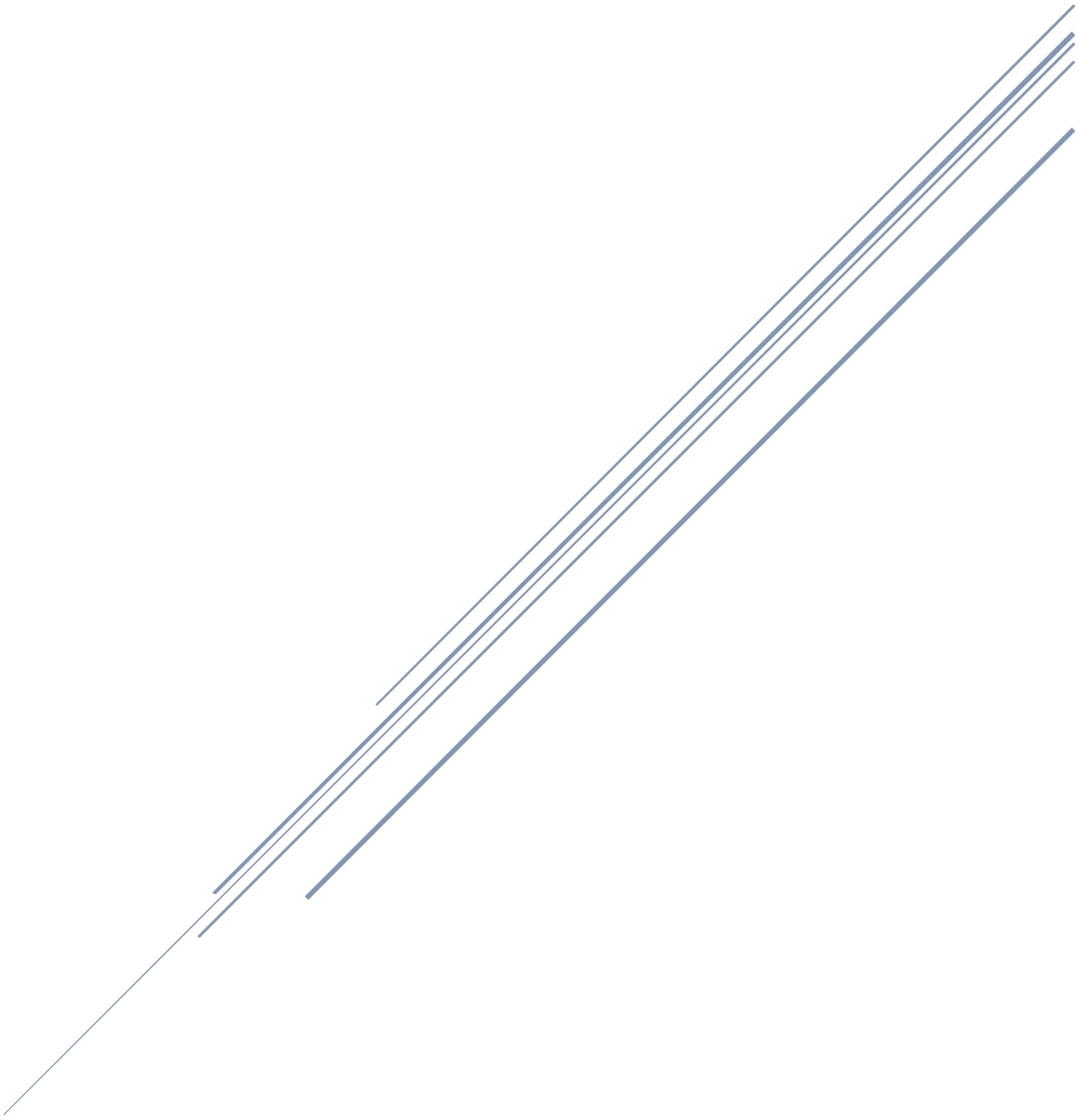
MRG Effitas
Efficacy Assessment & Assurance

# COMPARATIVE MALWARE PROTECTION ASSESSMENT

February 2018



2018.02.28

MRG Effitas
Efficacy Assessment & Assurance

# Table of Contents

# 1 Introduction

## 1.1 Executive summary

This report provides an independent comprehensive comparative assessment of enterprise endpoint protection products. In this assessment, we focused on executable malware. We used a wide spectrum of tests to cover advanced zero day threats that enterprise environments face.

This report contains the results of four test cases. The primary goal was to show the detection and prevention capabilities of new and unknown malicious executables.

The different test cases were a real world malware protection test, unknown/zero day malware protection test, longevity/holiday test and false positive tests.

In the malware protection test, we downloaded malware samples from URLs, and attempted to execute the samples. We also tested whether the product was able to late-block the malware, where the malware starts, but it is blocked at a later time during the test.

The unknown/zero day malware protection test was similar to the real world malware protection test, but in this test we selected samples which were not yet known to public malware file sharing services.

The longevity/holiday test simulated a user who was on vacation for 2 weeks, did not install any virus definition updates, and starts to browse the web two minutes after starting the machine, which was in sleep mode. This test provides insight into the decay rate of the protection effectiveness. Observing markedly lower protection rates in this test can indicate that the protection methods are very dependent on strict signature detection, or if machine learning is used that the detection model may be 'over fit' and not resilient when faced with new malware that is markedly different from what may have been common just a few days prior.

The malware protection, zero day test and longevity test were joined together in one chart called ITW test (in-the-wild malware). The Potentially Unwanted Applications (PUAs) are shown in a different chart.

For the false positive test, we collected a vast number of clean samples. We also focused on collecting rare and new samples, which are probably not yet known to vendors.

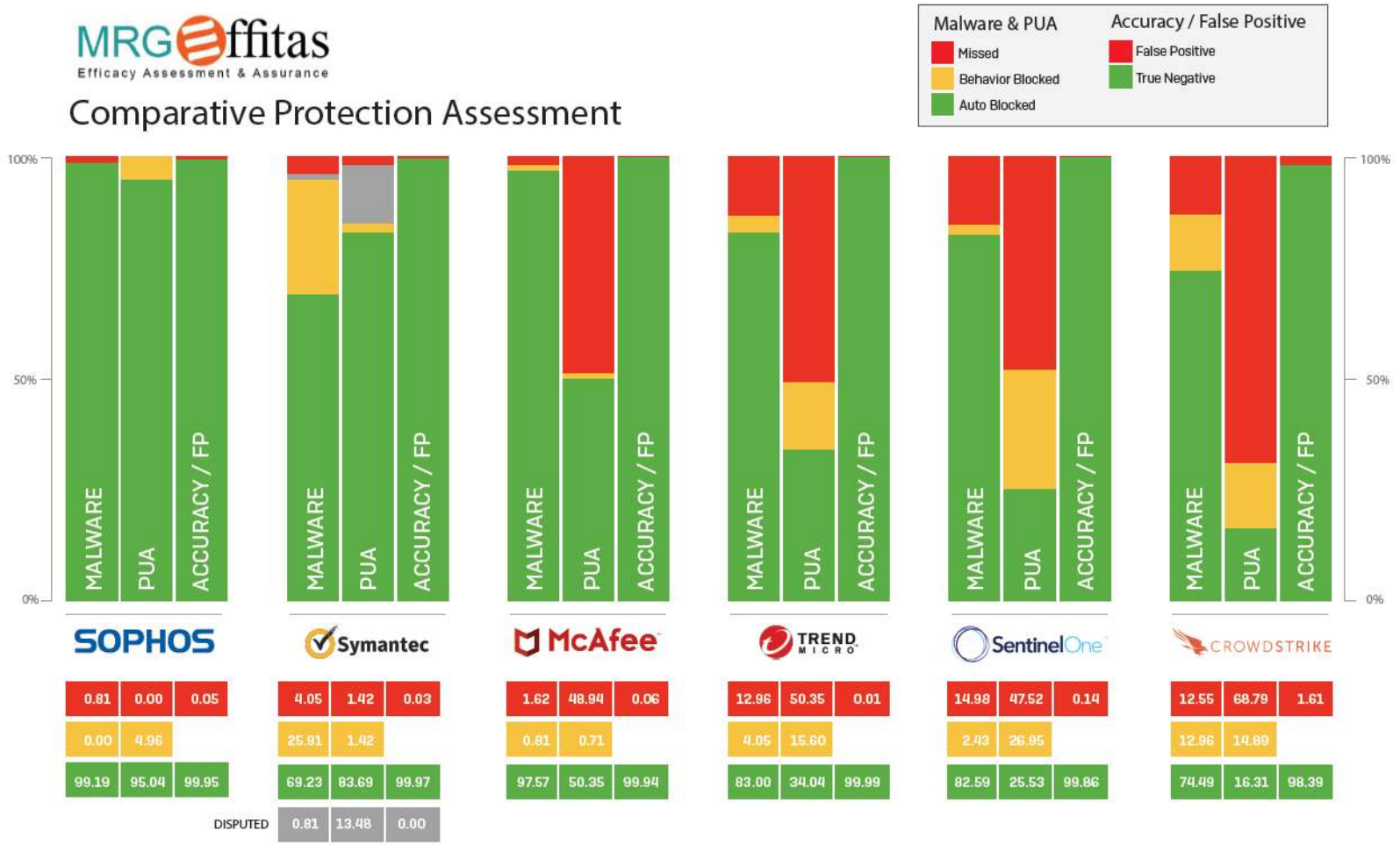
**Final results**

Based on the in-the-wild malware tests, Sophos Intercept X with Endpoint Advanced performed the best.

In the PUA test, Sophos Intercept X with Endpoint Advanced performed the best.

In the False Positive test, Trend Micro Smart Protection performed the best.

MRG Effitas
Efficacy Assessment & Assurance




Source: https://www.mrg-effitas.com/wp-content/uploads/2018/02/MRG_Comparative_2018_February_report.pdf

## 1.1 Test details

The target platform was Windows 10 64-bit, with Internet Explorer 11.

The test did not focus on the detection of non-PE files - things like documents, images and media files, or on the delivery vectors used to target a device, like email, web browsing, or direct exploit attacks on exposed network ports. The test was performed in January and February of 2018.

### 1.1.1 Malware protection test

Sample selection is of fundamental importance to this and all similar tests. Around 75% of the samples used were "live" and "in-the-wild", by which we mean that they reside at the URLs selected or created by the cybercriminals, and not from a time-lagged ITW list. As these are live ITW samples, they represent current zero-day threats that can be an issue with sample verification. For the remaining 25% we used an internal website to download the threats in order to simulate attackers changing to new URLs, and to test scenarios where a malicious URL is not used (e.g. hosted internal webmail).

There is no effective and reliable way to verify samples before testing that does not introduce possible artificial sample submission or delay, so all verification is conducted after testing. Tests performed using samples that are later proven to be invalid are excluded from the results. MRG Effitas selects the type of samples to be used, based on the following criteria:

1. Prevalence – they are widespread and therefore represent the most common threats.
2. Growth – they may be few now, but our research shows they are rapidly expanding.
3. Innovation – they employ innovative techniques to counter security measures.

In our test, the samples were downloaded first to a cache proxy. When a sample was found to be valid for the test, the cache proxy was used. The advantage of this is that even if the original URL is not stable or does not serve the same malware between sessions, the same sample is delivered to all the products during the test.

We collected live samples of in-the-wild financial malware, ransomware, PUA, Remote Access Trojans (RAT), coinminers and rootkits, and started the malware on an already protected system. Following these principles we used 388 samples in total.

### 1.1.2 Real-world Unknown/'zero-day' Malware

In this test we used malware samples that have not yet been submitted to public malware file-sharing services (used for analysis and collection by security vendors) and are less than 24 hours old. The test included 113 portable executable (PE) files from this test set.

### 1.1.3 Longevity (holiday) test

We tested a scenario where a person is on holiday and then comes online after a break. The offline period of the environment was two weeks. The time between the system resuming after sleep and the malware starting was two minutes, and cloud and network connectivity was provided. 42 samples from the overall test were tested in this way. This test provides insight into decay rate of the protection effectiveness. Observing markedly lower protection rates in this test can indicate that the protection methods are very dependent on strict signature (reactive) detection, or if machine learning is used that

the detection model may be 'over fit' and not resilient when faced with new malware that is markedly different from what may have been common just a few days prior.

### 1.1.4 FP rate for common legitimate software

We also determined the FP rate for common legitimate software. Common software is software published over one year ago by a legitimate software publisher, and is not classified as a PUA. The legitimate software includes samples that were created by a variety of large and small vendors, and includes both signed and unsigned software. Our test size was 10,000 samples.

Our plan was to use static scanning on all the samples, but some vendors do not support static scanning of a folder. For these products, the sample was started by a script.

### 1.1.5 FP rate for new and rare legitimate software

In this test we determined the FP rate for new and rare legitimate software. The software had been published less than one week prior to the test and was not from a common software manufacturer. Our test size was 100 samples.

The two false positive-test results are joined in one graph.

## 1.2 Test environment

The test environment was Windows 10 64-bit.

Our original plan was to use Windows 10 64-bit Fall Creators (1709) update for all the vendors. Unfortunately, at the time of the start of the test, not all vendors supported the Windows 10 64-bit Fall Creators update. For these products, we used the Windows 10 64-bit November (1511) update.

For the test, our plan was to use the Edge browser, but unfortunately not all vendors supported protecting Edge. Actually, some vendors even block the start of the Edge browser. Because of this, all products were tested with the Internet Explorer 11 browser.

We used Virtualbox as a virtualization platform. Our test system is hardened against virtualization /sandbox aware malware, and malware can't detect the presence of a test/sandbox/virtualization system.

4 Gbyte of RAM, 2 processor core, a Gigabit Ethernet and 60 Gbyte of disk space was dedicated for the OS and the security product.

## 1.3 Tested Software

All tested products had current updates (except in the holiday test), and were connected online to simulate the scenario where the end user is most likely to encounter malware delivery mechanisms, like browsing, emails and active adversary attacks from other devices. All products were using a default policy.

List of tested software:

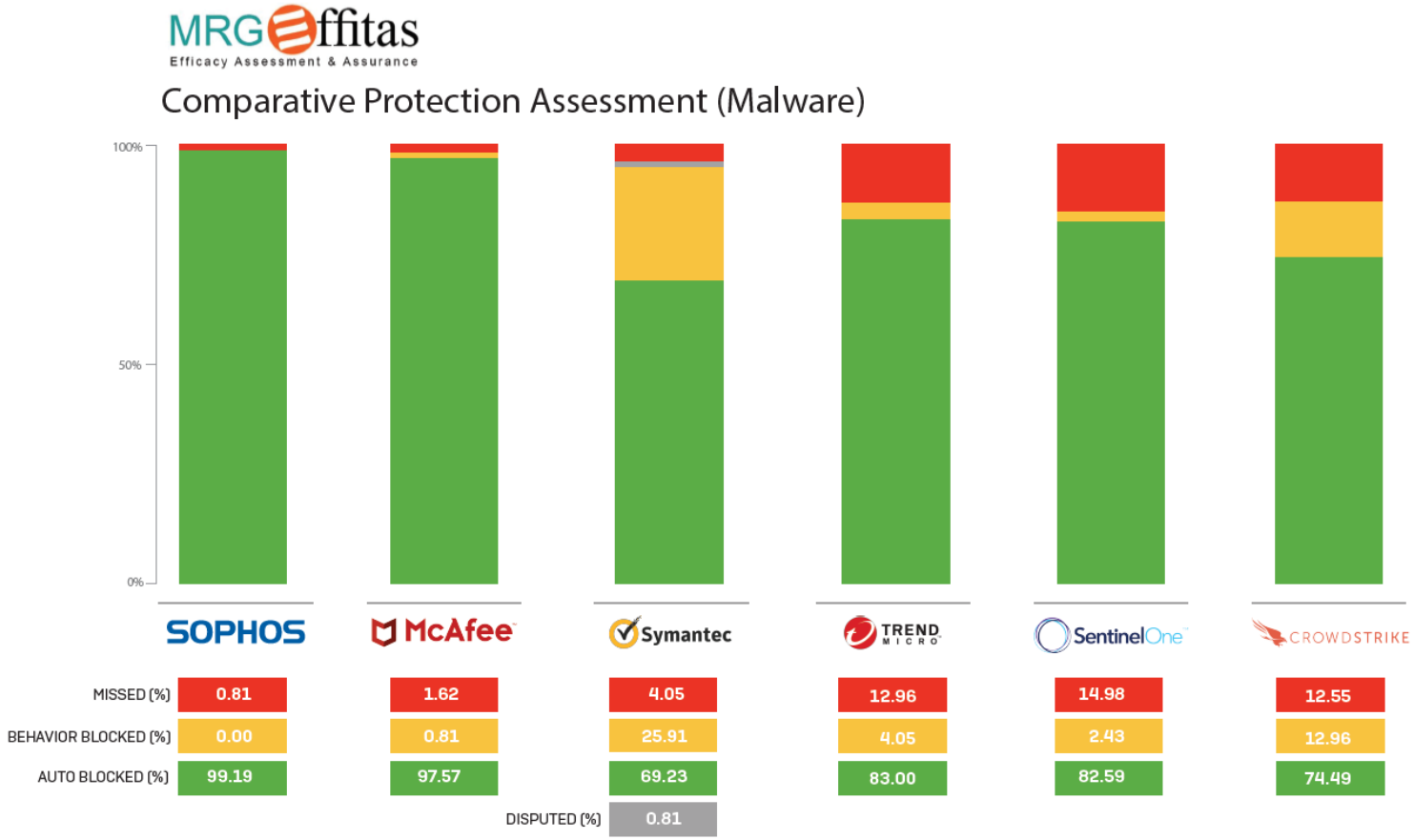
- CrowdStrike Falcon Prevent 3.8.5907.0
- McAfee Endpoint Threat Protection (Threat Prevention module) 10.5.2.2072[1]
- SentinelOne Endpoint Protection 1.8.4.3694
- Sophos Intercept X 2.0.0 with Endpoint Advanced (EA 10.8.1)
- Symantec Endpoint Protection 14.0.3752.1000
- TrendMicro Smart Protection for Endpoints 13.943.00

[1] The plan was to test McAfee **Complete** Endpoint Threat Protection but we could not get a license for it. Using the Complete version instead of the standard version could have changed the results

# 2 High-level results

## 2.1 In-the-wild malware protection results

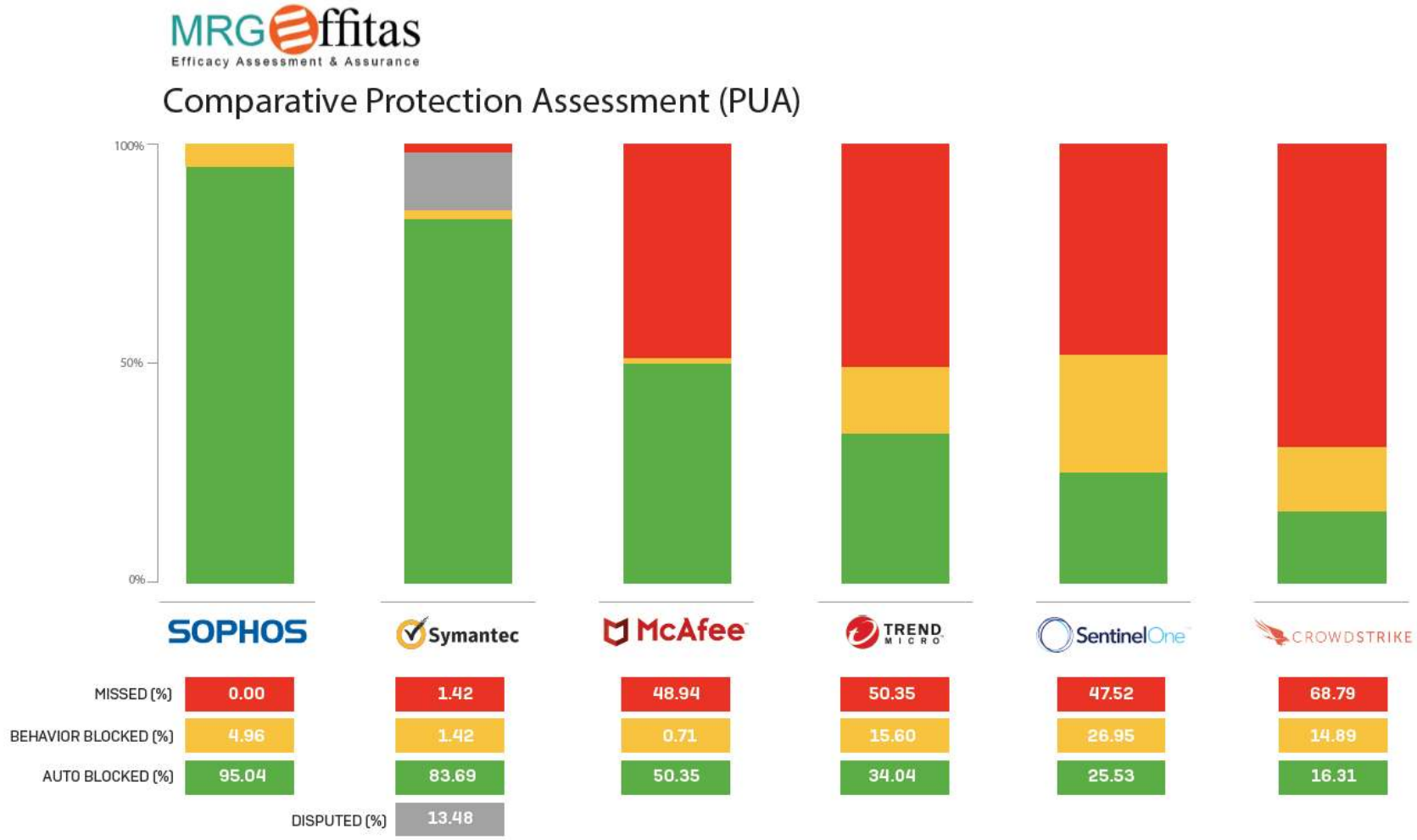
- Detect known PE malware and potentially unwanted applications
- Detect real-world ‘zero-day’ PE malware that has not been submitted to public malware sharing services
- Longevity of protection provided – back from holiday after two weeks scenario




Source: https://www.mrg-effitas.com/wp-content/uploads/2018/02/MRG_Comparative_2018_February_report.pdf

## 2.2 PUA protection test results

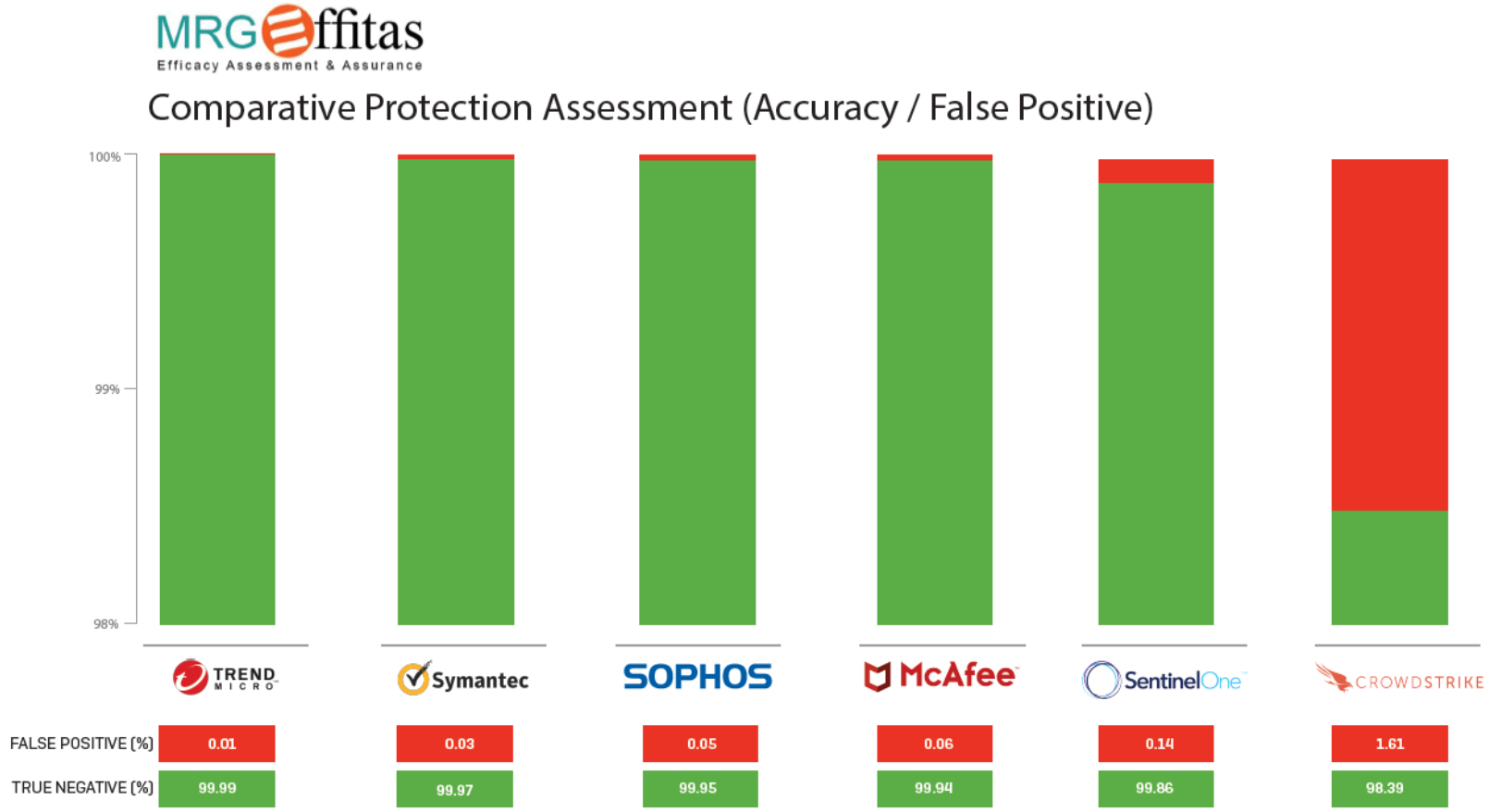
The following chart includes the test samples where the sample is a Potentially Unwanted Application (PUA), adware or a PUA coinminer. During this test, we saw more coinminer samples than ever before.




Source: https://www.mrg-effitas.com/wp-content/uploads/2018/02/MRG_Comparative_2018_February_report.pdf

## 2.3 False Positive test results

- Avoid detecting known legitimate software as malicious or potentially unwanted
- Avoid detecting new and rare legitimate software as malicious or potentially unwanted




Source: https://www.mrg-effitas.com/wp-content/uploads/2018/02/MRG_Comparative_2018_February_report.pdf

# 3 Vendor feedback

Before this assessment was started, all the vendors in the original cohort were contacted and notified that their product had been proposed to be included.

Of all the vendors contacted, the following agreed to be voluntary participants:

- Sophos
- Symantec

The following vendors stated they did not want to participate:

- Cylance (with CylanceProtect)
- McAfee (with Complete Endpoint)
- Microsoft (with Defender ATP)
- Palo Alto Networks (with Traps)

Effitas is unable to elaborate as to why some vendors who requested not to participate were removed from the cohort and why some were not.

Symantec also disputed 17 samples, saying that “these samples were downloaded from an internal IP address. This doesn't give our URL reputation technologies a fair chance at evaluating all contextual information surrounding the threat. Threats are usually downloaded externally on the Internet.” These test results were marked differently in the report.

# 4 Conclusion

Based on the in-the-wild malware tests, Sophos Intercept X with Endpoint Advanced performed the best.

In the PUA test, Sophos Intercept X with Endpoint Advanced performed the best.

In the False Positive test, Trend Micro Smart Protection performed the best.

# 5 Test methodology

Methodology used in the assessment:

1. Windows 10 64 bit operating system was installed on a virtual machine, all updates were applied and third party applications installed and updated according to our "Average Endpoint Specification"[i]
2. An image of the operating system was created.
3. A clone of the imaged systems was made for each of the security applications used in the test.
4. An individual security application was installed using default settings[ii] on each of the systems created in step 3 and then, where applicable, updated.
5. A clone of the system as at the end of step 4 was created.
6. Each live URL test was conducted by:
    a. Downloading a single malicious binary from its native URL using Microsoft Internet Explorer to the desktop, closing Microsoft Internet Explorer and then executing the binary.
7. The system under test was deemed to have been protected if:
    a. The security application blocked the URL where the malicious binary was located.
    b. The security application detected and blocked the malicious binary whilst it was being downloaded to the desktop.
    c. The security application detected the malicious binary when it was executed according to the following criteria:

        It identified the binary as being malicious and either automatically blocked it, or postponed its execution and warned the user that the file was malicious and awaited user input.
8. The system under test was deemed to have been infected if:

    The security application failed to detect or block the binary at any stage in step 6 and allowed it to be executed. If the malware was detected and fully or partially blocked at a later stage, we marked this as a behavior block.
9. Testing was conducted with all systems having internet access.
10. Each individual test for each security application was conducted from a unique IP address.
11. All security applications were fully-functional unregistered versions or versions registered anonymously, with no connection to MRG Effitas.
12. All testing was conducted during January and February, 2018.
13. As no user-initiated scans were involved in this test, applications relied on various technologies to detect, block and remediate threats. Some of these technologies were: background scanning, startup scanning, scheduled scanning, system monitors, etc. A scheduled scan was used only if enabled by default.

# 6 About MRG Effitas

MRG Effitas is a UK-based, independent IT security research organisation that focuses on providing cutting-edge efficacy assessment and assurance services, the supply of malware samples to vendors and the latest news concerning new threats and other information in the field of IT security.

MRG Effitas' origin dates back to 2009 when Sveta Miladinov, an independent security researcher and consultant, formed the Malware Research Group. Chris Pickard joined in June 2009, bringing expertise in process and methodology design, gained in the business process outsourcing market.

The Malware Research Group rapidly gained a reputation as the leading efficacy assessor in the browser and online banking space and, due to increasing demand for its services, was restructured in 2011 when it became MRG Effitas, with the parent company Effitas.

Today, MRG Effitas has a team of analysts, researchers and associates across EMEA, UATP and China, ensuring a truly global presence.

Since its inception, MRG Effitas has focused on providing ground-breaking testing processes and realistically modelling real-world environments in order to generate the most accurate efficacy assessments possible.

MRG Effitas is recognized by several leading security vendors as the leading testing and assessment organisation in the online banking, browser security and cloud security spaces and has become their partner of choice.

Our analysts have the following technical certificates:

Offensive Security Certified Expert (OSCE), Offensive Security Certified Professional (OSCP), Malware Analysis (Deloitte NL), Certified Information Systems Security Professional (CISSP), SecurityTube Linux Assembly Expert, SecurityTube Python Scripting Expert, Powershell for Penetration Testers, Certified Penetration Testing Specialist (CPTS), Computer Hacking Forensics Investigator (CHFI), and Microsoft Certified Professional (MCP).

[i] AES includes Adobe Flash, Reader, Java, Microsoft Office 2010, IE & VLC Player. All Microsoft components were fully updated; all third-party components were out of date by three months.
[ii] During installation of the security application, if an option to detect PUAs was given, it was selected.