

MODULO 03

Detection potenziata dall'AI

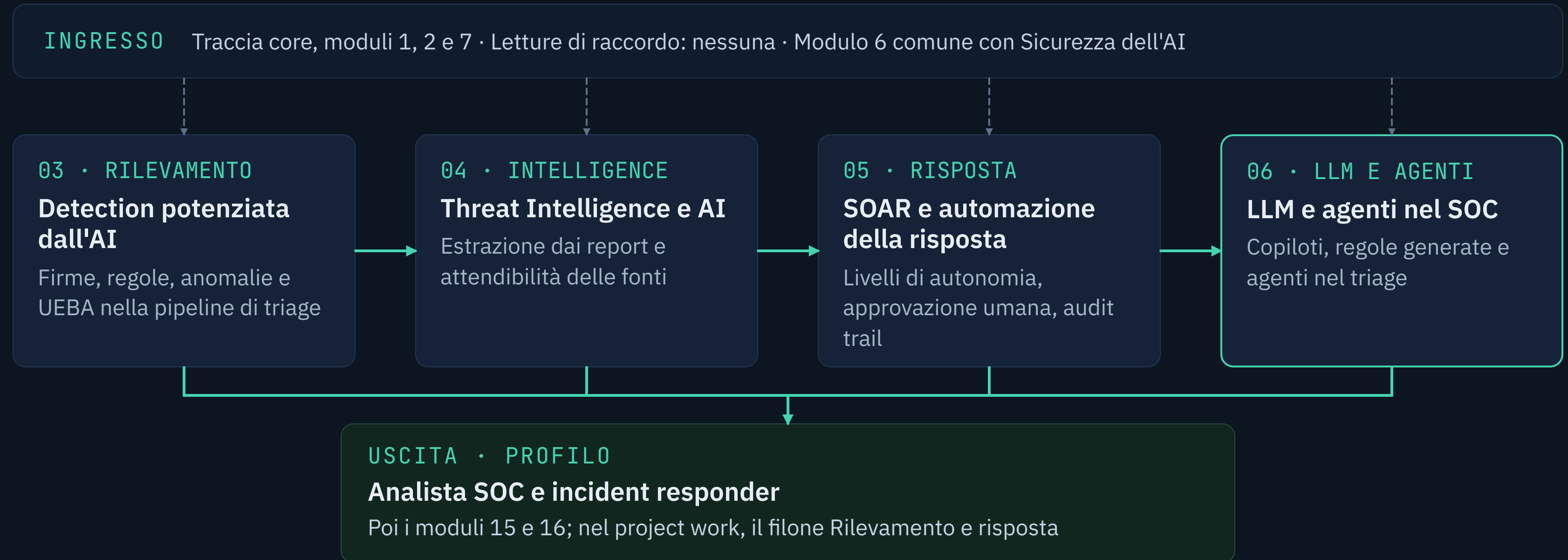
Anomaly detection, analisi comportamentale (UEBA) e triage automatizzato degli alert: contributi e limiti dell'AI nella fase di rilevamento.

UEBA

MITRE ATT&CK / D3FEND

SIEM ML

La specializzazione Detection & Response



Abstract e obiettivi formativi

Abstract

Il modulo analizza come le tecniche di anomaly detection e di analisi comportamentale trasformino la fase di rilevamento: baseline, scoring del rischio, riduzione dei falsi positivi e triage automatizzato degli alert, con integrazione nei SIEM e mappatura su MITRE ATT&CK e D3FEND.

Obiettivi formativi (Bloom)

- **Analizzare:** confrontare detection basata su firme, regole e anomalie
- **Valutare:** misurare un sistema UEBA in termini di copertura ATT&CK
- **Creare:** progettare una pipeline di alert triage con componenti ML

Traccia core

01

02

07

Detection & Response

03

04

05

06

o

Digital Forensics

08

09

10

11

o

Sicurezza dell'AI

06

12

13

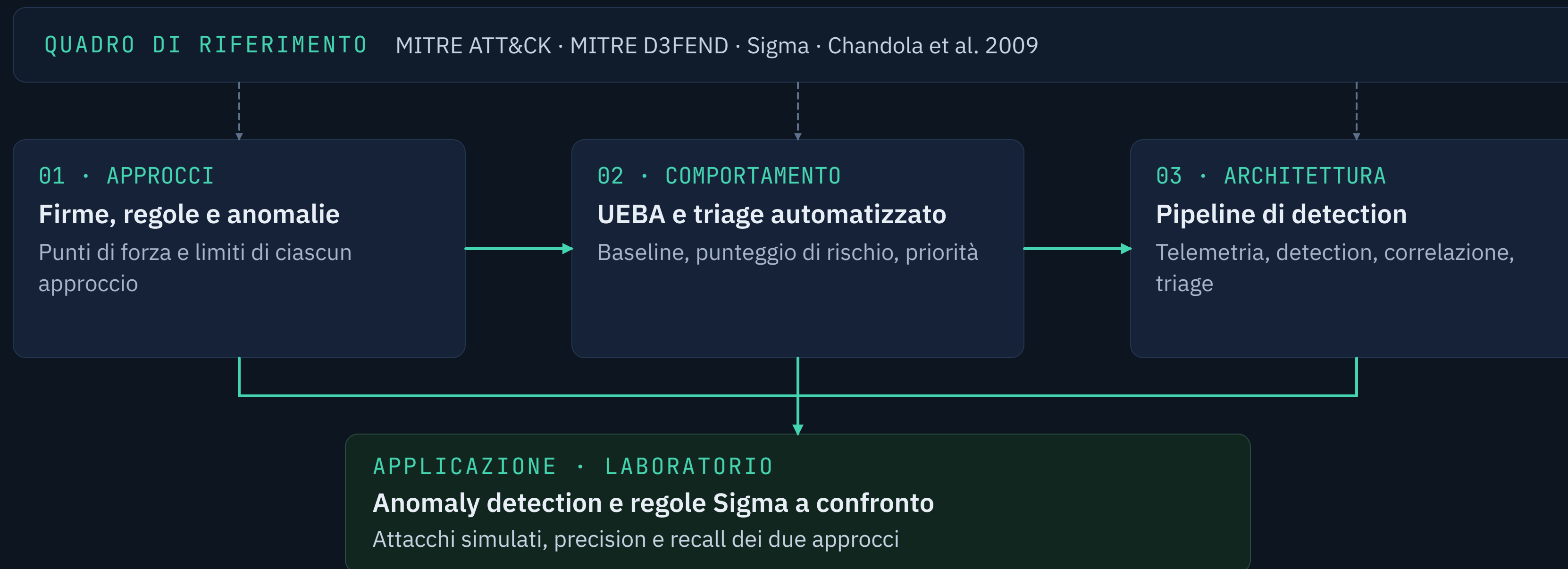
14

Traccia core

15

16

Il rilevamento assistito dal machine learning



Firme, regole e anomalie

Approccio	Come funziona	Rileva bene	Limite principale
Firme (IoC)	Confronto con indicatori noti	Minacce già catalogate	Non rileva varianti nuove o polimorfiche
Regole (Sigma)	Pattern comportamentali espliciti	TTP e comportamenti previsti dall'autore	Manutenzione, evasione mirata
Anomalie (ML)	Deviazione dalla baseline appresa	Novità e abusi interni	Falsi positivi, opacità

I tre approcci sono complementari, e la copertura si misura mappando le detection sulle tecniche MITRE ATT&CK e le contromisure su D3FEND.

UEBA e triage automatizzato

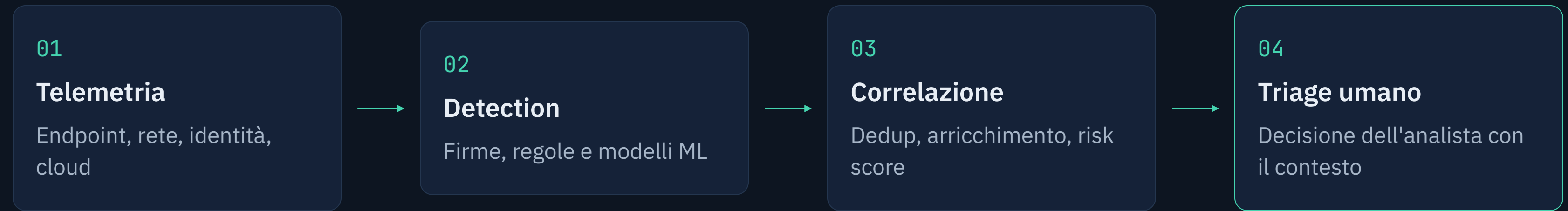
UEBA

- Baseline per utente ed entità: orari, host, volumi, geografie
- Clustering e autoencoder per lo scarto dalla norma
- Risk score cumulativo: eventi deboli che insieme diventano segnale
- Efficace su abusi interni e account compromessi

Alert fatigue e triage

- Con volumi elevati di alert la prioritizzazione diventa il collo di bottiglia
- ML per deduplicare, correlare e ordinare per rischio
- **Rischio**: un triage automatico troppo aggressivo produce falsi negativi silenziosi
- Esempi: correlazione degli incidenti di Microsoft Defender XDR, Attack Discovery di Elastic Security, risk-based alerting di Splunk ES

La pipeline di detection assistita



Misurare la copertura

Ogni detection si mappa su una tecnica ATT&CK; la matrice delle tecniche coperte, pesata per la rilevanza nel proprio threat model, misura la maturità del sistema.

Misura e manutenzione delle detection



Costruzione di una baseline comportamentale

La baseline è il modello di normalità rispetto al quale si misura lo scarto. Si costruisce per **entità**, cioè per utente, host o account di servizio, su una finestra storica abbastanza lunga da includere la stagionalità: settimane o mesi anziché giorni, perché l'attività cambia tra giorni feriali e fine settimana e nei periodi di chiusura contabile. Le dimensioni tipiche sono gli orari di attività, gli host e le aree geografiche di accesso, i volumi di dati trasferiti, le applicazioni usate e le modalità di autenticazione.

Scenario didattico

Si supponga che un amministratore, che di norma accede dalle 8 alle 18 da due postazioni, compaia alle 3 di notte, da una postazione diversa da quelle abituali, su un server che non aveva usato in precedenza. Lo scarto riguarda tre dimensioni insieme, e la combinazione di anomalie è un indizio più forte di ciascuna anomalia isolata.

Limiti

Se l'attaccante è già presente durante la fase di apprendimento, il suo comportamento entra nella baseline come normalità (Modulo 13). La baseline va inoltre aggiornata dopo ogni riorganizzazione, perché un cambio di ruolo legittimo produce scarti simili a quelli di un account compromesso.

Composizione e limiti del punteggio di rischio

In molti sistemi UEBA il punteggio di rischio è una somma pesata nel tempo: ogni anomalia aggiunge punti in proporzione alla sua rarità e alla criticità dell'entità coinvolta e, a seconda del prodotto, i contributi decadono progressivamente o escono da una finestra di osservazione se non seguono altri segnali. Il meccanismo intercetta le campagne lente e a basso profilo (low and slow), fatte di eventi singolarmente poco rilevanti che sfuggono a una regola applicata al singolo evento e che, accumulati sulla stessa entità nell'arco di una settimana, superano la soglia.

Scenario didattico

Accesso da una VPN insolita (+15), primo accesso a un file server del personale (+20), download di 2 GB fuori orario (+35), creazione di un archivio cifrato (+25): il totale di 95 punti supera la soglia di 80 e l'alert descrive una possibile esfiltrazione in corso.

Metodo

Un punteggio presentato senza la scomposizione nei suoi contributi non è verificabile né contestabile: l'analista deve vedere i quattro eventi che lo compongono. La trasparenza è un requisito operativo e, per le decisioni che riguardano persone, anche giuridico (Modulo 15).

Copertura ATT&CK come misura della detection

La capacità di rilevamento si misura con la quota di tecniche rilevanti per il proprio threat model coperte da almeno una detection. Si individuano gli attori e le campagne plausibili per il proprio settore e se ne estraggono le tecniche MITRE ATT&CK. Ogni detection esistente viene poi associata alla tecnica che rileva e il risultato si rappresenta sulla matrice, dove le tecniche rimaste scoperte indicano le priorità del detection engineering.

Scenario didattico

T1566 (phishing) è coperta da gateway di posta e sandbox, T1003 (OS credential dumping) dall'EDR. T1021.001 (Remote Desktop Protocol, sotto-tecnica di T1021 Remote Services) è invece scoperta, perché manca visibilità sul movimento laterale via RDP interno, e diventa l'obiettivo del successivo ciclo di lavoro.

Cautele

Una regola fragile su una sola fonte di log non equivale a tre segnali indipendenti. Qualificare la copertura per fonte di telemetria e verificarla con i test di Atomic Red Team evita una rappresentazione ottimistica della matrice.

Regole e modelli nella stessa architettura

Regole e modelli hanno proprietà complementari. Una **regola**, scritta in Sigma o come query di correlazione, è precisa, spiegabile, veloce e versionabile, ma rileva solo ciò che il suo autore ha previsto. Un **modello** generalizza e può segnalare comportamenti imprevisti, con più falsi positivi, maggiore bisogno di dati e minore spiegabilità. Una detection matura li combina: regole per i comportamenti noti e per vincoli di conformità come il controllo di ogni accesso fuori orario a un database critico, modelli per le anomalie e per ordinare per priorità l'output delle regole. Con il formato aperto Sigma una regola si scrive una sola volta e si converte per diversi SIEM, e la conoscenza di detection diventa condivisibile come il codice.

Esempio

Per rilevare l'uso di *certutil -urlcache* per scaricare file basta una regola Sigma, senza componenti di machine learning. Un movimento laterale lento eseguito con credenziali valide non è descritto da alcuna regola e può essere segnalato da un modello UEBA.

Ruolo dell'AI

Gli LLM producono bozze accettabili di regole Sigma a partire dalla descrizione di una tecnica (Modulo 6). Una regola generata va validata su log storici, misurandone i falsi positivi, prima del rilascio.

Alert fatigue e risk-based alerting

L>alert fatigue è l'effetto sugli analisti dell'elevato numero di falsi positivi prodotto dalla bassa prevalenza degli attacchi (Modulo 2, approfondimento 6): chi chiude ogni giorno centinaia di falsi positivi riduce progressivamente l'attenzione dedicata a ciascun alert. I segnali sono tempi di presa in carico crescenti, chiusure in blocco come benigni e regole disattivate per l'eccesso di rumore. Il **risk-based alerting** affronta il problema: invece di inviare ogni segnale, accumula punteggi di rischio per entità e genera un alert quando la somma supera una soglia, trasformando molti segnali deboli sullo stesso host in un unico alert con il contesto aggregato.

Caso documentato

Nella violazione subita da Target nel 2013 il sistema di rilevamento aveva generato alert sull'installazione del malware, che non furono gestiti in tempo. Il rapporto dello staff della commissione Commercio del Senato USA (2014) ipotizza, senza accertarlo, che siano stati presi per falsi positivi.

Metodo

Ogni regola e ogni modello hanno un costo in attenzione, stimabile come prodotto tra il numero di alert generati e il tempo medio di gestione, da confrontare con la quota di alert fondati. Una revisione periodica che corregge o disattiva le fonti più rumorose libera tempo degli analisti per gli alert rilevanti.

Detection engineering e detection-as-code

Il detection engineering tratta le detection come software sottoposto a un ciclo di sviluppo. Si parte da una minaccia prioritaria, individuata con la threat intelligence (Modulo 4) e mappata su ATT&CK, e si scrive la detection come codice versionato, con repository, revisione e integrazione continua (**detection-as-code**). La detection si verifica con attacchi simulati, con Atomic Red Team o in esercitazioni purple team, misurando falsi positivi e copertura. Dopo il rilascio se ne monitora il decadimento, perché ogni detection perde efficacia con il cambiare dell'avversario e dell'ambiente, e la mappa di copertura registra per ciascuna la data dell'ultima verifica.

Scenario didattico

Ciclo di lavoro su T1055 (process injection): la regola è scritta e salvata nel repository con la documentazione, un test automatico esegue il corrispondente test di Atomic Red Team in staging, la soglia è tarata su due settimane di log storici e la regola viene rilasciata. Sei mesi dopo il test periodico fallisce perché un aggiornamento dell'EDR ha cambiato il formato dei log.

Ruolo dell'AI

L'AI può accelerare ogni fase, dalle bozze di regole alla generazione di varianti di test e al triage del backlog. Restano decisioni umane la scelta delle minacce prioritarie e il giudizio su quando una detection è adeguata, perché dipendono dalla valutazione del rischio dell'organizzazione.

Anomaly detection e regole Sigma a confronto

Svolgimento (Security Onion / Elastic)

1. Creazione di un job di anomaly detection sui log di autenticazione
2. Scrittura delle regole Sigma equivalenti per gli stessi scenari
3. Generazione di attacchi simulati con Atomic Red Team
4. Misura comparata di precision e recall dei due approcci

Consegna

- Tabella precision/recall per approccio
- Mappatura delle detection su ATT&CK
- Analisi dei rilevamenti e dei falsi positivi aggiunti dall'ML

Lecture e autovalutazione

Lecture obbligatorie

- Chandola et al., *Anomaly Detection: A Survey*, ACM CSUR 2009
- Documentazione MITRE D3FEND
- ENISA Threat Landscape, edizione corrente

Domande di autovalutazione

1. Quali attacchi sfuggono alla detection basata su firme?
2. Come si costruisce una baseline comportamentale affidabile?
3. Quali rischi comporta un punteggio di rischio opaco?
4. In che modo il triage automatico può introdurre falsi negativi?
5. Come si misura la copertura ATT&CK di un set di regole?

FINE DELLA LEZIONE • MODULO 03

Grazie

Detection potenziata dall'AI • Corso «Incident Response potenziata dall'AI»

Domande e materiali: ricevimento del docente. Profilo e contatti:
vincenzocalabro.it



Vincenzo Calabrò

CYBER SECURITY • DIGITAL FORENSICS



VINCENZOCALABRO.IT