AP CYBERSECURITY • EXERCISE SHEET

3.5 Detecting Network Attacks

Name: ____________ Class: ________ Date: __________ **Total: 45 marks**

KEY WORDS anomaly-based 基于异常 • signature-based 基于特征 • nids 网络入侵检测系统 • nips 网络入侵防御系统 • threshold 阈值

Objective

Build the skills to answer exam questions on automated network detection: **NIDS** 网络入侵检测系统, **NIPS** 网络入侵防御系统 and **SIEM** 安全信息与事件管理, the choice between **signature-based** 基于特征 and **anomaly-based** 基于异常 detection, the role of **AI** 人工智能, and reading logs for network attacks.

You must be able to:

- distinguish a NIDS (alerts) from a NIPS (blocks) and a SIEM (correlates many sources)
- explain why AI is used on detection data, and what a probability threshold trades away
- choose a detection method from the four criteria: traffic volume, consistency, criticality, likelihood of novel attacks
- evaluate a method on speed, cost, false positives and false negatives
- name the indicator that identifies each network attack in a log

1 Worked examples

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

The three tools

All three analyse data collected from switches, routers, servers, firewalls and user computers. What separates them is what they do next.

Tool	On detecting malicious activity
NIDS	raises an **alert**; the traffic still passes
NIPS	**blocks** the traffic as well as alerting
SIEM	collects and correlates data from **many sources** - firewalls, NIDS/NIPS, device logs, application logs

A NIPS sounds strictly better, and is not: a false positive on a NIDS is a wasted alert, while a false positive on a NIPS **blocks legitimate traffic**. That trade is the usual “justify your choice” mark.

Why AI, and what the threshold costs

A medium-sized network produces far more log data than people can read. AI models classify data patterns as malicious or normal, and because they are **probabilistic** they report a **percentage likelihood**, not a verdict.

The organisation sets the alert **threshold**, and both directions hurt:

- **too high** - real attacks go undetected (false negatives)
- **too low** - analysts drown in false positives and start ignoring alerts

Choosing a method: four criteria

Criterion	Points to
High traffic volume	signature-based - it is more efficient
Consistent traffic patterns	anomaly-based - it needs a stable baseline to call something abnormal
Sensitive or critical network	the more thorough option, accepting cost and false positives
Novel attacks likely	anomaly-based - signature-based **cannot detect a new attack**

That last line is the one that decides most scenario questions: a signature only exists for an attack somebody has already seen.

Evaluating a method: four factors

Factor	Signature-based	Anomaly-based
Speed	faster, so response is faster	slower
Cost	lower	higher - more resources and tuning
False positives	almost none	many, especially before tuning
False negatives	high for novel attacks	lower for novel attacks

A **false negative** is the adversary getting past the system - which is why a network protecting critical data will accept false positives to reduce them.

Naming the attack from the log

Attack	The indicator
Evil twin	an SSID suspiciously similar to a legitimate local one; signal triangulation locates it
Jamming	no wireless device in one physical space can connect; electromagnetic interference on scanning
ARP poisoning	unusual ARP messages, particularly **duplicate MAC address** ARP packets
MAC flooding	a sudden surge of Ethernet frames with many **different** MAC addresses
DNS poisoning	hard to detect - an abrupt, otherwise inexplicable drop in traffic to the site
Smurf	a sudden increase in **ICMP** requests to the network's **broadcast address**

Network-based IoCs come from analysing traffic, often packet captures: unexpected source and destination IP addresses, ports and protocols.

2 Practice

2.1 State the difference between a NIDS and a NIPS. [2]

2.2 Explain what a SIEM adds that a single NIDS cannot provide. [2]

2.3 Explain why AI is used to analyse network detection data. [2]

2.4 An organisation raises its AI alert threshold from 60% to 95%.

(a) State the effect on false positives. [1]

(b) State the risk this creates. [1]

2.5 Choose signature-based or anomaly-based detection for each network, and justify.

(a) A very high-volume public web front end. [2]

(b) A defence contractor's design network, where new attack techniques are expected. [2]

(c) A factory control network whose traffic is identical every day. [2]

2.6 Explain why signature-based detection cannot detect a novel attack. [2]

2.7 State whether each is a false positive or a false negative, and which method it is characteristic of.

(a) An adversary's new technique passes the system unremarked. [2]

(b) A backup job at 3 a.m. triggers an alert every night. [2]

2.8 Name the attack indicated by each log observation.

(a) A surge of Ethernet frames with many different MAC addresses. [1]

(b) A wireless SSID almost identical to the company's own. [1]

(c) Duplicate MAC address ARP packets. [1]

(d) A spike of ICMP requests to the broadcast address. [1]

3 Exam-style questions

3.1 Which tool blocks malicious traffic rather than only alerting on it? [1]

- **A** NIDS
- **B** NIPS
- **C** SIEM
- **D** packet capture

3.2 Anomaly-based detection is most effective on a network with [1]

- **A** very high traffic volume
- **B** consistent traffic patterns
- **C** no sensitive data
- **D** a limited budget

3.3 A false negative occurs when [1]

- **A** legitimate traffic is flagged as malicious
- **B** an adversary bypasses the detection system
- **C** the AI threshold is set too low
- **D** the SIEM correlates two unrelated logs

3.4 An organisation's website suffers an abrupt and otherwise inexplicable drop in traffic. Which attack should be checked first? [1]

- **A** smurf
- **B** jamming
- **C** DNS poisoning
- **D** MAC flooding

3.5 A hospital runs two networks. Network A is the public guest Wi-Fi: very high traffic, constantly changing patterns, no patient data. Network B connects imaging machines and the patient record system: low, highly consistent traffic, and the most sensitive data the hospital holds. The security team has one budget and must justify a different approach on each. Recent logs from Network B show a surge of Ethernet frames carrying many different MAC addresses, and shortly afterwards several duplicate MAC address ARP packets.

(a) Name the **two** attacks the Network B log indicates, and state the indicator for each. [4]

(b) Explain what an adversary gains by running these two attacks in that order. [3]

(c) Recommend a detection method for Network A, justifying it against **two** of the four criteria. [3]

(d) Recommend a detection method for Network B, justifying it against **two** criteria, and explain why the false-positive cost is acceptable here when it might not be on Network A. [4]

(e) The team deploys AI classification across both networks with a single shared threshold. Explain why one threshold for both networks is a poor design. [3]

Solutions

2.1 both analyse network data to determine whether malicious activity is taking place; a NIDS alerts and the traffic still passes, while a NIPS also blocks it.

2.2 a SIEM collects and analyses data from **multiple sources** - firewalls, NIDS/NIPS, device logs, application logs; so it can correlate events that look harmless in one log but form an attack when seen together, which a single sensor cannot do.

2.3 computers log every user action and every sensor logs all traffic, so a medium-sized network produces far more data than analysts can read; AI algorithms classify these data patterns as malicious or normal at that scale.

2.4 (a) false positives fall. (b) real attacks may go undetected, because anything the model scores below 95% raises no alert.

2.5 (a) signature-based - it is more efficient at high traffic volume, and the patterns here are not consistent enough for an anomaly baseline. (b) anomaly-based - signature-based cannot detect a new attack, and novel techniques are expected. (c) anomaly-based - consistent traffic patterns give a reliable baseline, so genuine deviation stands out.

2.6 detection works by matching traffic to a stored signature of a known attack; a novel attack has no signature yet, so there is nothing to match and it passes unremarked.

2.7 (a) false negative; characteristic of signature-based detection. (b) false positive; characteristic of anomaly-based detection, which flags anything outside the baseline.

2.8 (a) MAC flooding. (b) evil twin. (c) ARP poisoning. (d) smurf attack.

3.1 B. A NIDS only alerts; a SIEM correlates; a packet capture only records.

3.2 B. It needs a stable baseline before a deviation means anything.

3.3 B. A is a false positive; the threshold in C would raise false positives, not negatives.

3.4 C. DNS poisoning is difficult to detect, and this drop in traffic is its documented indicator - visitors are being sent elsewhere.

3.5 (a) **MAC flooding** - a sudden surge of Ethernet frames with many different MAC addresses; **ARP poisoning** - duplicate MAC address ARP packets.

(b) MAC flooding overwhelms the switch's MAC address table, so the switch can no longer direct frames to one port and traffic reaches the adversary that would not otherwise; ARP poisoning then associates the adversary's MAC with a legitimate IP, letting them sit between the imaging machines and the record system to read or modify patient data in transit.

(c) **Signature-based**. Traffic volume is very high, where signature-based is more efficient; and the patterns constantly change, so an anomaly baseline would be unstable and produce endless false positives on a guest network the hospital does not control.

(d) **Anomaly-based**, or a hybrid. Traffic is highly consistent, which is exactly the condition anomaly-based needs; and the network carries the most sensitive data the hospital holds, so the criticality criterion favours the more thorough method despite its cost. A false positive here means investigating a legitimate imaging transfer - an hour of analyst time - whereas a false negative means patient records altered; on the guest network the same false-positive rate would generate constant alerts for no protective gain.

(e) The threshold is a trade between false positives and false negatives, and the two networks sit at opposite ends of it. Network B needs a **low** threshold, because a missed attack compromises regulated patient data; Network A needs a **higher** one, because its volume and variability would otherwise bury the team in alerts about guest devices carrying no sensitive data - and an analyst who learns to ignore alerts misses the ones that matter on both networks.