Testing anti-virus in Linux:
How effective are the solutions available for desktop computers?

Authors
Giuseppe Raffa, MSc (Royal Holloway, 2020)
Daniele Sgandurra, Huawei, Munich Research Center. (Formerly ISG, Royal Holloway.)

Abstract

Anti-virus (AV) programs are widely recognized as one of the most important defensive tools available for desktop computers. Regardless of this, several Linux users consider AVs unnecessary, arguing that this operating system (OS) is "malware-free". While Windows platforms are undoubtedly more affected by malicious software, there exist documented cases of Linux-specific malware. In addition, even though the estimated market share of Linux desktop systems is currently only at 2%, it is certainly possible that it will increase in the near future.

Considering all this, and the lack of up-to-date information about Linux-compatible AV solutions, we evaluated the effectiveness of some anti-virus products by using local installations, a well-known on-line malware scanning service (VirusTotal) and a renowned penetration testing tool (Metasploit). Interestingly, in our tests, the average detection rate of the locally-installed AV programs was always above 80%. However, when we extended our analysis to the wider set of anti-virus solutions available on VirusTotal, we found out that the average detection rate barely reached 60%. Finally, when evaluating malicious files created with Metasploit, we verified that the AVs' heuristic detection mechanisms performed very poorly, with detection rates as low as 8.3%.[a]

[a] This article is published online by Computer Weekly as part of the 2021 Royal Holloway information security thesis series `https://www.computerweekly.com/ehandbook/Royal-Holloway-Testing-antivirus-efficacy-in-Linux`. It is based on an MSc dissertation written as part of the MSc in Information Security at the ISG, Royal Holloway, University of London. The full thesis is published on the ISG's website at `https://www.royalholloway.ac.uk/research-and-teaching/departments-and-schools/information-security/research/explore-our-research/isg-technical-reports/`.

Introduction

Anti-virus (AV) software plays an important part in protecting end-users and networks from several types of malware. Consequently, installing and keeping up-to-date an AV program is widely considered an essential step in securing a vast range of computing devices regardless of the particular operating system. However, there seems to be a perception among Linux users that this OS can only be marginally affected by malicious software.

While it is undeniable that Linux desktop users are a lot less affected by malicious software compared to Windows users, Linux cannot be considered completely malware-free. There are, in fact, documented examples of Linux-specific malware infections, for example the one discovered in the 17.3 version of Linux Mint in 2016[1] Moreover, Linux systems are exposed to cross-platform threats as well, such as those coming from HTML, PDF and JavaScript.

To better understand the motivation of the evaluation presented in this paper, it is also important to emphasize that Linux is the most widely used OS for server computers. Statistics recently published by the SANS Institute, in fact, estimate that approximately 66% of the Internet web servers run Linux or another UNIX-based OS[2]. This is certainly a key feature of the modern IT industry, as it explains why

[1] M. Casserly, Tech Advisor Website, Does Linux need antivirus?, June 2018, `https://www.techadvisor.co.uk/feature/linux/does-linux-need-antivirus-3678945/`

[2] M. Koch, An Introduction to Linux-based malware, SANS Institute Information Security Reading Room, 2015, `https://www.sans.org/reading-room/whitepapers/malicious/introduction-linux-basedmalware-36097`

all the major companies selling Linux-compatible security products have been almost entirely focused on developing solutions for server applications rather than desktop systems.

Another factor that should be considered is the future of Windows. As Android and iOS platforms grow in popularity, the number of its users appears to be decreasing. In this evolving scenario, it is not therefore unreasonable to foresee an increase in Linux desktop systems over the next few years, even though their current market share is estimated to be only at 2%.

Considering the lack of up-to-date information about Linux-compatible AV solutions, we carried out our evaluation by performing a series of tests that aimed at:

- Measuring the detection rate of anti-virus programs installed in virtual machines and available through a well-known on-line malware scanning service.
- Assessing the AVs' effectiveness over a period of time to find out whether they are affected by regression. When this happens, an anti-virus is no longer capable of detecting malware samples that were successfully identified in the past.
- Evaluating the more advanced methods that nowadays AVs include to improve the standard signature-based detection mechanism. These methods are known as *heuristic* and in most cases rely on the identification of behavioural patterns.

Testing AVs with virtual machines

The first set of tests aimed at assessing the effectiveness of four anti-virus programs (ClamAV, Comodo, Dr Web and ESET NOD32). Installing several AVs on a given computer can cause problems though, as these pieces of software are very likely to interfere with each other. The first challenge of this work was therefore to prepare a suitable testing environment that allowed measuring the effectiveness of the AVs, but without requiring separate hardware platforms.

The standard solution to this issue is to use *virtualization*. This means creating one or more virtual machines, also known as *guests*, within a *host* OS, where a special software, called *hypervisor*, is available. The main features of our testing environments are shown in Table 1.

Virtual Machine

Emulated computer system with its own CPU, memory, network interface and storage, created on a physical hardware system by using specialized software.

Table 1: Summary of the host and guest systems features.

Host System Feature	Value	Guest System Feature	Value
Hypervisor	VirtualBox 6.1	Memory	2,048 MB
Memory	16 GB DDR4 2,400 MHz	Operating System	Ubuntu Linux 18.04 LTS
Operating System	Ubuntu Linux 18.04 LTS	Virtual Hard Disk Size	32 GB

It is also important to note that we originally considered a larger set of AVs, however, we decided to exclude some of them for the following two reasons:

- *Highly specialized products.* For example, Chkrootkit and Rootkit Hunter are well-known scanners, but they are server-oriented and routinely used by system administrators to detect only specific types of malware.
- *Discontinued products.* Several AV vendors, such as AVG, Avast, Bitdefender, F-Prot and Zoner, have decided to develop licence-protected products exclusively for Linux servers.

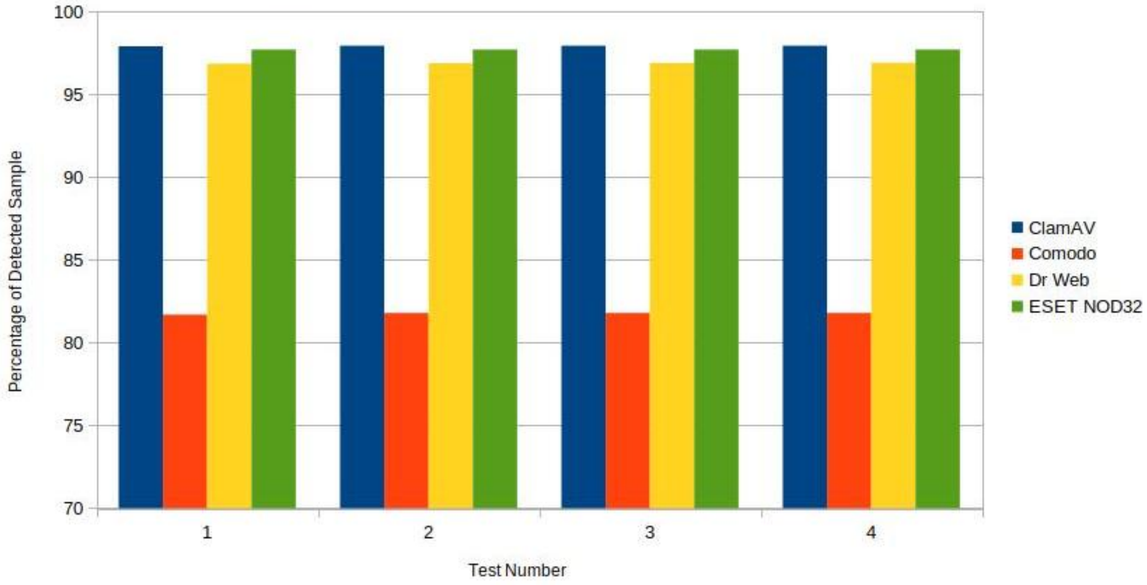

Figure 1: Percentage of detected samples vs test number

As far as the test methodology, recent studies[3] have stressed the importance of measuring the detection rate of AVs several times during an observation period to assess their update mechanisms. Furthermore, this approach enables the identification of regression effects, which are caused by the signature database updates as well as changes to the adopted heuristics.

We focused our analysis on a recent archive, downloaded from VirusShare, of malicious Linux Executable and Linkable Format (ELF) files, which included 43,553 samples. The selected AVs were then tested by executing four scans over the course of three weeks. Each scan was run after updating the anti-virus signature database.

As illustrated in Figure 1, the average detection rate ranged from 81.8% for Comodo to 97.9% for ClamAV. However, it is important to note that none of the AVs showed a 100% detection rate, in spite of the fact that our tests were carried out more than ten weeks after the malicious ELF files were made available in the on-line repository VirusShare.

Detection Rate

Percentage of malicious files correctly identified as malware by an anti-virus.

In addition, as shown in Figure 2, since we used more than 43,000 malware samples, even when the detection rate was high, the total amount of undetected malicious files was significant. For example, ClamAV, which was the AV with the best detection rate, did not detect on average 896 specimens, whilst the AV with the worst detection rate (Comodo) did not flag as malicious nearly 8,000 samples during the last test.

Two additional findings are worth mentioning:

[3] M. Botacin, F. Ceschin, P. de Geus, A. Grégio, We need to talk about antiviruses: challenges & pitfalls of AV evaluations, Computers & Security, Volume 95, August 2020, 101859.

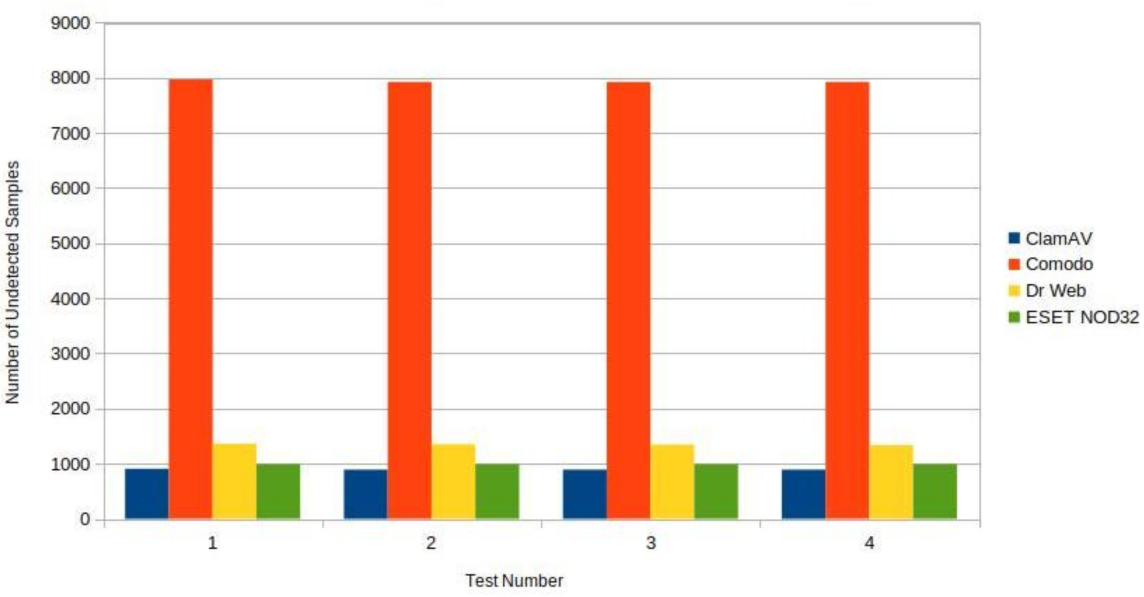

Figure 2: Number of undetected samples vs test number

- Since all the scans were executed after updating the AV signature databases, it was legitimate to expect a steady increase over time of the detection rates. However, only Dr Web showed this behaviour, though the detection rate for the last test (96.9%) was only 0.1% higher than the one recorded during the first. By contrast, the number of files flagged as malicious by ESET NOD32 remained constant during the observation period.
- As detailed in Figure 2, the number of undetected samples either remained constant or decreased. Therefore, we concluded that none of the tested products was affected by regression.

It is finally worth emphasizing that, differently from other Windows-focused studies, it was not possible to compare the figures presented in this section with those reported by specialized websites, e.g. AV Test and AV Comparatives, as they did not provide any information about Linux-compatible AV solutions.

Testing AVs with an on-line scanning service

The aim of the second set of tests was to evaluate the effectiveness of a larger number of AVs by exploiting an on-line malware scanning service. These are very useful tools for security researchers, as they allow testing malicious samples with several AVs by using a web-based interface or a scripting language. In particular, VirusTotal, which included 62 anti-virus products when this research was conducted, offers an Application Programming Interface (API) that we used to automatically scan 4,000 malware specimens from the VirusShare repository.

To ensure consistency with the evaluation of the locally-installed software, the malicious files were submitted twice over a period of two weeks to identify possible regression effects. In general, the VirusTotal AVs did not perform as well as the locally-installed products. The average number of detected malware samples was, in fact, approximately 2,395 out of 4,000. A further breakdown

Regression

An anti-virus software is affected by regression when it is no longer capable of detecting malware samples that were successfully identified in the past.

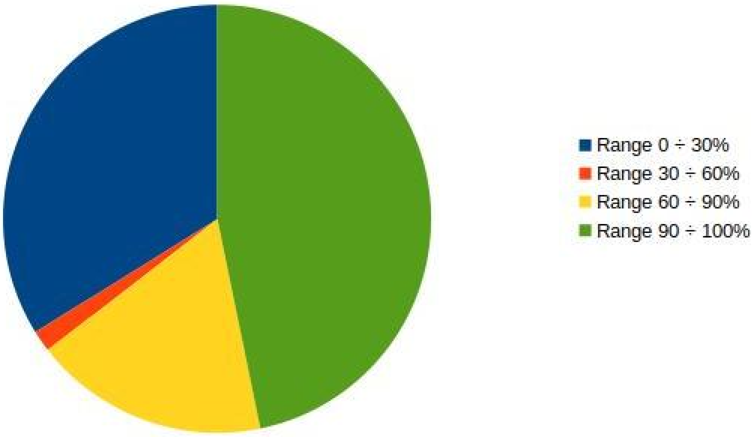

Figure 3: VirusTotal AVs average detection rate distribution (VirusShare malicious files)

of the detection rate figures, which is summarized in Figure 3, shows that nearly 50% of the VirusTotal AVs had a detection rate above 90%. This means that the overall average detection rate was significantly affected by a rather high number of AV programs that performed very poorly.

It is important to note that other similar studies have reported discrepancies between results obtained with locally-installed AVs and those provided by web-based services[4]. Although such inconsistencies can be attributed to inevitable differences in AV and signature database versions, no Linux-specific information appeared to be available in the literature. Therefore, the detection rates for the four locally-installed AVs were recomputed by considering the 4,000 malware samples tested with VirusTotal. The comparison of the results of local and on-line AVs showed only small discrepancies, with the largest being 1.9% (ClamAV).

Finally, differently from the tests presented in the previous section, our data analysis identified 13 AVs affected by regression, as the total number of files flagged as malicious during the second test was lower in comparison to the first.

Targeting AV-protected systems with a pentesting tool

To comprehensively assess both the locally-installed and the VirusTotal AVs, we tested them against malware samples created with the well-known pentesting (i.e., penetration testing) tool Metasploit. There are two reasons behind this strategy. Firstly, since this tool is open-source and widely used, we aimed at verifying that fully updated AV signature databases promptly include the information required to detect Metasploit-originated malicious files. Secondly, the goal of these tests was to measure the effectiveness of the AVs against crafted evasive malware samples created via Metasploit.

Signature

A virus signature can be informally defined as the fingerprint of a virus. More precisely, signatures are known patterns, e.g. unique data or portions of code, that are used by anti-virus software to identify malicious files.

It is important to observe that the latter includes Linux-compatible malware samples called *payloads*. We focused our attention on those creating a *reverse shell*, which is a technique commonly used to access a target system through a network connection started by the victim. This is a critical feature because:

[4] M. Botacin, F. Ceschin, P. de Geus, A. Grégio, We need to talk about antiviruses: challenges & pitfalls of AV evaluations, Computers & Security, Volume 95, August 2020, 101859.

- It enables the execution of the payload on machines that cannot be directly reached by the attacker.
- Network defences, such as firewalls, are routinely set up to prevent connections that start from outside the protected network. However, they far less frequently block outbound connections.

The considered attack model crucially relies on the fact that the malware samples are first made available on the target system, for instance through malicious websites or infected media drives, and then executed by the victim, which could be achieved via social engineering.

After combining six payloads with some Metasploit *encoders*, which are pieces of software capable of modifying the payloads to reduce the chance of detection, we created 24 malware samples. These were used to run three sets of tests. The first two were carried out by scanning the malicious files with the locally-installed AVs and VirusTotal, respectively. By contrast, during the third set of tests, the malware samples were executed within the AV-protected virtual machines to evaluate the effectiveness of the AVs' heuristic detection mechanisms.

The test results can be summarized as follows:

- *Scans with locally-installed AVs.* During these tests, the best detection rate was 41.7% (ESET NOD32) and two AV programs reported as malicious only two files out of 24 (Table 2). Surprisingly, eight samples were not detected by any of the tested anti-malware solutions, while only four malicious files were flagged more than once.
- *Scans with VirusTotal.* The results of these tests were surprisingly low as well, with an average detection rate of 16.9%. Although one AV was able to flag as malicious all the submitted samples, 32 AV programs detected no malicious file and the majority of them had a detection rate lower than 30% (Figure 4).
- *Execution within AV-protected virtual machines.* In these last tests, no AV product was able to detect malware samples that had not already been flagged during the scans with the locally-installed software. ClamAV and Dr Web, in fact, reported as malicious exactly the same files, whilst Comodo did not block any of the samples, despite having previously detected two of them. Finally, ESET NOD32 prevented the execution of only four out of the ten files initially flagged as malicious.

Table 2: Summary of AVs performance indicators (Metasploit malicious files)

AV Product	Total Number of Detections	Detection Rate (%)
ClamAV	6	25
Comodo	2	8.3
Dr Web	2	8.3
ESET NOD32	10	41.7

Conclusion

The main objective of this work was to assess the effectiveness of AV software solutions currently available for Linux desktop installations. This was achieved by evaluating both their signature-based and their heuristic detection mechanisms.

Over the course of three weeks, we have tested four locally-installed AVs against a repository of more than 43,000 malicious ELF files, to measure their detection rate and regression effects as well as to assess the efficacy of their update mechanisms. We have found that the average detection rate of the tested products was always well above 80% and that none of them was affected by regression.

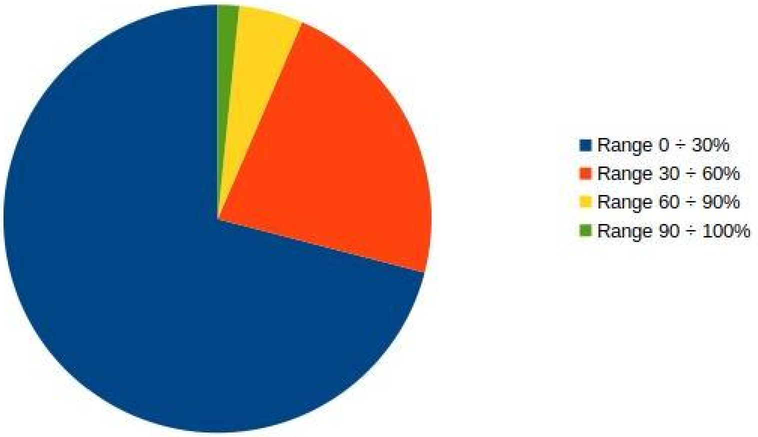

Figure 4: VirusTotal AVs average detection rate distribution (Metasploit malicious files)

Unexpectedly, though, only one of the locally-installed AV programs showed a steady increase over time in the number of detected malware samples.

We have also used the on-line malware scanning service VirusTotal to compare the effectiveness of 62 anti-virus products. This was achieved by using 4,000 malicious files, which were submitted twice over the course of two weeks. While the average detection rate of the on-line AVs barely reached 60%, nearly half of them featured a detection rate above 90%, with one third of AVs showing less than 30%. Interestingly, 13 out of 62 anti-virus products showed regression effects.

In addition, we have created 24 malicious files by using the penetration testing tool Metasploit. They were scanned and then executed to determine the effectiveness of the AVs' heuristic detection mechanisms. Contrary to expectations, the detection rate was as low as 8.3% and no anti-virus program was able to block the execution of samples that had not already been flagged during the initial scan. Our customized malware specimens were submitted to VirusTotal as well. The results show that the average detection rate was only 16.9% and that 32 out of 62 AVs did not report as malicious any of the submitted files.

The tests presented in this paper show that, while detection rates above 90% are achievable, the signature database update mechanisms should be reviewed and improved. Although only a minority of the tested AVs was affected by regression, the detection rates increased over time very marginally. Finally, another area where the considered AV solutions underperformed is heuristic detection, which did not provide any additional layer of protection to the end-user.

Biographies

Giuseppe Raffa holds a MSc in Electronic Engineering from the University of Pisa, Italy, and the MSc in Information Security from Royal Holloway. After spending more than fifteen years in the automotive industry developing Hardware-In-the-Loop systems for engine control units, he is continuing his studies within the RHUL Centre for Doctoral Training. He is mostly interested in penetration testing, malware analysis and development of innovative tools for cybersecurity.

Daniele Sgandurra was a Senior Lecturer in Information Security at Royal Holloway in the Information Security Group (ISG). His research interests are related to practical aspects of systems and software security, by focusing on attacks and defences at various architectural layers, as well as on threat detection. His recent work investigates the arms-race between malware and anti-malware systems and the usage of AI/ML within cyber-security. Daniele is now with Huawei at the Munich Research Center.

Series editor: S.- L. Ng