AP CYBERSECURITY • EXERCISE SHEET

4.4 Detecting Attacks on Devices

Name: _____ Class: _____ Date: _____ **Total: 43 marks**

KEY WORDS authentication logs 认证日志 • spraying 密码喷洒 • credential stuffing 撞库 • indicators of compromise (iocs) 入侵指标 • hash 散列值

Objective

Build the skills to answer exam questions on **indicators of compromise (IoCs)** 入侵指标, choosing and evaluating device detection controls, and reading **authentication logs** 认证日志 for password attacks.

You must be able to:

- define an IoC and classify one as host-based, file-based or behaviour-based
- choose a device detection method against performance, cost and criticality
- evaluate a method on speed, the attack phase it catches, and false positives versus ease of bypass
- tell password guessing, **spraying** 密码喷洒 and **credential stuffing** 撞库 apart from a log
- explain why an offline password attack cannot be detected

1 Worked examples

Study these first. Each one shows the method for a task used later.

■ What an IoC is, and which kind

Computing systems log system processes and settings, login attempts, file download attempts and user actions, and those logs let an investigator **reconstruct** what happened. An **indicator of compromise** is evidence that an adversary **has** compromised a device or network - past tense, not a prediction.

Three kinds, separated by **where you found it**:

Kind	Found by	Examples
Host-based	analysing logs and **configuration settings**	unusual files created or modified; unexpected processes or services; unauthorised configuration changes; unauthorised software installation
File-based	analysing the **files** on the device	a file whose **hash** matches known malware; a file name a known malware creates; a file path associated with malicious activity
Behaviour-based	analysing authentication and **access** logs	multiple failed logins; unusual login times or locations; unauthorised attempts to reach sensitive data; attempts to **elevate privileges**

The overlap trap: "multiple failed logins" is behaviour-based even though it comes from a log, while "an unexpected new service" is host-based even though it also comes from a log. Ask what the evidence is **about** - a *behaviour*, a *file*, or the *host's state*.

■ Choosing a device detection method

- **Performance** - detection tools consume memory and processing power; anomaly-based costs more of both, which matters on a device that must stay responsive.
- **Cost** - enough licences for every device adds up quickly across a fleet.
- **Criticality** - a device holding sensitive data or providing a critical service justifies the more thorough, more expensive method.

■ Evaluating the impact of a method

- **Speed and performance** - signature-based is generally faster, and that gap widens on a busy device.
- **Phase of the attack** - to act on a device, an adversary must first get past a combination of physical, network and device controls, so detection at an earlier phase limits the damage.
- **False positives versus ease of bypass** - most device-level tools are signature-based, which gives very few false positives but is easier for an adversary to bypass with something new.

■ Reading an authentication log

Auth logs record **every attempted login**. Three attacks look similar at a glance and are told apart by **how many accounts** and **how many passwords**:

Pattern in the log	Attack
One user, **many** wrong passwords	online password **guessing**
Many users, each tried within seconds, from **one** IP	password **spraying**
A series of **default user:password** pairs in quick succession	credential **stuffing**
An authorised user succeeding from an unexpected **location, IP or time**	the credentials are probably already compromised

Example. Forty different usernames, one attempt each, all from 203.0.113.7 inside ten seconds. Not guessing - only one password per account. Not stuffing - the passwords are not the vendor defaults. This is **spraying**: one common password tried against many accounts to stay under a lockout threshold.

Offline attacks cannot be detected at all. Once the adversary has copied the password file, the cracking happens on **their** computer, producing no entry in any log you own. That is why the detectable moment is the theft of the file - a host-based IoC - and not the attack itself.

2 Practice

2.1 Define an **indicator of compromise**. [2]

2.2 Classify each IoC as host-based, file-based or behaviour-based.

(a) A file whose hash matches known malware. [1]

(b) An attempt to elevate a standard user to administrator. [1]

(c) An unexpected service running at startup. [1]

(d) A login from an unusual country at 03:00. [1]

(e) An unauthorised change to a system configuration setting. [1]

2.3 State **two** things a computing system logs that help reconstruct an attack. [2]

2.4 Name the password attack from each log pattern.

(a) One username, sixty wrong passwords in two minutes. [1]

(b) Ninety usernames, one attempt each, all from one IP within seconds. [1]

(c) admin:admin, root:root, user:password tried in quick succession. [1]

2.5 Explain why an **offline** password attack cannot be detected. [2]

2.6 An authorised user's account logs in successfully from a country they have never visited.

(a) State what this indicates. [1]

(b) Explain why it is a behaviour-based rather than a host-based IoC. [2]

2.7 A school must choose a detection method for a fleet of 300 ageing student laptops.

(a) State **two** criteria that should drive the decision here. [2]

(b) Recommend a method and justify it. [2]

2.8 Explain the trade-off between false positives and ease of bypass for signature-based device detection. [2]

3 Exam-style questions

3.1 An IoC is best defined as [1]

- **A** a prediction that a device will be attacked
- **B** evidence that an adversary has compromised a device or network
- **C** a vulnerability found during a scan
- **D** a rule that blocks malicious traffic

3.2 Many users each attempting one login within seconds from a single IP address indicates [1]

- **A** credential stuffing
- **B** password spraying
- **C** an offline attack
- **D** online guessing of one account

3.3 Which is a **file-based** IoC? [1]

- **A** multiple failed login attempts
- **B** an unauthorised software installation
- **C** a file path known to be associated with malicious activity
- **D** a login at an unusual time

3.4 A device holding highly sensitive data is given anomaly-based detection despite the performance cost. This decision is justified mainly by [1]

- **A** cost
- **B** speed of detection
- **C** criticality of the device
- **D** the false positive rate

3.5 An analyst reviews one week of logs from a company file server. The authentication log shows, on Monday, 74 different usernames each tried once from 198.51.100.22 within eleven seconds; nothing succeeds. On Wednesday, one account belonging to a member of the finance team logs in successfully at 02:40 from an IP address in another country, and shortly afterwards the log records an attempt to elevate that account's privileges. On Thursday the server has a new scheduled task that no administrator created, and a file in a temporary directory whose hash matches a known remote access trojan.

(a) Name the attack on Monday and state the indicator that identifies it. [2]

(b) Classify **each** of the Wednesday and Thursday findings as host-, file- or behaviour-based, giving a reason for one of them. [4]

(c) Reconstruct the likely sequence of the attack across the week, naming the phase each stage belongs to. [4]

(d) Monday's attempt failed but Wednesday's succeeded. Explain what this suggests about how the finance account's password was obtained, and why that route left nothing in the logs. [3]

(e) The company runs signature-based detection on all servers. Explain **one** reason this was adequate on Thursday and **one** reason it was not adequate earlier in the week. [3]

Solutions

2.1 evidence that an adversary **has** compromised a device or network; found by analysing logs, configuration settings or files, and used to reconstruct what happened.

2.2 (a) file-based. (b) behaviour-based. (c) host-based. (d) behaviour-based. (e) host-based.

2.3 any two of: system processes and settings, login attempts, file download attempts, user actions.

2.4 (a) online password guessing. (b) password spraying. (c) credential stuffing.

2.5 the attack takes place on the **adversary's own computer**, using a copy of the password data they already hold; nothing is attempted against your system, so no authentication log entry is generated and there is nothing to detect.

2.6 (a) the user's credentials have probably been compromised. (b) the evidence is about the **behaviour** of an account - where and when it logged in - found in the authentication log, rather than about the state of the host, such as a new file, process or configuration change.

2.7 (a) any two of: **performance** - detection tools consume memory and processing power, and these laptops are ageing; **cost** - 300 licences is a large recurring expense; **criticality** - student laptops hold little sensitive data and provide no critical service. (b) **signature-based** - it is faster and uses fewer resources, so it will not slow ageing machines, and it costs less across 300 devices; the low criticality means the higher false-negative rate is an acceptable trade.

2.8 signature-based detection matches known patterns, so it produces almost no false positives and analysts trust its alerts; but an adversary only needs something the signature set has not seen to pass straight through, so it is easier to bypass than anomaly-based detection.

3.1 B. An IoC is evidence of a compromise that has already happened.

3.2 B. One attempt per account rules out guessing; non-default passwords rule out stuffing.

3.3 C. A is behaviour-based, B and D are host-based and behaviour-based respectively.

3.4 C. Criticality is the criterion that justifies accepting cost and performance impact.

3.5 (a) **password spraying** - many different users each attempted once, within seconds, from a single IP address.

(b) Wednesday's out-of-hours foreign login: **behaviour-based**. The privilege-elevation attempt: **behaviour-based**. Thursday's uncreated scheduled task: **host-based**. The trojan file matched by hash: **file-based**. Reason, for any one: e.g. the scheduled task is host-based because the evidence is an unauthorised change to the device's own configuration, not an account's behaviour.

(c) Monday's spraying is **reconnaissance/initial access** attempted and failed; Wednesday's successful foreign login is **initial access** achieved with valid credentials; the privilege-elevation attempt is **lateral movement**; Thursday's scheduled task and RAT file are

persistence, keeping access without needing to log in again.

(d) the spraying failed, so the password was not guessed on this system; it was most likely obtained elsewhere - phished from the user, captured by a keylogger, or cracked **offline** from a stolen password file, or reused from another site's breach. None of those touch this server's authentication process, so no failed-attempt entries appear and the first sign here is a **successful** login.

(e) Adequate on Thursday: the trojan was a known one, so its file **hash matched a signature** and signature-based detection catches it quickly and with almost no false positives. Not adequate earlier: a successful login with a valid password is not a signature of anything - it is only abnormal relative to the user's normal location and hours, which needs anomaly- or behaviour-based analysis; and detecting only at Thursday's persistence phase means the adversary already had two days inside, whereas detection at Wednesday's access phase would have limited the damage.