# Antivirus Myths and Facts

**By Helmuth Freericks**

# Antivirus Myths and Facts

## Introduction

As an active, veteran member of the antivirus community and a pioneer of one of the earliest antivirus companies, I have spoken with thousands of people with an interest in antivirus over the last two decades. Some were journalists tasked with testing antivirus products and publishing the results. Some were product managers or IT directors who were interested in integrating an antivirus engine into their product, service offering or IT infrastructures. Some wanted to compare multiple engines to make sure they were making the best decision. Many were simply satisfied when the antivirus worked, and considered it somewhat of a black box, not caring how or why it worked, or what sort of system resources it required.

The one thing that I consistently came across in all my dealings with this highly intelligent, security conscious group of people is that there were many myths about antivirus software, that have persisted over these many years. These myths can cause misunderstandings, and potentially cost organizations money and resources. For example, there are differences among testing methodologies that can cause one antivirus engine to appear more powerful than another, where in a real-life situation the opposite may be true.

After speaking one-on-one with people for nearly 20 years, sharing facts about antivirus testing and other key areas, I decided it was time to record these myths and facts in a formal way. This paper is designed to shed some light on the most common myths, and provide some seasoned advice on the most objective and effective ways to judge your antivirus engine.

## The most common myths are:

**1.** Antivirus software can only detect specific, known viruses.

**2.** If an antivirus has flagged or blocked a file, this file is definitely malware.

**3.** Third Party evaluations of AV engines are more accurate than tests run by AV vendors.

**4.** Testing an antivirus solution should be done by throwing as many viruses at it as possible.

## Myth: Antivirus software can only detect specific, known viruses



**Fact:** In the very early days of antivirus technology – around the late '80s to early '90s – antivirus solutions were designed to detect known malware. This solution was sufficient due to the length of attacks, which could last weeks or months; the relatively few number of viruses; the slow propagation of viruses, mostly via infected files and floppy disks (the Web as we know it today was not yet born).

The world today is very different, with attacks lasting hours or minutes, tens of thousands of new variants every day and widespread access to the Web. These factors mean millions of users can be exposed to viruses in an instant. To deal with this threat, most antivirus engines today use a variety of approaches and technologies to detect malware, especially as-yet-unseen malware. The technologies used by antivirus solutions can be roughly categorized into two types: reactive and proactive.

The more basic, reactive, approach to malware detection consists of using a hash or fingerprint of an infected file. Detection is almost always exact, meaning it is an almost sure way to detect a malicious sample. It can provide an exact match between a malware or malware family and its antivirus signature, which simplifies the malware cleanup process in the case of infection. However, due to the huge volume of malware prevalent today, exclusive use of signature-based detection is no longer feasible. Since every type of malware needs its own signature, definition file updates end up being too large, resulting in ineffective use of available resources. In addition, creation of signatures can often take too much time to be effective.

A more advanced, proactive, approach to virus detection is commonly referred to as heuristics. Heuristics can include a combination of techniques such as computer instruction emulation, detection of code behavior, rules, malicious links, etc. In essence it is behavior-based technology. As part of these methods and rules, a heuristic engine assigns scores to information and behavior found in the file. If a score exceeds a certain level, the engine indicates that malware has been detected. The advantage of this technology is that it is by nature proactive, and in most cases does not require a definition file update. Heuristic detection, however, is not exact, meaning that the chance for a false positive (incorrect identification as malware) is greater than with signatures.

> Most antivirus products today use a combination of reactive and proactive technologies

To ensure the most accurate virus detection, most antivirus products today use a combination of reactive and proactive technologies – signature/fingerprint/hash approaches and heuristics – in order to take advantage of each one's strengths and minimize weaknesses. There is, however, a general trend towards increased use of heuristics, due to their capability of early detection as well as their speed.

Testing the proactive capabilities of an antivirus is best done with the most recent virus sets. Some testing organizations actually disable the signature update capabilities of an antivirus in order to evaluate its detection of these newer viruses (one week old) for which it has not yet downloaded signatures.

# Myth: If an antivirus has flagged or blocked a file, this file is definitely malware

## Fact: While the goal of antivirus engines is of course to block files that contain malware, the reality is that other types of files may be mistakenly blocked, even if they do not contain viruses or other threatening code. A few incorrectly blocked files (aka "false positives") are not necessarily a major cause for concern, but should be taken into account when comparing antivirus engines (in addition to detection comparisons).

Another type of non-malware file that may also trigger a malware indication is "pseudo-malware." This consists of "corrupted" malware, "garbage" files and "intended malware."

Garbage and corrupted files are generally the product of changes made to malware code over the years due to handling by antivirus engines, incorrect file replication, or other types of file manipulation. After decades of malware handling, a vast number of files are distributed across the Internet containing parts of original malware. These modified files are generally inert and not able to replicate or cause any damage, yet may trigger an antivirus which would then flag it as malware.

Intended malware files are those that may be released by malware authors before they have been fully tested, and contain bugs which prevent them from "working" as they were intended.

Pseudo-malware may appear during testing of antivirus solutions since many sites that maintain malware collections also contain these files. For this reason, many antivirus companies detect these files as malware. On the other hand, some antivirus engines do not identify these types of files as malware, since after all, there is no risk associated with them.

Since many vendors detect and flag pseudo-malware, there are often significant differences of detection levels between various antivirus products. For example, in comparing two antivirus engines, one may appear to detect many more viruses than another, but upon closer analysis of the samples used in the test, it may become clear that the "viruses" the other engine "missed" were not actually harmful at all.

**Pseudo-malware can skew test results depending on the policies of the antivirus vendor.**

In other words, pseudo-malware can have an adverse effect on detection test results, causing a competent antivirus engine to perform badly, or a deficient antivirus engine to show very good test results, (without providing very good real-world protection). For this reason it is not enough to simply compare the number of viruses detected between antivirus engines.

## Myth: Third Party evaluations of AV engines are more accurate than tests run by AV vendors

## Fact: The testing and evaluation of antivirus software has always been a challenge. Test results are, of course, dependent on the knowledge and motivation of the tester. However, much of the value of the test lies in the quality of the test set of virus samples.

Most testers for publications or other third-party testing organizations have built impressive malware collections over time. Since these testers or organizations are not, however, necessarily part of the antivirus community, they often rely on questionable sources for their samples and their test sets often contain a high percentage of old viruses. The types of files used in testing can skew test results, as described in the previous section, obscuring which antivirus engine truly has the best detection.

Anti-malware vendors have different motives for their testing than third-parties, since they typically want to ensure that their products appear in the best light. On the other hand, antivirus vendors know exactly how to test and have the means to build a variety of high quality test sets. Many vendors receive over 10 million samples per year that they can use as part of their test sets.

Since a truly objective test of antivirus products is virtually impossible, what can be done to get a sense of an engine's capabilities? A reasonable impression of the true quality of a solution requires checking how well a product does in a real environment, and in particular with *new* viruses. A good option is a site like virustotal.com,

> **Test results are dependent on the knowledge and motivation of the tester, and the test samples.**

where people submit suspicious files on a daily basis and have the files scanned by a multitude of virus scanners. If a virus is detected by just one engine, it is often a false positive. If it is detected by ten or more engines then it is probably malware. It should be noted that VirusTotal uses command line scanners, which may not fully reflect the normal operation of the products from each of the companies. Of course, the only way to know for sure that something is malware is to run it in a fully protected environment and have a qualified virus analyst verify the malicious behavior.

Companies determined to do their own testing should start with a visit to the web site of the Anti-Malware Testing Standards Organization (AMTSO, [www.amtso.org](www.amtso.org)). This organization defines guidelines and best methods and approaches for malware testing. Also, if you are doing your own independent testing and get results that look strange (e.g. the product detects only a subset of your samples), it is best to contact the antivirus vendor and ask them to assist in determining whether anything is wrong with the testing methodology or sample set. You will find that in most cases antivirus vendors will be happy to evaluate your approach and test set.

Antivirus vendors cannot provide you with a virus sample set to test since antivirus community agreements prohibit this. However, it is perfectly acceptable for antivirus vendors to receive a sample set from the public to determine whether the set is suitable for testing, and whether the testing approach needs to be modified to get more objective results.

## **Myth:** Testing an antivirus solution should be done by throwing as many viruses at it as possible

**Fact:** Whenever an antivirus product is evaluated, it is important to test with a mix of good files and malware to ensure that the product not only detects viruses, but that it also does not generate false positives, or at least a very low number. False positives can be a serious problem, in particular on desktop environments, as oftentimes those files that are falsely accused of being viruses are actually needed by the operating system. Improper removal of these system files can lead to an unusable computer requiring significant effort to restore or even total loss of data.

Testing using clean (non-malicious) files also has another important purpose and that is to test the performance and resource utilization of the antivirus solution.

Performance testing is equally as important as detection testing, since performance differences translate into actual costs – an antivirus engine that requires four servers is twice as expensive to run as another one that requires only two servers. In order to test actual antivirus performance or system impact, a test using *only clean files* should be run. The reason for this is that most antivirus engines in production spend most of their time scanning legitimate files, and only a small time scanning malware, so it is important for the antivirus engine to be able to scan such legitimate files quickly and efficiently. Typical users receive malware only occasionally, so the time an antivirus engine takes to scan infected files is minimal in comparison to its overall performance.

> **Performance testing is equally important as detection testing, since performance differences translate into actual costs**

# Conclusion

When entering into an evaluation of antivirus engines, the path to conclusive results requires patience and knowledge. The design of the testing process and the test files used will directly influence the outcome. A few key points to bear in mind:

- Use a mixture of new and old files, and a mixture of bad and good files
- Consider turning off signature updates for a few days to a week, to test the engine's proactive capabilities
- Do not just compare "number detected," but rather divide the detected files into categories like garbage/corrupted files, virus files, and clean files in order to accurately compare detection rates of various engines
- When using third-party test results, it is important to understand how those results were achieved
- Besides testing virus detection levels, it is crucial to evaluate performance and/or system utilization, since this has the largest impact on cost and overall satisfaction levels from the system
- For assistance with antivirus testing, it may seem counterintuitive to request help from an antivirus vendor(s), however these are the people who know the most about malware and product testing, and can assist in making sure the test is valid and will bring the desired results

## About the Author

Helmuth Freericks is an industry pioneer in the anti-malware field, having co-founded one of the earliest antivirus companies – Command Software – in 1984, long before computers were ubiquitous, and before computer viruses were a household concept. At Command, Mr. Freericks served in various executive roles over the years, starting as Vice President of Research and Development and Chief Technology Officer, and finally as CEO before the company was acquired by Authentium in 2002. While at Authentium, he served as CEO of Global Risc, an Authentium subsidiary, and later as CTO and then Chief Science officer at Authentium. He joined Commtouch in 2010 with the acquisition of the Command division of Authentium, and serves as General Manager, Anti-Malware Solutions, bringing with him extensive anti-malware knowledge, as well as experience in managing all aspects of software product research, design, implementation and quality assurance. Commtouch Command Antivirus is integrated into the offerings of leading vendors and service providers including Google, McAfee, Microsoft, and Websense.

## About Commtouch

Commtouch® (NASDAQ: CTCH) provides proven Internet security solutions, including messaging security, Web security and antivirus, to more than 150 security companies and service providers for integration into their solutions. Commtouch's Command Antivirus – acquired from Authentium in 2010 – utilizes a multi-layered approach to provide award winning malware detection and industry-leading performance. Commtouch's GlobalView™ and patented Recurrent Pattern Detection™ (RPD™) technologies are founded on a unique cloud-based approach, and work together in a comprehensive feedback loop to protect effectively in all languages and formats. Commtouch technology automatically analyzes billions of Internet transactions in real-time in its global data centers to identify new threats as they are initiated, enabling our partners and customers to protect end-users from spam and malware, and enabling safe, compliant browsing. The company's expertise in building efficient, massive-scale security services has resulted in mitigating Internet threats for thousands of organizations and hundreds of millions of users in 190 countries. Commtouch was founded in 1991, is headquartered in Netanya, Israel, and has a subsidiary with offices in Sunnyvale, California and Palm Beach Gardens, Florida.

Visit us: www.commtouch.com and blog.commtouch.com
Email us: info@commtouch.com
Call us: 650 864 2000 (US) or +972 9 863 6888 (International)

**comm touch**®

**Real Security. In Real Time.**