Contents

[9.8] Disfunzione del Team Umano-IA 1

[9.8] Disfunzione del Team Umano-IA

1. Definizione Operativa: Il deterioramento della collaborazione effettiva, della comunicazione e della chiarezza dei ruoli tra i membri umani del team di sicurezza e gli agenti IA, portando ad azioni non allineate, sforzi duplicati, o compiti critici che vengono omessi.

2. Metrica Principale e Algoritmo:

- **Metrica:** Tasso di Conflitto di Compiti (TCR). Formula: `TCR = N_azioni_conflittuali / N_coppie_azione_umano_IA_totali`.
- **Pseudocodice:**

```
def calculate_tcr(ai_actions, human_actions, start_date, end_date):
    # Ottenere azioni sullo stesso avviso entro una finestra di tempo breve
    conflicting_pairs = []
    time_window = timedelta(minutes=5)

    for h_action in human_actions:
        for a_action in ai_actions:
            if (h_action.alert_id == a_action.alert_id and
                abs(h_action.timestamp - a_action.timestamp) < time_window and
                h_action.action != a_action.action):
                conflicting_pairs.append((h_action, a_action))

    # Stimare le coppie potenziali totali per un rapporto
    total_potential_pairs = ... # Calcolo complesso basato su assegnamenti e azioni sovrap
    # Un approssimativo più semplice: Numero totale di avvisi lavorati sia da IA che uman

    N_conflicts = len(conflicting_pairs)
    # Utilizzare un approssimativo denominatore più semplice
    N_alerts_worked = count_alerts_worked_by_both(start_date, end_date)

    if N_alerts_worked > 0:
        TCR = N_conflicts / N_alerts_worked
    else:
        TCR = 0

    return TCR
```

- **Soglia di Avviso:** `TCR > 0.1` (Le azioni conflittuali si verificano su più del 10% degli avvisi lavorati da entrambi).

3. Fonti Dati Digitali (Input dell'Algoritmo):

- **API di SOAR/SIEM:** Log dettagliati di tutte le azioni intraprese da agenti IA e analisti umani (`actor_type`, `alert_id`, `action`, `timestamp`).

4. Protocollo di Audit Umano-Umano: Eseguire un workshop che simula una risposta a un incidente. Includere l'IA come membro del team. Osservare il flusso di lavoro: Gli umani sanno cosa sta facendo l'IA? L'IA mina i comandi umani? C'è confusione sui ruoli?

5. Azioni di Mitigazione Consigliate:

- **Mitigazione Tecnica/Digitale:** Migliorare il "gioco di squadra" dell'IA facendole trasmettere le sue intenzioni e azioni chiaramente tramite un canale di notifica dedicato (es. "Sto prioritizzando l'avviso X").
- **Mitigazione Umana/Organizzativa:** Definire chiaramente il ruolo dell'IA in charter di team (es. "L'IA è un assistente per lo screening, non per il processo decisionale finale").
- **Mitigazione di Processo:** Progettare playbook chiari che specificano quale agente (umano o IA) è responsabile di quale azione specifica in uno scenario dato per prevenire sovrapposizione e conflitto.