristics. This is why behaviour blockers should be considered a complement to the other features of a security product (multi-layer protection), and not a replacement.

What about the cloud? Even several weeks later, a number of the malware samples used were still not detected by some cloud-dependent products, even when their cloud-based features were available. Consequently, we consider it a marketing excuse if retrospective tests - which test the proactive protection against new malware - are criticized for not being allowed to use cloud resources. This is especially true considering that in many corporate environments the cloud connection is disabled by the company policy, and the detection of new malware coming into the company often has to be provided (or is supposed to be provided) by other product features. Cloud features are very (economically) convenient for security software vendors and allow the collection and processing of large amounts of metadata. However, in most cases (not all) they still rely on blacklisting known malware, i.e. if a file is completely new/unknown, the cloud will usually not be able to determine if it is good or malicious.

The awards are given by the testers after consulting a number of statistical methods, including hierarchical clustering². This time we based our decisions on the following scheme:

	Proactive Protection Rates			
	Under 50%	Cluster 3	Cluster 2	Cluster 1
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

² http://en.wikipedia.org/wiki/Hierarchical_clustering

3. False alarm test

To better evaluate the proactive detection capabilities, the false-alarm rate has to be taken into account too. A false alarm (or false positive [FP]) occurs when an antivirus product flags an innocent file as infected. False alarms can sometimes cause as much trouble as real infections.

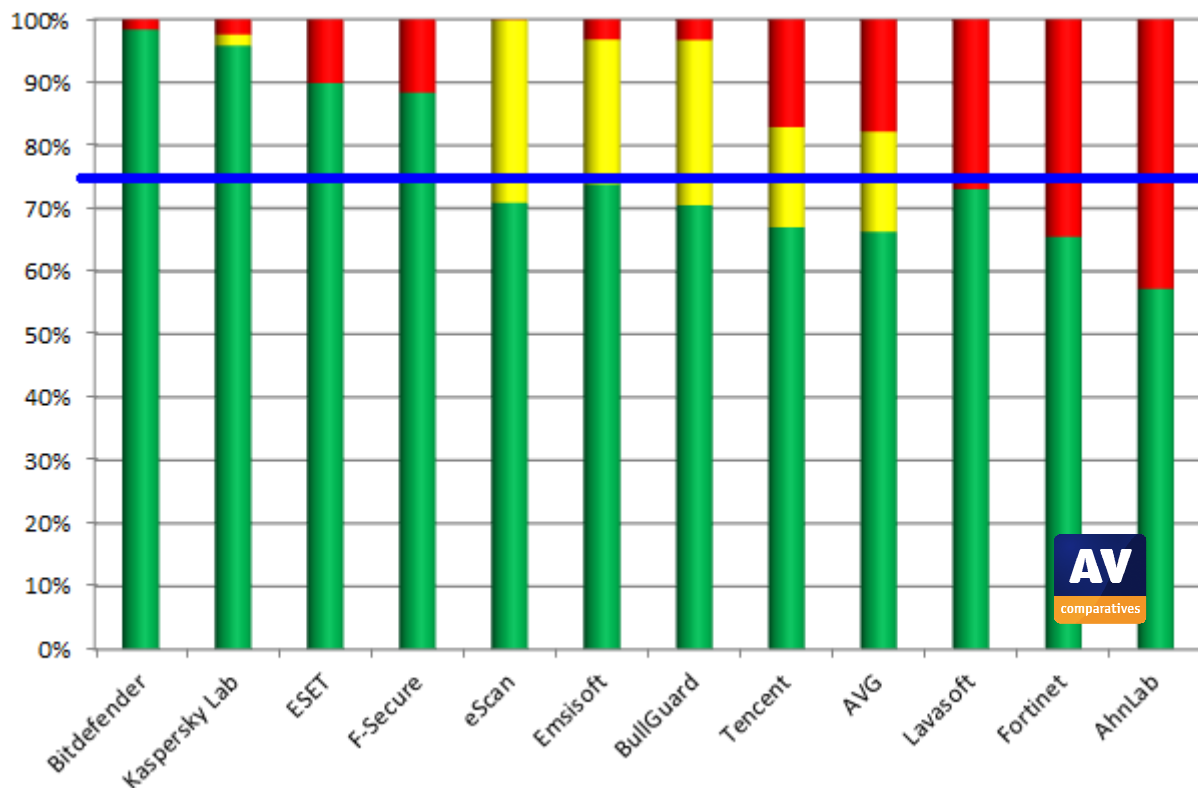
The false-alarm test results were already included in the March test report. For details, please read the report, available at http://www.av-comparatives.org/wp-content/uploads/2014/04/avc_fps_201403_en.pdf

Very few false alarms (0-1):	ESET, Microsoft
Few false alarms (2-10):	Fortinet, Bitdefender, BullGuard, Emsisoft, Lavasoft, Tencent, eScan, F-Secure, AVG, Kaspersky Lab
Many false alarms (over 10):	AhnLab

A small behavioural false-alarm test using 200 most downloaded/common software packages released in February did not bring up any additional false alarms.

4. Test Results

The table below shows the proactive protection capabilities of the various products. The awards given (see page 7 of this report) consider not only the protection rates against new malware, but also the false alarm rates.



Key:
Green = blocked/protected
Yellow = user dependent
Red = not blocked/compromised
The blue line indicates the results of Microsoft Security Essentials

5. Summary results

The results show the proactive (generic/heuristic/behavioural) protection capabilities of the various products against new malware. The percentages are rounded to the nearest whole number.

To know how these antivirus products perform with updated signatures and cloud connection against prevalent malware files, please have a look at our File Detection Tests of March and September. To find out about real-life online protection rates provided by the various products, please have a look at our ongoing Whole-Product Dynamic “Real-World” Protection tests. Readers should look at the results and decide on the best product for them based on their individual needs. For example, laptop users who are worried about infection from e.g. infected flash drives whilst offline should pay particular attention to this Proactive test.



Below you can see the proactive protection results over our set of new and very prevalent malware appeared in-the-field (1,200 malware samples):

	Blocked	User dependent ³	Compromised	Proactive Protection Rate	False Alarms	Cluster
Bitdefender	1180	-	20	98%	few	1
Kaspersky Lab	1150	21	29	97%	few	1
ESET	1078	-	122	90%	very few	1
F-Secure	1060	-	140	88%	few	1
eScan	851	348	1	85%	few	1
Emsisoft	886	275	39	85%	few	1
BullGuard	846	314	40	84%	few	1
Microsoft	900	-	300	75%	very few	2
Tencent	804	191	205	75%	few	2
AVG	795	191	214	74%	few	2
Lavasoft	876	-	324	73%	few	2
Fortinet	786	-	414	66%	few	3
AhnLab	687	-	513	57%	many	3

³ User-dependent cases were given a half credit. Example: if a program blocks 80% of malware by itself, plus another 20% user-dependent, we give it 90% altogether, i.e. 80% + (20% x 0.5).

6. Awards reached in this test

The following awards are for the results reached in the proactive/behavioural test:

AWARDS	PRODUCTS
	Bitdefender Kaspersky Lab ESET F-Secure eScan Emsisoft BullGuard
	Tencent AVG Lavasoft
	Fortinet
	AhnLab

Microsoft security products are not included in the awards page, as their out-of-box protection is (optionally) included in the operating system and is therefore out-of-competition.

7. Copyright and Disclaimer

This publication is Copyright © 2014 by AV-Comparatives®. 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, prior to any publication. AV-Comparatives 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 can be 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 (July 2014)