PROJECT_03 – Detection Logic & Explanation

Security Monitoring & Incident Response

Overview

This document explains the detection logic used to identify suspicious activity, classify incidents, and trigger alerts during the Security Monitoring & Incident Response simulation. The logic is based on common SOC detection practices, correlation of multiple data sources, and behavioral indicators of phishing attacks.

---

1. Data Sources Used for Detection

The following log sources were analyzed to detect and validate security incidents:

- Email Security Gateway logs
- Firewall and Web Proxy logs
- Threat Intelligence / Blacklists
- User activity indicators (click behavior)

Using multiple data sources allows correlation of events and helps identify complete attack chains rather than isolated alerts.

---

2. Detection Logic for Phishing Emails

Rule Logic

An alert is triggered when an inbound email contains one or more external links and matches suspicious characteristics.

Indicators Used

- External URL embedded in email content
- Sender domain mismatch or impersonation
- Use of urgency or fear-based language
- URL shortening services (e.g., bit.ly)
- Brand impersonation (Amazon, Microsoft, HR systems)

Example Logic

IF email.direction = inbound
AND email.contains_external_link = true

AND sender_domain NOT IN trusted_domains
THEN trigger phishing alert

---

3. False Positive Identification Logic

Not all alerts indicate malicious activity. An alert is classified as a False Positive when:

- Sender domain is verified and legitimate
- URL belongs to a trusted internal or partner domain
- No malicious indicators are found
- No suspicious outbound activity is observed in proxy or firewall logs

Example

An HR onboarding email triggered a phishing alert due to the presence of an external link. However, further investigation confirmed that the sender domain and URL were legitimate and related to normal business activity. Therefore, the alert was closed as a False Positive.

---

4. True Positive Detection Logic

An alert is classified as a True Positive when one or more of the following conditions are met:

- Sender domain impersonates a legitimate organization
- URL is listed in threat intelligence or blacklist feeds
- User interaction with a malicious link is confirmed
- Firewall or proxy logs show blocked outbound attempts to malicious destinations

Example Logic

IF phishing_email_detected = true
AND outbound_connection_attempted = true
AND destination_url IN blacklist
THEN confirm true positive phishing incident

---

5. Alert Correlation Logic

Multiple alerts related to the same indicators are correlated to identify an attack chain.

Correlated Events

1. Inbound phishing email detected
2. User clicks a malicious link

3. Outbound request to a malicious domain is blocked by the firewall

This correlation confirms attacker intent and increases incident severity.

---

6. Severity Classification Logic

Severity	Criteria
Low	False positive, no threat indicators
Medium	Suspicious email, no user interaction
High	Confirmed phishing, user click or outbound attempt

Alerts involving user interaction, credential harvesting risk, or attack progression are escalated to **High severity**.

---

7. Response Trigger Logic

Once a True Positive is confirmed, the following response actions are triggered:

- Block sender domain and malicious URL
- Notify the affected user
- Remove the phishing email from all mailboxes
- Investigate endpoint activity
- Reset credentials if user interaction is confirmed

Conclusion

The detection logic used in this project demonstrates a layered defense approach combining technical controls, behavioral analysis, and alert correlation. This methodology reflects real-world SOC workflows and demonstrates the analyst's ability to distinguish between false positives and confirmed threats, enabling timely detection and effective response to phishing attacks.

Prepared by: Lala Alili
Date: February 25, 2026