la rilevazione automatica di problemi/anomalie

Intrusion Detection Systems e affini

“problema” vs. “anomalia”

- sono sinonimi nel linguaggio comune
- il termine **anomalia** è usato tipicamente in “anomaly detection” con un significato specifico: “qualcosa di diverso rispetto ad una normalità ben definita”
- in queste slides, per **problema** intendiamo qualcosa di più generale
  - es. attacchi, configurazioni errate, pacchetti inattesi, sono tutti problemi che vorremmo rilevare
  - in cui la “normalità” non è detto che sia chiaramente definita

aree di applicabilità di ciò che segue

- log/event audting
  - input: sequenza degli eventi o linee di log (real-time o off-line)
- Network Intrusion Detection Systems (NIDS)
  - input: pacchetti di rete
- Host Intrusion Detection Systems (HIDS)
  - input: system calls, pacchetti di rete, email, pagine web, ecc.
- Intrusion Prevention Systems
  - vedi NIDS
- antivirus
  - vedi HIDS

l'automatismo necessario

la rilevazione automatica dei problemi prevede…

1. **vagliare** una grande quantità di dati
   - pacchetti di rete
   - "eventi" o righe di log
     - anche da migliaia di fonti distinte
   - system call
     - normalmente non loggate perché troppe e troppo veloci
   - **intervento umano impossibile**
     - per la quantità di dati e la velocità con cui devono essere processati
2. **riconoscere** le "*parti problematiche*"
   - **difficile da automatizzare**
     - perché richiede una definizione formale che può non esistere

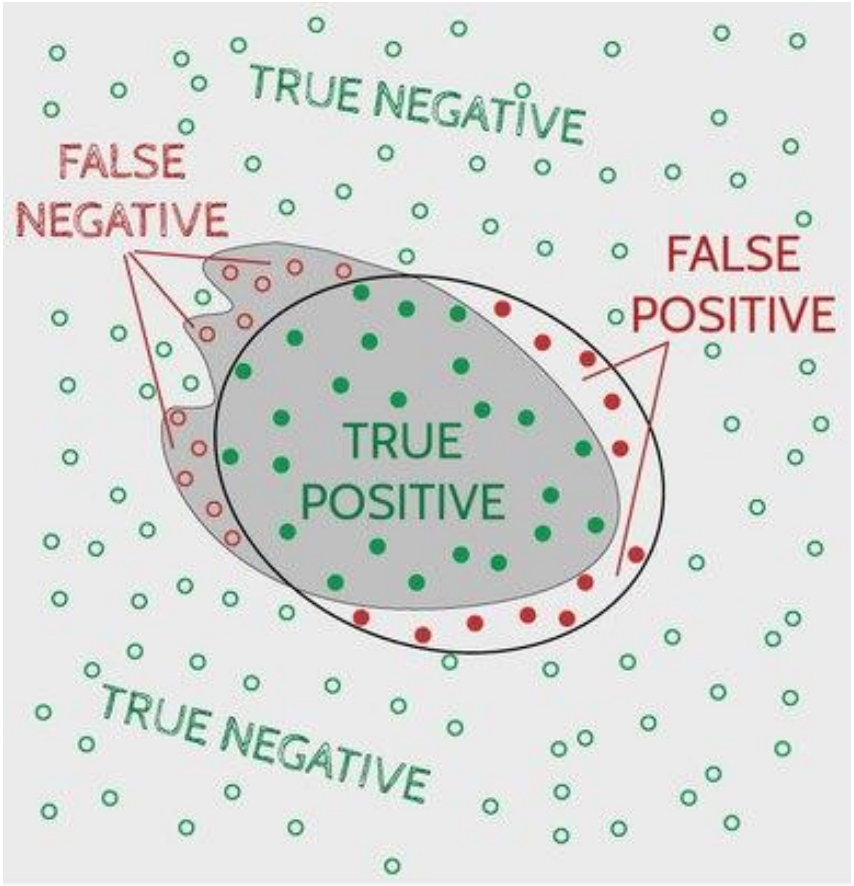
è un *problema* o no?

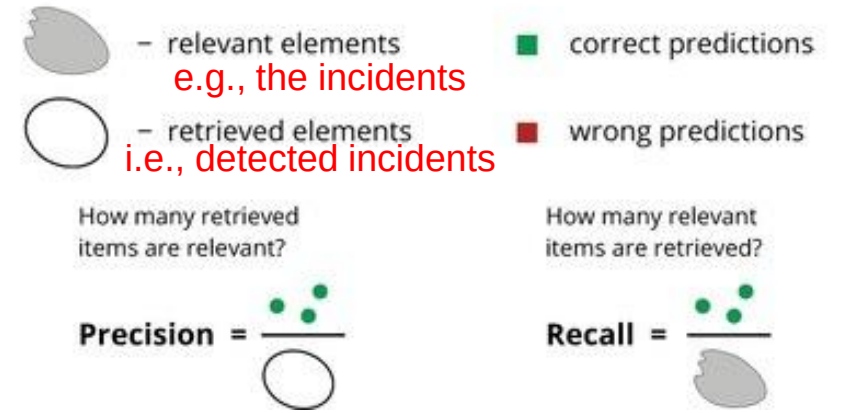
- la definizione di cosa è un *problema* è…
  - **soggettiva**
    - persone diverse potrebbero avere opinioni diverse
  - **variabile nel tempo**
    - per le esigenze dell'organizzazione
    - per nuovi attacchi
    - per cambiamenti delle apparecchiature o del business
  - **potenzialmente ambigua**
    - in mancanza di una definizione formale
      - es. tale definizione potrebbe essere data in un piano di sicurezza
    - la definizione formale potrebbe essere incompleta

il grande problema dei “falsi”

- **falsi positivi**: il sistema **rileva** problemi che non esistono
  - misurati in % su totale degli allarmi forniti
- **falsi negativi**: il sistema **NON rileva** problemi che dovrebbe rilevare
  - misurati in % sul totale dei problemi da rilevare... **molto difficili da misurare!**
    - Infatti raramente conosciamo tutti i problemi che vogliamo rilevare (ciò avviene solo in situazioni di laboratorio molto controllate)
- tutti i sistemi possono sbagliare dando falsi positivi o negativi
  - idealmente vorremmo 0% di falsi positivi e negativi

false positive/negative vs. precision & recall





Cybersecurity terminology vs. AI terminology

F=false, T=true
P=positive, N=negative

% of FP=
FP/(TP+FP)=
1-precision

% of FN= FN/incidents=
FN/(FN+TP)=
1-recall

cosa è peggio tra falsi positivi e negativi?

- in intelligenza artificiale: dipende dall'applicazione
- **in cybersecurity: i falsi negativi**
- **falsi positivi**
  - noiosi
  - Costosi
  - possono innescare una procedura per migliorare la rilevazione (taratura del sistema di rilevazione)
- **falsi negativi**
  - **pericolosi!**
  - **attacchi o problemi non rilevati possono essere molto costosi**
  - poiché non sono rilevati difficile avere una procedura per migliorare la rilevazione
    - ciò è possibile solo se c'è un'altra fonte di rilevazione di problemi e comunque sempre nei limiti dell'accuratezza di quest'ultima

due approcci

- **rule-based**
  - **regole formalizzate** in qualche linguaggio
    - es. espressioni regolari o linguaggi ad-hoc
    - spesso contengono una *signature* dell'attacco
      - **signature** = sottostringa di pacchetto o file che caratterizza l'attacco
  - **matching automatico**
    - es. con automi a stati finiti
- **anomaly-based (anomaly detection)**
  - apprendimento del "**comportamento normale**"
    - tecniche di AI (machine learning) o statistiche
    - alle volte possono richiedere approcci big data.
  - rilevazione di comportamenti "**anomali**"
    - cioè che deviano rispetto al comportamento normale appreso

rule-based

- tutti gli strumenti basati su regole vanno “tarati” per le esigenze del caso
  - aggiungendo e/o togliendo regole
- la taratura...
  - richiede risorse umane
    - la taratura viene fatta da un “amministratore”
  - **complessa e quindi può essere errata**
  - si basa sull'esperienza e quindi richiede tempi lunghi ed è **costosa**
    - il sistema non sarà efficace sin da subito
- **molto più facile ridurre i falsi positivi che i falsi negativi!**
  - perché i falsi negativi non sono evidenti
  - notare che vorremmo ridurre i falsi negativi più che quelli positivi!

rule-based

- l'aggiornamento del set di regole può essere…
  - condiviso tra molti utenti in modo da fare economia di scala
  - con download automatico delle nuove regole
- sistemi obbligatoriamente associati ad un servizio in cloud e **non sono configurabili esplicitamente**
  - semplicità
  - riduzione costi
  - nessuna consapevolezza per l'utente

Efficiency of rule-based detection

- Growing number of malwares → large rule/signature DBs.
- Naïve matching is computationally expensive
  - linear scan per each signature

- In practice: **single scan to check for all signatures!**
- Based on Finite Automata (FA) and Regular expressions (regex) theory
  - Signatures DB = huge regex mapped on a deterministic FA

- Used in rule-based anti-viruses, intrusion detection systems, etc. , and their evolutions

anomaly-based

- generalmente hanno due modalità operative
  - **learning mode (o training mode)**: per apprendere il comportamento normale
  - **detection mode**: per rilevare anomalie
- **rappresentazione interna non accessibile**
  - potrebbe non essere “esplicita”, es. bayesian network, neural network, SVM, ecc.
- nei sistemi puramente anomaly-based è generalmente impossibile effettuare tarature manuali mirate
  - è necessario un re-training

anomaly-based

- learning mode: **input deve essere rappresentativo**
  - non deve contenere attacchi!
  - deve contenere **tutti i casi "normali"**
- quando il sistema cambia comportamento deve essere rilanciato il learning
- **continuous learning**
  - il sistema **si adatta lentamente ai cambiamenti**
  - rilevato solo l'inizio dell'anomalia, poi il sistema assorbe l'anomalia nel comportamento normale appreso
  - **non rileva attacchi molto "lenti"**

anomaly-based

- difficile da adottare in pratica in ambienti IT standard
  - nuove applicazioni e protocolli generano allarmi, necessario re-training
- adottabile in ambienti molto statici
  - es. sistemi di controllo industriale, sistemi militari

approccio ibrido

- una combinazione dei due precedenti
- la combinazione può essere fatta in vari modi
  - anomaly-based con possibilità di inibire certi allarmi mediante regole (per evitare falsi positivi)
  - unione di allarmi derivante dall'approccio rule-based e anomaly-based
    - per rilevare anomalie non presenti nelle regole
    - per avere un livello maggiore di "certezza" quando si attiva una regola

parametri di qualità/prestazioni

- **accuratezza** nella rilevazione dell'intrusione
  - nessuno o pochi falsi positivi
  - nessuno o pochi falsi negativi (più pericolosi!)
- **tempestività** nella rilevazione dell'intrusione
- **throughput** massimo
  - ... senza che l’elevato carico introduca maggiori falsi negativi
- **semplicità** di configurazione
- **integrazione** con sistemi di contrasto immediato all'attacco (*intrusion prevention systems*)

la valutazione è difficile

Obiettivo 1: dato un sistema di rilevazione automatica dei problemi, dare una **misura della sua efficacia**

Obiettivo 2: dati due sistemi di rilevazione automatica dei problemi dire **quale dei due è "migliore"**

- problema della copertura dello "spazio dei problemi/attacchi da rilevare" (approccio black-box)
  - problemi importanti domani possono non essere noti oggi
  - l'importanza di un problema non definita oggettivamente
- spesso non si conoscono i metodi di funzionamento interno, le regole o gli algoritmi sottostanti (approccio white-box)
  - quindi non si riescono a fare ragionamenti specifici
- **non esiste metodologia efficace di valutazione o comparazione né black-box né white-box**

UEBA

- User and Entity Behaviour Analysis
- A currently used commercial buzzword that expresses some form of learning-based anomaly detection technique
- Entity = devices, applications, systems
- Behaviour = login hours and locations, amount of data transferred, contacted systems, etc.