Public

Breach Response Testing Methodology

Contents

1. Test framework

The test framework collects threats, verifies that they work against unprotected targets and exposes protected targeted 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 adversary techniques and tools used to emulate real world threat actors. Test cases are applied to the Threat Verification Network (TVN).

1.2 Threat Verification Network (TVN)

Threats sourced from the TMS are sent to vulnerable target system to ensure the validity of each test case. The TVN is set up to emulate different environments. This includes, but is not limited to, Windows endpoints, Windows servers and Linux-based devices.

1.3 Initial Attack Vectors and Targeted Systems

The following attack vectors are considered valid.

a) Private e-mail attachments (social engineering attacks)
b) Private direct-download web threats (social engineering attacks)
c) Private exploit-based web threats (exploitation attacks)
d) Access with compromised credentials (using credentials stolen via spear phishing attacks, enabling initial access to target devices)
e) Previously compromised endpoints (replicating an attacker with foothold on a network that was established before the tested security product was deployed)

The targeted systems will be configured for the listed attack vectors.

1.4 Scenario selection

Scenarios used in this test are subject to change to reflect real world threat events. This helps identify products with coverage against multiple threat actors throughout a period of time. Scenarios are specified in each test plan and any subsequent reports.

Scenarios are based on publicly available information. SE Labs maps key points of attacks to MITRE's ATT&CK Matrix for Enterprise to help readers of our reports identify strengths and weaknesses in the tested products.

1.5 Product configuration

The configuration of the products is made available to readers of any testing reports. Decisions to configure products differently to recommended policies are explained.

2. Legitimate sample selection

Non-malicious website URLS and applications files are used to check for false positive detection. Candidates for legitimate sample testing include newly released applications and internally developed applications/ scripts that a system administrator may use in their environment.

Potentially unwanted programs, which are not clearly malicious but exhibit dubious privacy policies and behaviours, are excluded from the test. The candidates are split evenly into pre-installed and newly introduced applications.

Security products are deployed to a system with the legitimate candidates already installed. The products are expected to allow the user normal interaction with each test case. The products are expected to allow new applications to be installed with no or minimal friction.

3. Measuring success

This test methodology allows for a variety of approaches to security provision. Solutions with preventive and remediating capabilities are rated as described in section 3.1.1. Pure Endpoint Detection and Response (EDR) products are rated as described in section 3.12.

Detection of an attack technique does not equivalate to protection against that technique. The rating paradigm for each product enrolled in the test is decided during the configuration phase.

3.1 Evaluated techniques

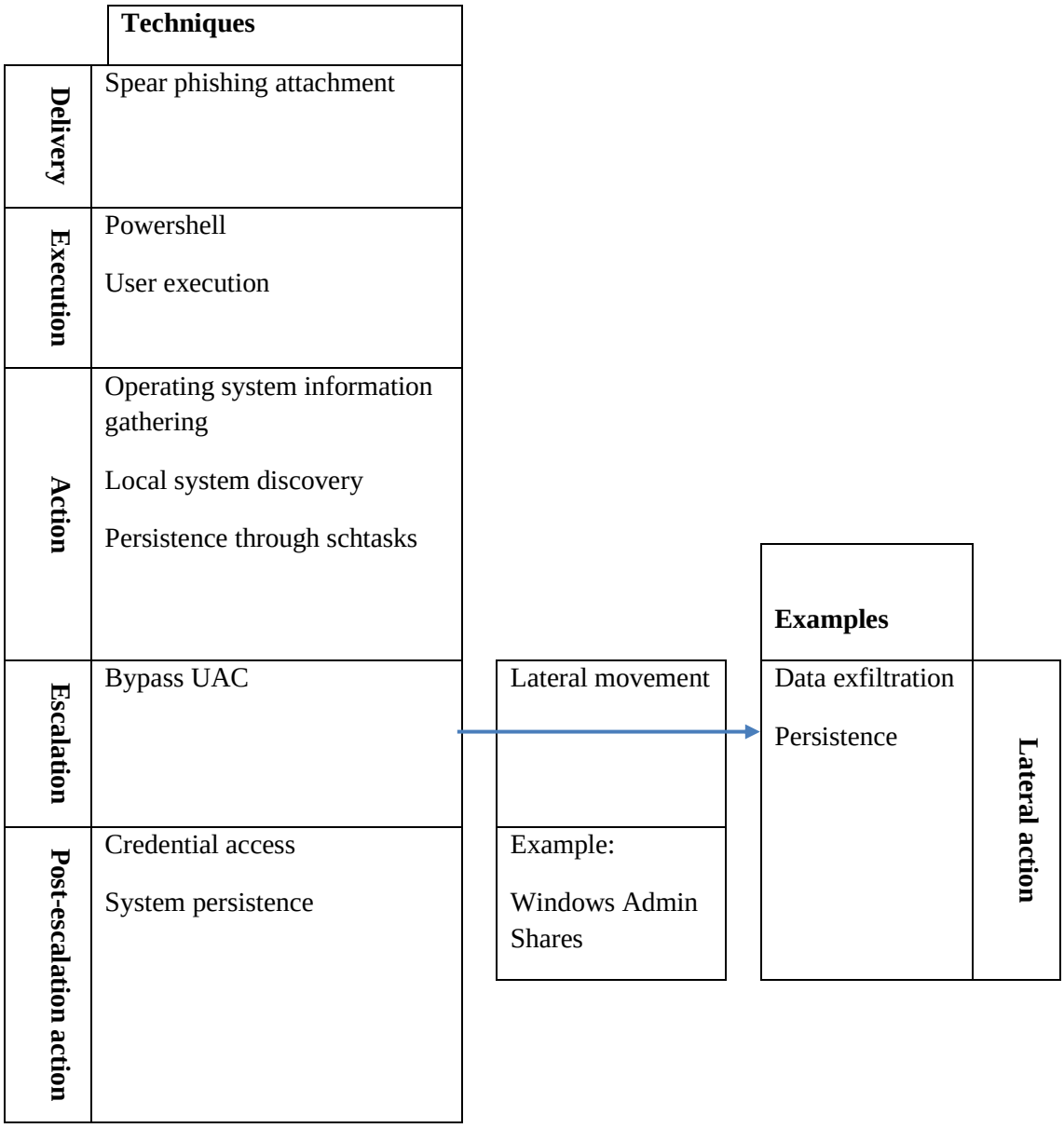
For each attack scenario under test a MITRE ATT&CK-style matrix showing the techniques used in the test is presented in the report.

3.1.1 Solutions with preventative and remediating capabilities

Each test case is measured against a typical threat chain. Depending on the scenario, the product may be subjected to different techniques at each stage of the attack.

The following illustration shows some examples:

3.1.1.1 Threat Chain Examples



Each test case starts with a maximum available four points. For each successful stage of the attack penalties will be applied (see Scoring Penalties below). The penalty for each attack stage will only be applied when the attacker is able to execute harmful actions on the target.

A fully successful attack will yield a score of -5. Stages that are most impactful on the compromised organisation incur heavier penalties than those that affect only the initially compromised endpoint.

3.1.1.2 Scoring Penalties

Attack Stage	Delivery	Execution	Action	Escalation	Lateral Movement	Lateral action	Post Escalation Action
Rating	-0.5	-0.5	-1	-2	-2	-2	-1
Exceptions	If the initiating file is removed after an insignificant time frame (< 60 sec) and it has not caused any damage to the system, the penalty is removed.		N/A				

3.1.1.3 Attack stage definitions

Delivery

Successful download of a malicious file.

Execution

Manual or exploit-driven launch of the downloaded malicious file.

NOTE: In case of tested drive-by exploit techniques, if the attack is successful in creating a connection to its command and control server this stage is considered successful.

Action

Successful execution of unauthorised and potentially malicious commands on the initial target system.

Examples include:

- Discovering current local processes
- Discovering current network connections
- Recording keystrokes
- Discovering system information
- Persistence via schtasks
- Downloading files from the Target

Escalation

Gaining elevated privileges above the standard user level. Examples of techniques possible at this stage:

- Access Token Manipulation
- Bypass User Control

Post-Escalation Action

Actions that require escalated privileges. Examples include:

- Credential Access
- System persistence
- Token impersonation

Lateral Movement

Gaining access to a secondary endpoint after the initial compromised target. The technique used for this will vary depending on the attack emulated.

Lateral Action

Successful actions taken on the secondary targeted machine.

Associating the rating with MITRE ATT&CK framework

To allow for better understanding of the test results in accordance with the MITRE ATT&CK framework, reports include tables showing the progress of each attack in a similar way to the example below:

Security Product X	**Initial Access**	**Execution**	**Privilege Escalation**	**Exfiltration**
Test case 1	Technique a	Technique c	Technique x	Technique z
Test case 2	Technique a	Technique c	Technique x	Technique t
Test case 3	Technique b	Technique y	Technique x	Technique u

Key

Unprotected	The product offered no notification or protection capabilities against this technique.	Detected but not protected	The product detected the technique but did not protect against it.
Prevented	The product blocked the prerequisite for the technique, which was not tested.	Detected and protected	The product detected and protected against the technique.

3.1.2 Solutions with Endpoint Detection and Response or similar characteristics

Each stage of the attack is broken down into a MITRE ATT&CK-style matrix containing information about the coverage of the different techniques. Tested products are expected to show clear notification of each malicious behaviour. While direct references to the listed techniques in the MITRE ATTT&CK matrix are valuable, they are not requirement for detection credit to be given.

(If a drive-by compromise technique is tested, the notification from the tested solution does not need to refer to 'T1189', MITRE's specific reference code for that technique. Descriptive language in the notification is enough to recognise this technique.)

The results for the products tested under this paradigm are presented in the following way:

Execution	Persistence	Privilege Escalation	Defence evasion	Credential access
Technique 1	Technique 1	Technique 1	Technique 1	Technique 1
Technique 2	Technique 2	Technique 2	Technique 2	Technique 2
Technique 3	Technique 3	Technique 3	Technique 3	Technique 3

Key

Out of scope	Not detected	Detected

3.2 Detection categorisation

How the product itself categorises its detection of the threat. For example, using a 'signature' or 'heuristics'.

3.3 Reported threat details

How the product reports the threat when detected. For example, the threat's name or an attack type.

3.4 Unsuccessful threat detection

When the product fails to detect the threat, this is recorded.

3.5 Successful execution of legitimate applications

Products are expected to allow legitimate applications to execute and operate without problems. A rating modifier is applied to each application according to its general prevalence. Internally developed applications and scripts are considered very low impact because they are not very prevalent globally. The following table shows the rating modifiers:

Impact category	Very high	High	Medium	Low	Very low
Rating modifier	5	4	3	2	1

3.6 Sub-optimal handling of legitimate applications

How the product categorises and allows or hinders the application. For example, it might generate a false positive result by classifying the application as being malware, or it might block installation with or without warning (see 3.7 User interaction, below). It might misclassify and/ or block activity before or after installation.

3.7 User interaction

The security product may interact with the user when a malicious or legitimate application is analysed. Details of these interactions, such as those below, are recorded:

a) Pop-up information messages (even if not requiring a response).
b) Requests for action (the tester will take the default option or follow a testing policy of 'naïve user' if no default is provided).
c) Default suggestions made by the product.
d) Time-out details (e.g. if an alert/ request for action disappears/ takes a default action after *n* seconds of no user response).

3.8 Administrator interaction

If applicable to a product, its administrator dashboard can contain vital information about the malicious actions taken on a protected endpoint. Cloud or on-premise control panels are acceptable. Details are recorded alongside any client-side notifications.

3.9 Anomalies

Testers record any strange or inconsistent behaviour shown by the product.

4. Change log

13/01/2020	v1.0	Document created
10/02/2020	v1.01	Minor typographical corrections

SE LABS LTD
4 Cromwell Court, New Street, Aylesbury, Buckinghamshire, HP20 2PB, United Kingdom.
Registered in England: 9688006.
Tel: +44(0)20 3875 5000; Email: aux@selabs.uk