MRG Effitas 360 Degree Assessment & Certification

Q4 2016

Contents

Introduction

MRG Effitas has a core focus on efficacy assessments in the anti–financial fraud space, but we also publish more traditional "Real World" detection tests. An example of such a test is our "Time to Detect Assessment Q4 2013" (Project 37).

This assessment measured the ability of security products to protect an endpoint from a live infection, and, in the event of a system being compromised, the time taken to detect the infection and remediate the system. The time-to-detect-and-remediate component relied on each security product being manually forced to conduct a scan every thirty minutes over a 24-hour period.

For 2014, it was decided that a new approach was needed as the methodology applied in previous tests did not reflect how a security product would be used on an endpoint in the Real World. In practice, many security applications will only detect an infection during a reboot/startup or if a scheduled scan has been set by default.

For this assessment, time-to-detect will employ a methodology based on the infected endpoint being re-scanned once during a 24-hour period.

The methodology employed in this test maps more closely to Real World use, and although it may not be a 100% accurate model of how an "average" system is used, it gives a more realistic assessment of a security product's ability to detect and remediate an infected endpoint.

This Programme is called a "360 Assessment" since it deals with the full spectrum of malware instead of just financial malware. In the 360 Assessments, trojans, backdoors, ransomware, PUAs, financial malware and "other" malware are used.

Executive Summary

This Certification Programme is designed to serve as a reflection of product efficacy based on what we have previously termed "metrics that matter".

In many of our previous tests, particularly those that have focused on financial malware, we started with the assumption that the endpoint has already been compromised. Being the world's largest supplier of early-life malicious binaries and malicious URLs, and from our own simulator development, we know that all endpoints can be infected, regardless of the security solution employed.

For us, a product's ability to block initial infection (although critical in most cases) is not the only metric that matters. One also needs to measure the time taken for the security product to detect malware on a system and remediate it.

When conducting these tests, we tried to simulate normal user behaviour. We are aware that a "Real World" test cannot be conducted by a team of professionals inside a lab because we understand how certain types of malware work, how malware attacks and how such attacks could be prevented. Simulating normal user behaviour means that we paid special attention to all alerts given by security applications. A pass was given only when alerts were straightforward and clearly suggested that malicious action should be blocked.

We tested a group of internet security suites and complementary security applications. With these, it is very important to note that the best choice for an average user is to keep things very simple and for the product not to present many pop-up alerts or questions.

Out of seventeen products we tested, thirteen managed to meet the specification to attain our Q4 2016 360 certification award, these being **avast! Internet Security, Avira Internet Security, Bitdefender Internet Security, ESET Smart Security, Kaspersky Internet Security, Panda Internet Security, SurfRight HitmanPro, Symantec Norton Security, ThreatTrack Vipre Internet Security, Trend Micro Maximum**, **Watchdog Anti-Malware, Webroot SecureAnywhere, Zemana Anti-Malware.**

All other security applications failed the test in that they were unable to detect the malware and/or remediate the system even after the end of a 24-hour period.

Certification

In order to attain a quarterly MRG Effitas 360 Degree certification award, a security application must either protect the system from initial infection (behaviour protection or block within 24 hour) (a level 1 pass) or in at least 97% of all cases detect any missed malware and fully remediate the system before or on the first user initiated rescan (a level 2 pass). Applications that meet this specification will be given certification for that quarter.

Under the MRG Effitas 360 Degree Assessment & Certification, the following products were certified for Q4 2016:

Certified (level 1): **avast! Internet Security, Kaspersky Internet Security**

Certified (level 2): **Avira Internet Security, Bitdefender Internet Security, ESET Smart Security, Panda Internet Security, SurfRight HitmanPro, Symantec Norton Security, ThreatTrack Vipre Internet Security, Trend Micro Maximum**, **Watchdog Anti-Malware, Webroot SecureAnywhere, Zemana Anti-Malware**

The Purpose of this Report

Since its inception in 2009, MRG Effitas has strived to differentiate itself from traditional testing houses by having its primary focus on providing "*efficacy assessments"* and not just performing *"tests"*.

Traditionally, testing of security software has centred on measuring a product's ability to detect malware. Testing has evolved rapidly over the last two to three years as most labs, under the guidance of AMTSO (of which MRG Effitas is a member) strived to conduct "Real World" testing.

Although there is no absolute definition of this kind of testing, loosely speaking, it involves the introduction of malware to an endpoint through a realistic vector, such as a browser or USB memory stick. Real World testing mostly involves "dynamic testing" (i.e. the malware is executed and then the ability of the security product to block the malware is measured).

Several testing labs also conduct "System Rescue" tests. These assess a security product's ability to remediate a pre-infected endpoint.

Whilst both types of tests are useful and yield valid and meaningful data, MRG Effitas wanted to merge these tests and also go one step further by measuring the time security products take to detect infections and remediate the endpoint.

To make testing more akin to Real World scenarios, no manual scanning was conducted. Instead, the system was re-scanned once a day (exactly 24 hours after the system was compromised), thereby giving security applications the opportunity to detect infections on restart.

As we have stated in our previous test reports, all malware has one primary objective, and that is to make money for the cybercriminals.

Measuring initial detection rates and the time taken to detect active malware is important, particularly in today's threat landscape with the mix of malware that is prevalent.

As we have repeated in our previous financial malware test reports, the longer a cybercriminal can have their malware on a system, the greater the opportunity for them to be able to capture private user information including banking passwords and social media credentials, etc.

There has been an increase in the prevalence of ransomware, such as "CryptoLocker", which, once active on the system, holds the user at ransom to decrypt system data or unlock the system in some other way (interestingly, the most common way CryptoLocker is installed on an endpoint is via Zeus infections).

For these types of malware, it is initial detection that is of the greatest importance, since the vast majority of security solutions will be unable to rescue an encrypted or locked system. (In other internal tests, we have found that Webroot SecureAnywhere was in fact able to undo the encryption performed by some ransomware.)

In providing these quarterly certifications, the MRG Effitas 360 Assessment & Certification Programme is the *de facto* standard by which security vendors, financial institutions and other corporations can attain the most rigorous and accurate determination of a product's efficacy against the full spectrum of malware that is prevalent during the period.

Tests Employed

In this assessment (Q4 2016), we ran the following tests:

In the Wild 360 / Full Spectrum Test

Approximately 50% of the malicious URLs used in this test were compromised legitimate websites which served malware. We believe that such URLs pose the greatest danger to users as this is the place where they least expect to get infected. 10% of the URLs pose as fake porn websites serving visitors with various types of malware. The remaining 40% of the URLs come from our regular honeypots or, in case of ransomware and financial malware in particular, we used URLs from newly-discovered distribution sites.

Malware delivered by URLs used in this test can be considered as Zero Day in the true meaning of that phrase. This posed a great challenge to all participants as new variant samples such as Locky (Ransomware) TeslaCrypt (Ransomware), Dridex (Banking Trojan) and many others caused most damage.

It is our opinion that Ransomware currently poses the greatest threat to users, for this reason we choose to use more URLs serving this threat than before.

Because of the wide spectrum of malware used in this project and the freshness of the samples, we used a smaller set than usual.

Applications that didn't protect the system from file encrypting ransomware cannot be certified because they could not remediate the threat; as files usually cannot be decrypted.

Our testing environment supports the use of VM aware malware, this is the reason why we were able to use more sophisticated threats which wouldn't run on Virtual Machines.

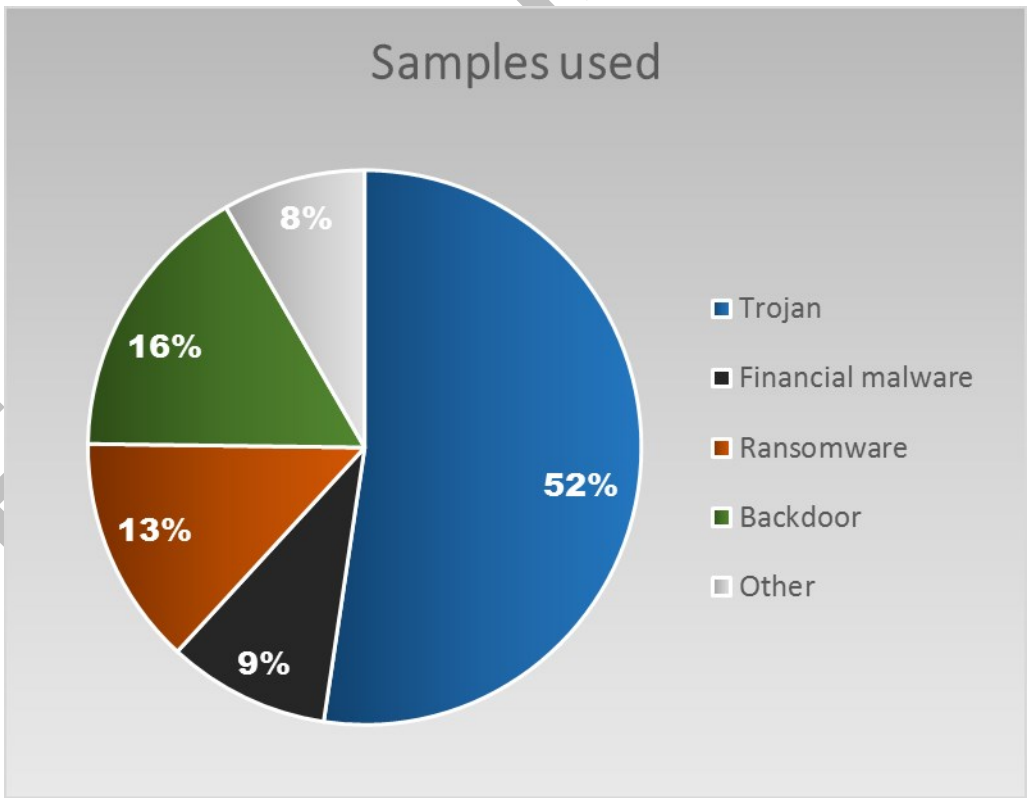
10% of the threats used in this test were introduced to the system via USB flash memory sticks. These samples came originally from live URLs, but inside archives.

Testing was conducted as per the methodology detailed in Appendix 1. In total, 359 live ITW samples were used. The stimulus load comprised the following: 187 trojans, 59 backdoors, 34 financial malware samples, 48 ransomware samples, and 31 others.

Security Applications Tested

- avast! Internet Security 12.3.2280
- AVG Internet Security 16.131.7924
- Avira Internet Security 15.0.23.58
- Bitdefender Internet Security 2017 21.0.21.976
- ESET Internet Security 10.0.369.0
- Kaspersky Internet Security 17.0.0.611 (b)
- Malwarebytes Anti-Malware 2.2.1.1043
- McAfee LiveSafe 15.1.156
- Microsoft Windows Defender 4.10.14393.0
- Panda Internet Security 17.0.1
- SurfRight HitmanPro 3.7.15 - Build 281
- Symantec Norton Security 22.7.1.32
- ThreatTrack Vipre Internet Security 9.3.6.3
- Trend Micro Maximum Security 11.0.1186
- Watchdog Anti-Malware 2.20.186.576
- Webroot SecureAnywhere 9.0.13.62
- Zemana Anti-Malware 2.50.2.133

Malware sample types used to conduct the tests.

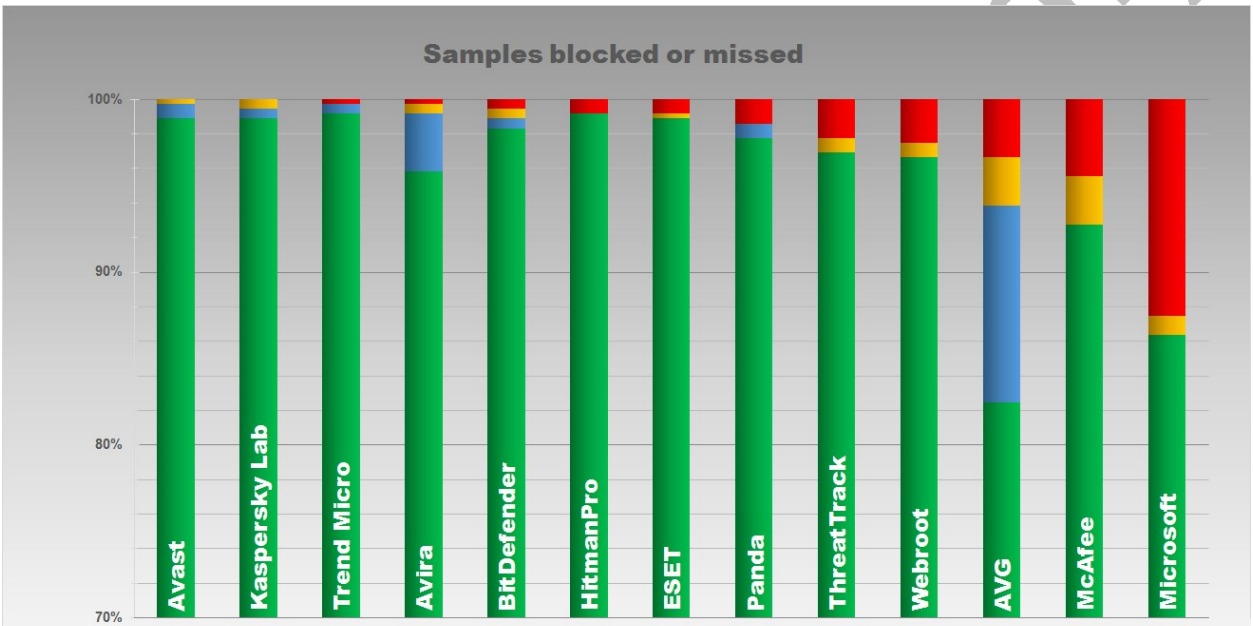


Test Results

The tables below show the results of testing under the MRG Effitas 360 Q4 Assessment Programme.

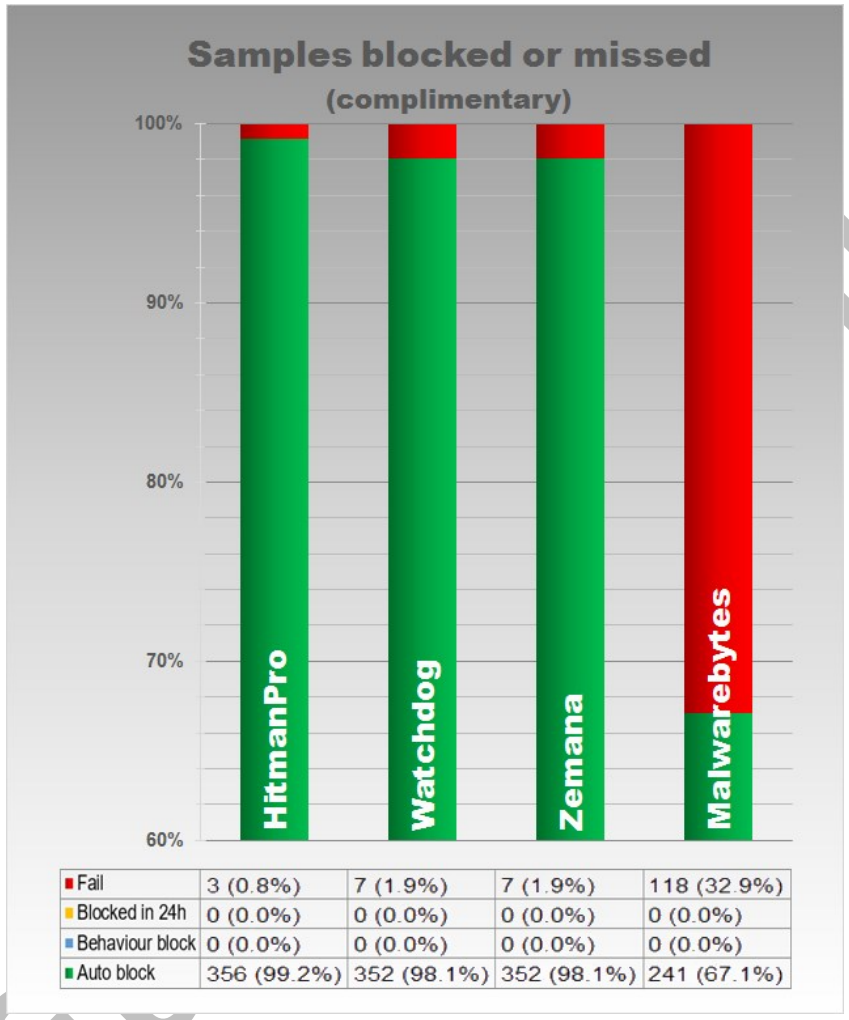
Q4 2016 In the Wild 360 / Full Spectrum Test Results

The table below shows the initial detection rates of the security products. This table is sorted by smallest amount of failures.



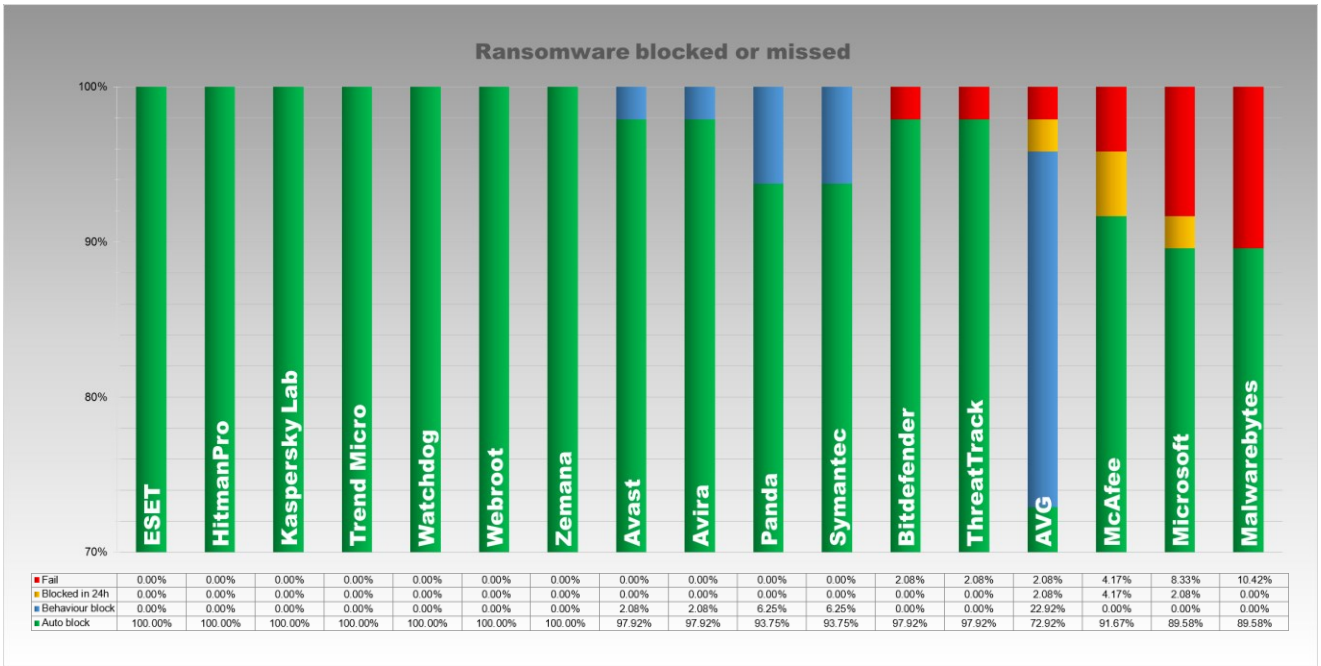
	Avast	Kaspersky Lab	Trend Micro	Avira	BitDefender	HitmanPro	ESET	Panda	ThreatTrack	Webroot	AVG	McAfee	Microsoft
Fail	0 (0.0%)	0 (0.0%)	1 (0.3%)	1 (0.3%)	2 (0.6%)	2 (0.6%)	3 (0.8%)	4 (1.1%)	8 (2.2%)	9 (2.5%)	12 (3.3%)	16 (4.5%)	45 (12.5%)
Blocked in 24h	1 (0.3%)	2 (0.6%)	0 (0.0%)	2 (0.6%)	2 (0.6%)	12 (3.3%)	1 (0.3%)	9 (2.5%)	3 (0.8%)	3 (0.8%)	10 (2.8%)	10 (2.8%)	4 (1.1%)
Behaviour block	3 (0.8%)	2 (0.6%)	2 (0.6%)	12 (3.3%)	2 (0.6%)	56 (15.6%)	0 (0.0%)	8 (2.2%)	0 (0.0%)	0 (0.0%)	41 (11.4%)	0 (0.0%)	0 (0.0%)
Auto block	355 (98.9%)	355 (98.9%)	356 (99.2%)	344 (95.8%)	353 (98.3%)	289 (80.5%)	355 (98.9%)	338 (94.2%)	348 (96.9%)	347 (96.7%)	296 (82.5%)	333 (92.8%)	310 (86.4%)

The table below shows the initial detection rates of the on-demand security products. This table is sorted by smallest amount of failures.



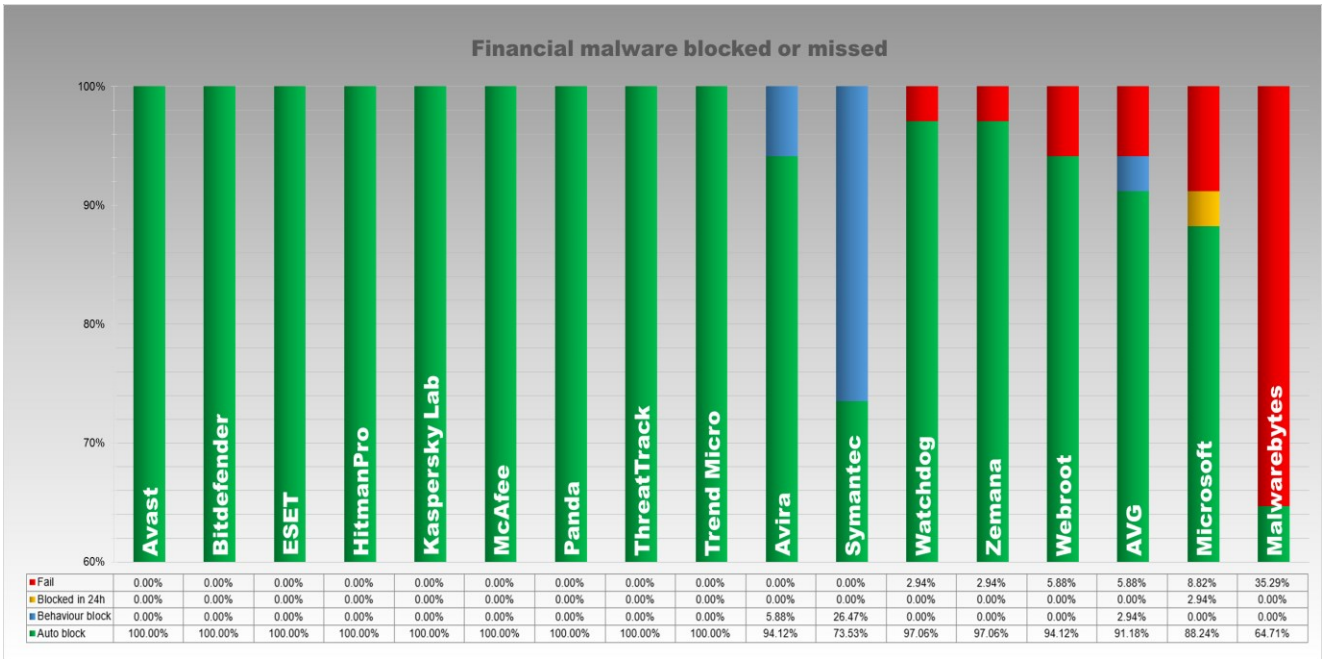
	HitmanPro	Watchdog	Zemana	Malwarebytes
Fail	3 (0.8%)	7 (1.9%)	7 (1.9%)	118 (32.9%)
Blocked in 24h	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Behaviour block	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Auto block	356 (99.2%)	352 (98.1%)	352 (98.1%)	241 (67.1%)

The table below shows the initial detection rates of the security products for ransomware. This table is sorted by smallest amount of failures.



	ESET	HitmanPro	Kaspersky Lab	Trend Micro	Watchdog	Webroot	Zemana	Avast	Avira	Panda	Symantec	Bitdefender	ThreatTrack	AVG	McAfee	Microsoft	Malwarebytes
Fail	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.08%	2.08%	2.08%	4.17%	8.33%	10.42%
Blocked in 24h	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.08%	4.17%	2.08%	0.00%
Behaviour block	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.08%	2.08%	6.25%	6.25%	0.00%	0.00%	22.92%	0.00%	0.00%	0.00%
Auto block	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	97.92%	97.92%	93.75%	93.75%	97.92%	97.92%	72.92%	91.67%	89.58%	89.58%

The table below shows the initial detection rates of the security products for financial malware. This table is sorted by smallest amount of failures.



	Avast	Bitdefender	ESET	HitmanPro	Kaspersky Lab	McAfee	Panda	ThreatTrack	Trend Micro	Avira	Symantec	Watchdog	Zemana	Webroot	AVG	Microsoft	Malwarebytes
Fail	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.94%	2.94%	5.88%	5.88%	8.82%	35.29%
Blocked in 24h	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	2.94%	0.00%
Behaviour block	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	5.88%	26.47%	0.00%	0.00%	0.00%	2.94%	0.00%	0.00%
Auto block	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	94.12%	73.53%	97.06%	97.06%	94.12%	91.18%	88.24%	64.71%

Understanding Grade of Pass

- **Level 1** = All threats detected on first exposure or via behaviour protection or detected in 24 hours.

  avast! Internet Security

  Kaspersky Internet Security

- **Level 2** = At least 97% of the threats detected and neutralised / system remediated before or on the first rescan.

  Avira Internet Security

  Bitdefender Internet Security

  ESET Smart Security

  Panda Internet Security

  SurfRight HitmanPro

  Symantec Norton Security

  ThreatTrack Vipre Internet Security

  Trend Micro Maximum

  Watchdog Anti-Malware

  Webroot SecureAnywhere

  Zemana Anti-Malware

- **Failed** = Security product failed to detect all infections and remediate the system during the test procedure.

  AVG Internet Security

  Malwarebytes Anti-Malware

  McAfee Internet Security

  Microsoft Windows Defender

Appendix 1
Methodology Used in the 360 Assessment & Certification Programme Q4 2016

Methodology used in the assessment:

1. Windows 10 64 bit operating system was installed on a virtual machine[i], all updates were applied and third party applications installed and updated according to our "Average Endpoint Specification"[ii]
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[iii] on each of the systems created in 3. and then, where applicable, updated.
5. A clone of the system as at the end of 4. was created.
6. Each live URL test was conducted by:
    a. Downloading a single malicious binary from its native URL using Microsoft Edge to the desktop, closing Microsoft Edge and then executing the binary.
    b. The security application blocked the URL where the malicious binary was located.
    c. The security application detected and blocked the malicious binary whilst it was being downloaded to the desktop.
    d. 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.
7. 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 6. and allowed it to be executed.
8. Testing on infected systems continued for 24 hours. The system was rescanned once, exactly 24 hours after the system was compromised.
9. Remediation performance of an application was determined by manual inspection of the system in contrast to its pre-infected state and not by the logs and reports of the security application itself.[iv]
10. Testing was conducted with all systems having internet access.
11. Each individual test for each security application was conducted from a unique IP address.
12. All security applications were fully-functional unregistered versions or versions registered anonymously, with no connection to MRG Effitas.
13. All testing was conducted during Q4 2016.
14. 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.

[i] VM hardware spec is 4GB RAM & 2 core processor.
[ii] AES includes Adobe Flash, Reader, Java, Microsoft Office 2010, Edge & VLC Player. All Microsoft components were fully updated; all third-party components were out of date by three months.
[iii] During installation of the security application, if an option to detect PUAs was given, it was selected.
[iv] This is because in some instances, an application will claim to have removed an infection, but actually failed to do so and was still active on the system.