Anti-Virus Comparative

Whole Product Dynamic Test

Protection Offered by Internet Security Suites

Language: English
December 2009
Last Revision: 16th December 2009

www.av-comparatives.org

Table of Contents

Products Tested

- avast! Free 5.0
- AVG Internet Security 9.0
- AVIRA Premium Security Suite 9.0
- BitDefender Internet Security 2010
- eScan Internet Security 10.0
- ESET Smart Security 4.0
- F-Secure Internet Security 2010
- G DATA Internet Security 2010
- Kaspersky Internet Security 2010
- Kingsoft Internet Security 9+
- McAfee Internet Security 2010
- Microsoft Security Essentials 1.0
- Norman Security Suite 7.2
- Symantec Norton Internet Security 2010
- Trustport PC Security 2010

Introduction

The goal of this Whole Product Dynamic Test is to compare the protection offered by various security solutions, by testing them under real-world conditions. There has been a lot of talk in the past years about such tests and their value for home users. Some issues related to these tests are that they are very expensive to perform (due to the time and personnel required) and difficult to replicate. Nonetheless, such tests are very important and show the ability of the various security products to protect the users against malware.

This is our first public whole product dynamic test, and although our goal was to test many more samples, due to time/resources restrictions and some unexpected issues, we had to cut down to the relatively small number of around 100 test cases. Based on the experience and issues observed during this test, as well as feedback from AV vendors, AV-Comparatives will start providing whole product dynamic tests regularly starting from 2010, using a much greater number of test samples (to increase the statistical relevance), further infection vectors and improved reproducibility, by developing an automated system in co-operation with the Institute for Informatics and Quality Engineering of the University of Innsbruck.

Products included in this test

The products tested are listed on the previous page. In Whole Product Dynamic Tests we use the security suite products offered by the vendors. If such a suite is not available, as is currently the case with Avast and Microsoft, their results have to be considered, as per their request, as "non-competitive" (although we did anyway not consider e.g. firewalls). Sophos decided not to participate in this test as their business oriented product is used differently compared to the other consumer based products. All products are tested using their default/recommended settings and with latest product and signature updates at the time of testing.

Test cases used

We included 100 test cases in this test. A test case is a website containing a malicious script or exploit (pointing to malware) or malicious file. Based on threat statistics, nowadays over 70% of malware is delivered through websites carrying malicious scripts or exploits (drive-by downloads) and nearly 20% by social engineering tactics pointing to websites where users can manually download malicious software (the remaining percentage comes through other infection vectors). Furthermore, most infected websites are currently on Chinese domains. Our test-case selection took also this into account; we used mainly websites with exploits and malicious scripts, and only 15 links pointing directly to executable malware; about 30 sites were on Chinese domains. The URLs were collected by using our own in-house crawler; to avoid bias, we did not use any publicly available services which deliver malicious URL feeds. For security reasons, we do not publish malicious URLs. We also took care to do not include several URLs leading to the same malware, in order to have a variety of test cases in the test. Althought we used 15-20 freshly collected URLs each testing day, we want to make clear that using newly discovered infected websites does not necessarily mean that we used "zero-day exploits/malware". The test goal is not to confront the security products against zero-day malware – the goal is to represent a realistic picture of the security products as experienced by most home users in the real world, when using the product and surfing the Internet.

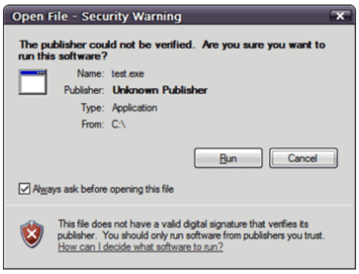
Test system used

In order to reflect the most common system used by home users in the world (and consequently also the most frequently targeted by malware, through exploits etc.), we did not use the latest versions of the operating system or applications. We ran the test under Windows XP Professional SP3 (basic Service Pack), using Internet Explorer 7 (which according to Internet statistics is one of the most prevalent versions, after IE6[1]), and Adobe Acrobat Reader 8 (even a large number of visitors of our website, whom we considered to be more security-aware, still seem to be using outdated software, as many of them have contacted us to say that their version of Acrobat Reader is not able to read our PDF reports, which require at least version 8). Also, not the latest, but statistically the most commonly used versions of Java, FlashPlayer, etc. were used. In summary, it can be said that we used a system that is about a year out of date – in the real-world, even more-outdated systems may be prevalent, but we preferred not to use very old software. Having said this, we want to emphasize to users the importance of always keeping all their software (not just security software) up-to-date, as many exploits etc. would not work on updated/patched software versions. We will continue to observe the usage statistics and switch to newer software versions, as well as Windows 7, as soon as they become the most prevalent systems.

Excursion: Security provided within the operating system and the browser

As this report is aimed towards home users, we thought it would be a good chance to provide some information about the security features included in their operating system and browser.

First of all: please update your operating system and browser to the latest versions - do not ignore or turn off the automatic updates! However, although many people consider their OS to be at fault in the event of malware infections, in reality the reason for an infection is usually actions taken by the users themselves - not just failure to keep their software up-to-date. Almost every time a new, unknown program is launched, a Windows prompt appears, warning the users of the risk of executing the file. Furthermore, a similar warning message appears if a user downloads a file from the Internet and wants to run it. A few examples are given below:



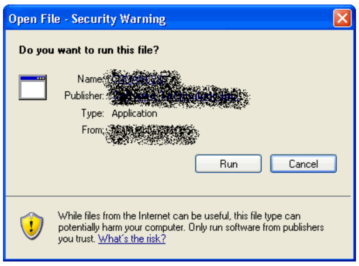
The publisher could not be verified. Are you sure you want to run this software?

Name: test.exe
Publisher: Unknown Publisher
Type: Application
From: C:\

☑ Always ask before opening this file

This file does not have a valid digital signature that verifies its publisher. You should only run software from publishers you trust. How can I decide what software to run?

Open File - Security Warning: *"The publisher could not be verified. Are you sure you want to run this software? – The file does not have a valid digital signature that verifies its publisher. You should only run software from publishers you trust."*



Do you want to run this file?

Name: [redacted]
Publisher: [redacted]
Type: Application
From: [redacted]

While files from the Internet can be useful, this file type can potentially harm your computer. Only run software from publishers you trust. What's the risk?

Open File - Security Warning: *"Do you want to run this file? - While files from the Internet can be useful, this file type can potentially harm your computer. Only run software from publishers you trust."*

[1] http://marketshare.hitslink.com/browser-market-share.aspx?qprid=2&qpmr=40&qpdt=1&qpct=3&qptimeframe=Y

Most users would ignore such warning messages and execute the files anyway, as they are used to seeing such warnings. Also, with today's social engineering tactics, users can easily be tricked into launching applications in spite of any warnings. This is why it is also important for security software not to rely on users' decisions to provide security; users expect the security software to do this for them. A good security product should clearly state if a file is malicious or not, and if it thinks that it is malicious, it should not allow the user to execute the file (or at least have "block/quarantine" as default choice option). If a product has very often to rely on user decisions whether a file is save to run or not, there is no big benefit for home users in using such a product.

Additionally, Internet Explorer 8 comes with SmartScreen Filter, blocking many malicious files while browsing the internet. Even Google and Mozilla Firefox block dangerous websites. Looking at all these security features available, one may wonder how people still manage to infect their machines. The problem is not always the products or technologies, in most cases it is the users' fault.

How we tested

The Whole-Product-Dynamic Test is not a simple "detection" test as usual, it is more a "protection/prevention" test. The test mimics malware reaching and executing on a user's machine, as it happens in the real world (e.g. by visiting a website with a malicious payload such as drive-by downloads/exploits, or by being fooled into downloading a malicous file by social engineering tactics). This means that not only the signatures, heuristics and in-the-cloud detections are evaluated, but URL-blockers, Web reputation services, exploit-shields, in-the-cloud heuristics, HIPS and behavioral detection are also considered. Firewall warnings when the malware was already running and just trying to connect to the outside world were considered a fail. We browsed to websites with exploits/drive-by downloads, and also to a few websites with malicious files that we downloaded and executed. The criteria for success/failure is independent of the technology used by the products. What matters is that the products provide reliable protection to the user, ideally without requiring any user decisions as to whether something is malicious or not.

For the Whole-Product-Dynamic Test we used 16 identical physical PCs (not virtual machines), with identical hardware, software and OS configuration (administrator account). Each PC had one security product installed. We used the security suite product of each vendor where available, evaluating the overall protection provided. Products were always up-to-date and had a live Internet connection, as in the real world. Each machine had its own IP address. We used the default settings of the products. The test started on the 16th November 2009. Each day we tested about 15 or 20 test cases (new URLs with fresh/relevant exploits/malware, but taking care not to use URLs which deliver identical malware) gathered from our own crawler. As each machine had to be inspected, and all machines returned to their original state (which meant waiting until all machines were ready for the next threat), it took nearly 12 hours each day for 4 people to perform the tests (although we developed tools to speed up some procedures). Each test case was first verified by browsing to it on an unprotected system (with no security software installed), in order to see if the sample was valid and did something in the test environment. After that, all 15 security products were updated before browsing to any test-case. We took care that the site exposed all the machines to the same threat. All URLs were browsed to at the same moment, and screenshots taken in the event that the security product reacted; otherwise, we checked to see if the product had taken any action silently, or if the threat had been successful in compromising the machine (i.e. the security product had failed).

URLs which e.g. delivered different malware to the machines, or went down during the test, were excluded afterwards. Due to this, the test ran until the 26th of November, and the number of valid test cases for the report was reduced to 100.

Although we saved a lot of data, reproducing dynamic tests is a difficult task, especially if done the way we always use: on physical machines and without any simulated environments, but still taking care that no malware breaks out of the test network. Furthermore, some products do not provide logs for everything they do, and in-the-cloud products can deliver different results if tested at different times (or even in different countries). During the test we observed issues which need to be addressed in further dynamic tests and taken into account in our automated dynamic testing model. We are working with the University of Innsbruck to develop this model. We plan to have it done and use it as soon as 2010, allowing us to use a much larger number of test cases, improved logging, reproducibility and also additional attack vectors (like email, IM, P2P, USB, etc.). In some cases it requires some vendors to change or improve their products too, to make it possible for testers to automate such testing and support them with a standard.

In most cases the security products took the appropriate action by themselves (this is what we usually call deterministic – either something is malicious and should be blocked, or it is not); in a few cases they asked the user what to do, but suggested blocking the threat as the default option. We always took the default action when asked, and considered suggestions to block as “success/protected” (if the machine was indeed really protected, as we did not blindly trust what the product claimed). If no default option was available and the warning indirectly suggested that the program/activity/website might be dangerous, we chose “block”; the same action was also applied in the false alarm/noise test, if such a warning had appeared during the test with actual malware. If during the malware test-cases there were no warnings where the user had to decide (because no default option was given), the warning without a default option was not considered as false alarm during the test with clean applications.

Firewall warnings/pop-ups were not taken into account, because they usually just announce that a program (which might well be a known clean application) is trying to connect to the outside world, and ask the user what to do. Some firewalls, such as from AVG, AVIRA, Bitdefender, F-Secure and Trustport are in our opinion still a bit too chatty, requiring user interaction/decisions. Kingsoft is a particular offender, and even suggests blocking the connections of well-known, important programs. AVIRA sometimes gives a firewall warning, but suggests allowing the connection. In our opinion, in such cases AVIRA should perform the suggested action (allow) by itself, without bothering the user.
The Host-Based Intrusion Prevention Systems (HIPS) of F-Secure and especially G DATA may sometimes warn about system configuration changes made by applications, but with the default option being to allow them (also because such changes are very often observed even during clean software installations). Our view is that the products should just carry out the suggested action (allow) by themselves, and not ask the user to make a decision. If users are often confronted with such warnings, even during the installation of known, safe applications, they may get used to allowing these changes, and do the same thing without thinking in the few cases when they really are executing malware.

Symantec's Norton Download Insight messages were not taken into account in the tests. Download Insight uses new reputation technology from Symantec to block malicious files and warn on files where the reputation is not yet established. If we would have considered it, Symantec would have protected also against the one malware it missed.

False Alarm/Noise Test (“Oversensitivity”)

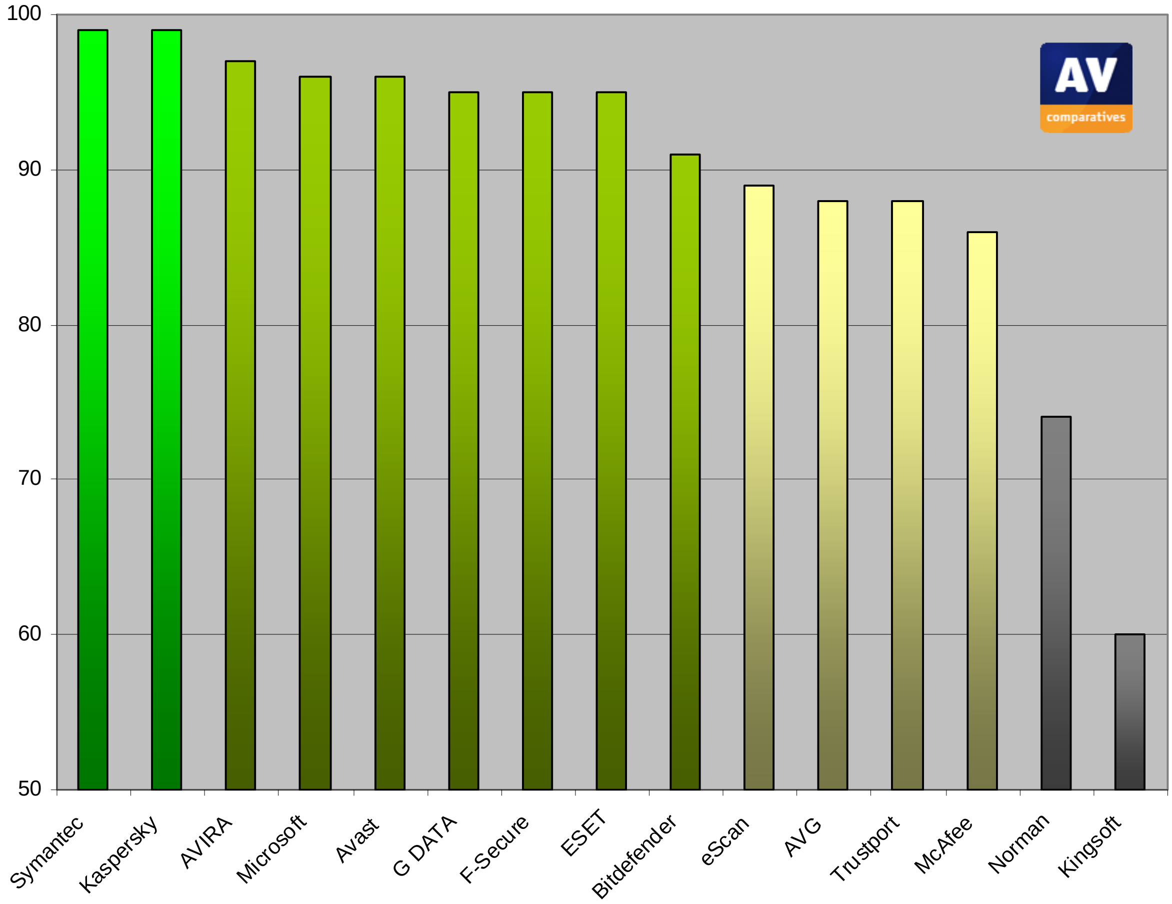
To provide a balanced test of user experience, we wanted to include also a false alarm/noise test, to see if the protection features of the security products might be oversensitive, and show the same alerts while browsing safe websites, and installing or using clean applications. We tested 40 clean test-cases, randomly choosen from various download portals. We browsed to the websites, downloaded, unarchivied where necessary, installed and ran/used the installed applications to check them for functionality, and to see if the security products interfered. This test alone was a great deal of work.

Most products did not interfere, while some other products (like ESET, F-Secure, Kaspersky and Symantec) had only one case, where they blocked automatically a clean application. Initially we wanted to penalize based on one FP only, but we came to the conclusion that this wouldn’t be statistically meaningful enough to degrade products and mark them as “oversensitive”.

A look inside our dynamic testing lab:

Test Results

		Threats blocked
1.	Symantec, Kaspersky	99 out of 100
2.	AVIRA	97 out of 100
3.	Microsoft, Avast	96 out of 100
4.	G DATA, F-Secure, ESET	95 out of 100
5.	Bitdefender	91 out of 100
6.	eScan	89 out of 100
7.	Trustport, AVG	88 out of 100
8.	McAfee	86 out of 100
9.	Norman	74 out of 100
10.	Kingsoft	60 out of 100



In our opinion, the above results show that despite the good protection features build into the security products, users should never expect to be automatically 100% protected just by using them. Furthermore, the more security a user expects or wants, the more the usability may decrease, and the noise due to oversensitive protection features (or chatty products which ask for user interaction/decisions) may increase.

Award levels reached in this test

AV-Comparatives provides a 4-level-ranking-system (Tested, STANDARD, ADVANCED and ADVANCED+).

AWARDS	PRODUCTS (in no specific order) [2]
AV comparatives ADVANCED+ ★★★ DYNAMIC PROTECTION TEST DEC 09	✓ Symantec ✓ Kaspersky
AV comparatives ADVANCED ★★ DYNAMIC PROTECTION TEST DEC 09	✓ AVIRA ✓ Microsoft ✓ Avast ✓ G DATA ✓ F-Secure ✓ ESET ✓ BitDefender
AV comparatives STANDARD ★ DYNAMIC PROTECTION TEST DEC 09	✓ eScan ✓ TrustPort ✓ McAfee[3] ✓ AVG
AV comparatives TESTED DYNAMIC PROTECTION TEST DEC 09	✓ Norman ✓ Kingsoft

Protection			
<80%	80 – 90%	90 – 98%	98 – 100%
tested	STANDARD	ADVANCED	ADVANCED+

The above scoring system is an attempt to rate the results. We will change/adapt/improve it in the next tests. Considering that in the whole-product dynamic tests the products are tested as a whole, and various protection features come into play, we expect very good scores from the products if they are to receive the ADVANCED+ award. We also want readers to understand that as AV-Comparatives only includes good products in its main tests, even a STANDARD award is already a good score; ADVANCED is very good and ADVANCED+ exceptional.

[2] We suggest considering all products with the same award to be as good as each other.

[3] Tested McAfee product is about to be replaced soon by a newer version. We could not test this new version due to unfortunate timing of the release with respect to this testing.

Copyright and Disclaimer

AV-Comparatives e.V. and its testers cannot be held liable for any damage or loss which might occur as a 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 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. (December 2009)