SE Labs

# Endpoint Protection Testing Methodology

Methodology version 1.21; Updated: 27/11/2023

## Contents



## Change Log

Version 1.21, Updated 27/11/2023
Added scoring details.

Version 1.2, Updated 17/10/2019
Updated enterprise false positive testing process to include pre-installed applications. Removed incorrect claim that "automatic submission of data to vendors is disabled".

Version 1.1, Updated 01/12/2017
Test platform is upgraded from Windows 7 to Windows 10. Hardware is upgraded with SSD hard disks. Threat selection is now officially bound to real-world prevalence and does not require specific threat types.

## 1.0 Test framework

The test framework collects threats, verifies that they will work against unprotected targets and exposes protected targets to the verified threats to determine the effectiveness of the protection mechanisms.

### 1.1 Threat Management System (TMS)

The Threat Management System is a database of attacks including live malicious URLs; malware attached to email messages; and a range of other attacks generated in the lab using a variety of tools and techniques. Threats are fed to the Threat Verification Network (TVN).

### 1.2 Threat Verification Network (TVN)

When threats arrive at the Threat Verification Network they are sent to Vulnerable Target Systems in a realistic way. For example, a target would load the URL for an exploit-based web threat into a web browser and visit the page; while its email client would download, process and open email messages with malicious attachments, downloading and handling the attachment as if a naïve user was in control.

Replay systems are used to ensure consistency when using threats that are likely to exhibit random

behaviours and to make it simpler for other labs to replicate the attacks.

### 1.3 Target Systems (TS)

Target Systems (TS) are identical to the Vulnerable Target Systems used on the Threat Verification Network, except that they also have endpoint protection software installed.

### 1.4 Threat selection

All of the following threats are considered valid for inclusion in the test, although the distribution of the different types will vary according to the test's specific purpose:

a) Public exploit-based web threats (exploitation attacks)
b) Public direct-download web threats (social engineering attacks)
c) Public email attachment threats (exploitation and social engineering attacks)
d) Private exploit-based web threats (exploitation attacks)
e) Private direct-download web threats (social engineering attacks)
f) Private email attachment threats (exploitation and social engineering attacks)

Public threats are sourced directly from attacking systems on the internet at the time of the test and can be considered 'live' attacks that were attacking members of the public at the time of the test run. Multiple versions of the same prevalent threats may be used in a single test run, but different domain names will always be used in each case.

Private threats are generated in the lab according to threat intelligence gathered from a variety of sources and can be considered as similar to more targeted attacks that are in common use at the test of the test run.

All threats are identified, collected and analysed independently of security vendors directly or indirectly involved in the test.

The full threat sample selection will be confirmed by the Threat Verification Network as being malicious.

False positive samples will be popular and non-malicious website URLs as well as applications downloaded directly from their source websites where possible.

### 1.5 Target System details

The Target Systems are identical Windows PCs specified as below.

Each system has unrestricted internet access and is isolated from other Target Systems using Virtual Local Area Networks (VLANs).

Each system runs Windows 10 (64-bit), updated with security patches.

Popular but vulnerable third-party applications installed.

If a security product requires an updated file from Microsoft the tester will install the necessary file.

A web session replay system will be used when exposing systems to web-based threats. This provides an accurate simulation of a live internet connection and allows each product to experience exactly the same threat. All products have real-time and unrestricted access to the internet.

Products run with the default settings. Additional logging may be enabled if requested by the vendor of the product in question. Vendors of business software are invited to make configuration recommendations.

All products are updated fully using the latest definitions, patches and any other available updates. These updates are made immediately prior to each exposure to a threat or legitimate application. Products may be upgraded to the latest version, if the version changes during the test period.

### 1.6 Target System specification

Specification: Intel Core i3-4160 3.6GHz processor; 4GB RAM; 500GB 7200 RPM SATA SSD hard disk

# 2.0 Threat selection and management

### 2.1 Sample numbers and sources

The Target Systems will be exposed to a selection of threats. These are weighted heavily (~75 per cent) towards public threats as judged by SE Labs to be prevalent at the time of testing. These may be web threats, email attachments or deliverable by some other realistic vector.

A smaller set of the samples will include targeted attacks delivered by web download, exploitation or as email attachments. There may also be some threats found via alternative routes, such as internet messaging (IM) or peer-to-peer (P2P) networks.

### 2.2 Sample verification

Threats will be verified using Vulnerable Target Systems, as outlined above (see *1.0 Test framework*).

Threat verification occurs throughout the test period, with live public threats being used on shortly after they are verified as being effective against the Vulnerable Target Systems on the Threat Verification Network.

In cases where a threat is initially verified to be effective, but which is found not to be effective during testing (e.g. its C&C server becomes unavailable) the threat sample will be excluded from the test results of each product.

### 2.3 Attack stage

Threats will be introduced to the system in as realistic a method as possible. This means that threats found as email attachments will be sent to target systems in the same way – as attachments to email messages. Web-based threats are downloaded directly from their original sources. These downloads occur through a proxy system that includes a session replay service to ensure consistency.

Public threats that run on the Target System are allowed 10 minutes to exhibit autonomous malicious behaviour. This may include initiating connections to systems on the internet or making changes to the system to establish persistence.

# 3.0 Legitimate sample selection

Non-malicious website URLs and application files are used to check for false positive detection. The number of these URLs and files will match the number of malware samples used. Candidates for

legitimate sample testing include newly released applications, ranging from free software to the latest commercial releases.

When testing business grade products, the delivery method of the legitimate applications reflects real-world conditions. A system image with all business applications installed is created and the product under test is then installed on this new corporate image. If the product performs any full disk scanning during the installation process, any detections resulting from this will be noted. After the product is deployed each application will be executed for at least 60 seconds, with as many features of the application used as possible.

Potentially unwanted programs, which are not clearly malicious but that exhibit dubious privacy policies and behaviours, will be excluded from the test.

# 4.0 Measuring success

The following occurrences during the attack stage will be recorded.

### 4.1 The point of detection

(e.g. before/after execution).

### 4.2 Detection categorisation, where possible

(e.g. URL reputation, signature or heuristics).

### 4.3 Details of the threat, as reported by the product

(e.g. threat name; attack type).

### 4.4 Unsuccessful detection of threats.

### 4.5 Legitimate files allowed to run without problems.

### 4.6 Legitimate files acted on in non-optimal ways

(e.g. accusations of malicious behaviour; blocking of installation) and at what stage (e.g. before/after execution).

### 4.7 User alerts/interaction prompts such as:

a) Pop-up information messages (even if non-interactive).
b) Requests for action (take default option or follow testing policy of 'naïve user' if no default provided).
c) Default suggestions.
d) Time-out details (e.g. record if an alert/request for action disappears/takes a default action after n seconds of no user response).

### 4.8 When an initial attack or attacker succeeds in downloading further malicious files, such downloads will be recorded along with the product's behaviour (if any).

This additional data will be presented alongside the main results, clearly labelled as representing a second attack. For statistical purposes, detection rates of these files will not be automatically added to the overall totals for each product (although doing so after the event will be possible).

### 4.9 Any anomalies

(e.g. strange or inconsistent behaviour by the product.)

# 5.0 Measuring product effectiveness

Each Target System is monitored to detect a product's ability to detect, block or neutralise threats that are allowed to execute. Third-party software records each Target System's state before, during and after the threat exposure stage. These results show the extent of an attacker's interaction with the target and the level of remediation provided by the product being tested.

The same level of monitoring occurs when introducing legitimate URLs and files when assessing false positive detection rates.

Products are scored according to how they handle both public and targeted threats. The scoring details for targeted threats are more nuanced than for public threats because the testers have direct control of the targeted attacks and can turn a basic attack into a more detailed breach.

The following tables show how we score different outcomes.

### 5.1 Public threat scoring

| Result | Description | Score |
|---|---|---|
| Detected | If the product detects the threat with any degree of useful information, we award it one point. | +1 |
| Blocked | Threats that are disallowed from even starting their malicious activities are blocked. Blocking products score two points. | +2 |
| Complete Remediation | If, in addition to neutralising a threat, the product removes all significant traces of the attack, it gains an additional one point. | +1 |
| Neutralised | Products that kill all running processes associated with the test case ‘neutralise’ the threat and win one point. | +1 |
| Persistent Neutralisation | This result occurs when a product continually blocks a persistent threat from achieving its aim, while not removing it from the system. | -2 |
| Compromised | If the threat is successful in achieving its malicious activities, the product loses five points. This loss may be reduced to four points if it manages to detect the threat (see Detected, above), as this at least alerts the user, who may now take steps to secure the system. | -5 |

## 5.2 Targeted threat scoring

| Result | Description | Score |
|---|---|---|
| Access | If any command that yields information about the target system is successful, this score is applied. Examples of successful commands include listing current running processes, exploring the file system and so on. If the first command is attempted and the session is terminated by the product without the command being successful, the score of Neutralised (see above) will be applied.<br>This stage serves as proof that the connection between the target and the attacker is functional. | -1 |
| Action | If an activity causes significant damage to the targeted user or system, the product loses one point. Examples include data exfiltration, introduction of new files and modification of existing files. | -1 |
| Escalation | The attacker attempts to escalate privileges to above the level of a standard user. If successful, an additional two points are deducted. | -2 |
| Post-Escalation Action | After escalation, the attacker attempts actions that rely on escalated privileges. These actions can include attempting to steal credentials, achieving persistence and deploying ransomware. If any of these actions are successful, then a further penalty of one point is deducted. | -1 |