SE Labs

Predictive malware response testing methodology

Methodology version 1.0; Created 17/01/2018

Contents

1.0 Introduction

This methodology provides a way to test the ability of anti-malware file scanners to detect and protect against unknown threats. This approach to retrospective testing uses a combination of known threats, detected in the past, and security products deployed in states and/ or versions as were available in the time prior to the first known detections of the threats used.

Crucially products are tested in an offline state, disallowing modern updates and queries to internet-hosted services.

The threats used represent well-known, highly-impactful attacks recognised by businesses and the wider media as being of interest to organisations facing opponents including organised criminals and both competitive and nation-state-related attackers.

2.0 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.

2.1 Threat Management System (TMS)

The Threat Management System is a database of attacks including a range of attacks that meet the criteria of *2.4 Threat selection.* Threats are fed to the Threat Verification Network (TVN).

2.2 Threat Verification Network (TVN)

When threats arrive at the Threat Verification Network they are sent to Vulnerable Target Systems and attempts are made to execute them. Those that fail to execute are excluded from the test.

2.3 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.

2.4 Threat selection

Threats represent well-known, highly-impactful attacks recognised by businesses and the wider media as being of interest to organisations facing opponents including organised criminals and both competitive and nation-state-related attackers.

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 direct-download web threats (social engineering attacks)
b) Public email attachment threats (exploitation and social engineering attacks)
c) Private direct-download web threats (social engineering attacks)
d) Private email attachment threats (exploitation and social engineering attacks)

Public threats are sourced from a variety of sources, including respected public malware repositories, data leak websites, cooperating organisations and from collections independently created and curated by SE Labs.

Private threats are variations of the public threats and are generated in the lab according to threat intelligence gathered from a variety of sources.

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 legitimate, popular applications, excluding those likely to be classified as "potentially unwanted", "risky" or otherwise undesirable, such as those listed in the AppEsteem Deceptor list."

2.5 Target System details

The Target Systems are identical Windows systems specified as below.

Each system is disconnected from the internet and is isolated from other Target Systems using Virtual Local Area Networks (VLANs).

Each system runs Windows 7 (64-bit), updated with security patches available up to Service Pack 1.

Popular but vulnerable third-party applications installed include Adobe Flash Player, Adobe Reader, Apple QuickTime and Oracle Java (32-bit).

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

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 deployed with the appropriate state and/ or version according to the time period being simulated in the test.

2.6 Target System specification

Specification: Virtualised; 4GB RAM

3.0 Threat selection and management

3.1 Sample numbers and sources

The Target Systems will be exposed to nine sets of threats. Each set compromises a real malware executable as discovered following a well-known, highly-impactful attack as described in *2.4 Threat selection*. Each such attack will be further changed in various realistic ways to create four further functionally-similar attacks with changes designed to evade detection.

3.2 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.

In cases where a threat is initially verified to be effective, but which is found not to be effective during testing (e.g. due to an environmental issue such as lack of internet connection) the threat sample will be excluded from the test results of each product.

3.3 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. As there is no internet access allowed to the tested systems, threats that originally arrived from internet-hosted websites will be copied to the targets using alternative ways.

NOTE: This test methodology applies to anti-malware file scanners. It is not suitable for testing products with multiple layers of security that might rely on internet-hosted services to provide live threat protection information.

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

4.0 Legitimate sample selection

Non-malicious and application files are used to check for false positive detection. The number of these 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.

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

5.0 Measuring success

The following occurrences during the attack stage will be recorded.

5.1 The point of detection

(e.g. before/after execution).

5.2 Detection categorisation, where possible

(e.g. Signature or heuristics).

5.3 Details of the threat, as reported by the product

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

5.4 Unsuccessful detection of threats.

5.5 Legitimate files allowed to run without problems.

5.6 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).

5.7 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).

5.8 Secondary payloads

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).

5.9 Any anomalies

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

5.10 Earliest predicted date

Products are deployed in various states and/ or versions, ranging from older versions through to more recent versions. The latest version of each product is exposed to a threat and the results measured. If a detection and/ or protection is determined the threat is then re-used with increasingly older versions of the product until either all versions have been tested and found to be successful or a version is unable to detect and/ or protect against the threat. The 'earliest predicted date' corresponds to the date stamp on the oldest successful product version.

6.0 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 files when assessing false positive detection rates.