MI-PB-18

Detekce síťových útoků a anomálií, prevence útoků, statistické aspekty detekce útoků.

Detekce a prevence útoků

Aktivní obrana: ofenzivní techniky, ale v defenzivním přístupu
Pasti v systému, na které útočník narazí

Cyber Deception: Proces oklamání útočníka tak, aby byl zpomalen a zmaten -- aby $t_{\text{detekce}} + t_{\text{reakce}} < t_{\text{útok}}$

Právní dopady:

- Matení a obfuskace bez problému
- Protiútok -- konzultovat s právním oddělením
- Používat varovné bannery, podmínky používání
- Entrapment (přesvědčení útočníka k nelegální činnosti, kterou by jinak nespchal -- nelegální)
- Enticement (útočník by trestnou činnost spáchal stejně -- honeypots)

Fáze:

- **Annoyance:** plýtvání útočníkovým časem
- **Attribution:** zjištění, kdo útočí (Google, DNS tools, port scan, credential harvesting, location)
- **Attack:** spuštění kódu na útočníkově systému

OODA cyklus:
Útočníkův cyklus je napřed oproti obráncovi. Rychlejší cyklus vyhrává

- Observe
- Orient
- Decide
- Act

Kill Chain:
Posloupnost činností, kterými útočník postupuje

- **Reconnaissance:** zjišťování informací, identifikace a výběr cílů (narušení: klamné informace)
- **Weaponization:** Spojení malwaru pro vzdálený přístup s exploitem, tvorba doručitelného payloadu (narušení: falešné nasměrování)
- **Delivery:** přenos payloadu do cíle (narušení: odvrácení)
- **Exploitation:** spuštění škodlivého kódu (narušení: oklamání útočníka, aby si myslel že malware byl spuštěn)
- **Installation:** instalace backdooru do systému (narušení: obfuskace)
- **Command & Control:** komunikace malwaru a vnějšího serveru (narušení: odchycení komunikace)
- **Actions on Objective:** útočník získá cílová data/přístupy/... (narušení: co největší pozdržení)

Obfuskace prostředí:

- Změna identifikace web serveru
- Změna TCP/IP protokol stacku v OS

Pasti na spidery/crawlery:

- Poskytnout jim random linky, nekonečně rekurzivní adresáře
- Nedělat na zvenčí přístupném serveru (zřídit si `robots.txt`)

Honeypots:

- Objekty zřízené k tomu, aby s nimi interagoval útočník
- S honeypotem nepracuje žádná komponenta systému/sítě -- jakákoliv interakce znamená útok
- Honeynet: síť honeypotů
- Honeytables: tabulky v DB s nesmyslnými daty
- Honeyports: porty sloužící k blacklistování útočníkova systému

Útočník používá proxy: Cíl: na jeho systému spustit aplikaci, která proxy nepoužije (Office, Flash, Java...) $\Rightarrow$ získání skutečné IP adresy

Statistické aspekty detekce útoků

Založené na statistickém rozdělení síťového provozu

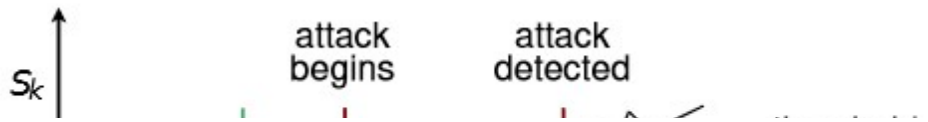
Nejjednodušší statistický model: spočtení parametrů hustoty pravděpodobnosti pro každou známou třídu provozu, otestovat neznámý vzorek a určit, do které třídy patří

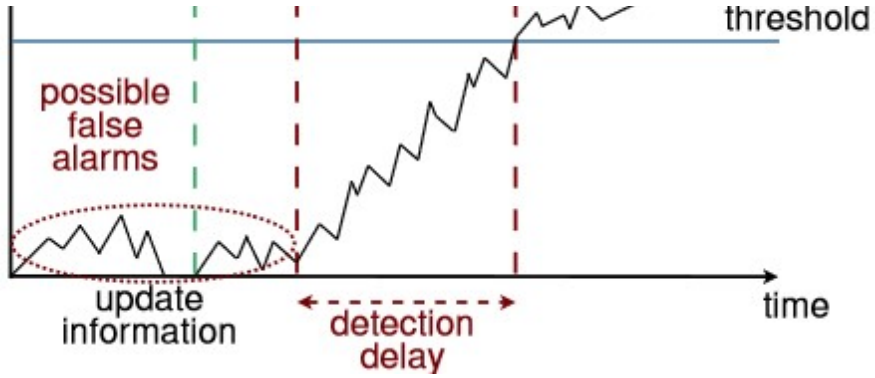
Parametrický test: předpokládá znalost rozdělení a odhadu jeho parametrů z daných dat
Neparametrický: nepředpokldá znalost rozdělení ani parametrů

Ne-statistický přístup: Protokoly deterministické -- detekovat anomálie lze stavovou analýzou
Statistický přístup: Útoky probíhají náhodně v neznámých časech a vedou ke změnám statistických vlastností některých pozorovatelných charakteristik

Detekce útoků jako Change-Point Detection:

- Detekce změn v rozdělení s fixním zpožděním, udržení falšených poplachů na dané úrovni
- Pozorovaná sekvence náhodných proněmmých $X_1, ..., X_n$ pozorovaná v časech $t_1, ..., t_n$ (např. počet deautentizačních rámců, počet neúspěšných připojení, ...)
- Změna v rozdělení se projeví v neznámém indexu λ:
  - Změna odpovídá anomálii v čase t_λ
  - $P_k = P(\lambda = k)$
  - P_0: rozdělení před změnou
- V čase τ byl spuštěn "alarm" (detekce změny):
  - Zpoždění detekce: $ADD_\lambda(\tau) = E_\lambda(\tau - \lambda | \tau \geq \lambda)$
  - Poměr falešných poplachů: $FAR(\tau) = \frac{1}{E_0(\tau)}$
- Podmíněná pravděpodobnostní hustota:
  - Před změnou ($n < \lambda$):
    $p_0(X_n | X_1, ..., X_{n-1})$
  - Po změně ($n \geq \lambda$):
    $p_1(X_n | X_1, ..., X_{n-1})$
  - Log-likelyhood ratio:
    $Z_{n,\lambda} = \sum_{k=\lambda}^{n} \log \frac{p_1(X_k | X_1, ..., X_{k-1})}{p_0(X_k | X_1, ..., X_{k-1})}$
- Klasická Change-Point Detection
  - Statistika založená na Log-likelyhood Ratio:
    - Page's Cumulative sum:
      $U_n = \max_{1 \leq \lambda \leq n} Z_{n,\lambda}$,
      alarm v čase $\tau_{CU}(h) = \min\{n \geq 1 : U_n \geq h\}$
    - SR prodecure:
      $R_n = \sum_{\lambda=1}^{n} \exp\{Z_{n,\lambda}\}$,
      alarm v čase $\tau_{SR}(h) = \min\{n \geq 1 : \log R_n \geq h\}$
  - Detekce, když statistika překočí daný práh
  - Test hypotézy, že se změna objevila v čase λ versus že se žádná změna neobjevila





- Při iid obě metody minimalizují worst-case průměrné zpoždění detekce (ADD)

Zahlazování:

- Exponential Weighted Moving Average (EWMA):
  - Krátkodobé vrcholy - produkce falešných poplachů
  - Zahlazení pozorovaných charakteristik při daném koeficientu zahlazení

Výkonnostní metriky:

- Test Power:
  - $PWK^k = P(\tau \leq k | \lambda \leq k)$
  - *Šance, že když anomálie nastala v čase menším než k, bude i detekována v čase menším než k*
  - Chceme co nejvyšší
- Probability of False Alert:
  - $PFA^k = P(\tau \leq k | \lambda > k) = P_0(\tau \leq k)$
  - *Šance, že anomálie byla detekována před časem k, když přitom nastala až po čase k*
  - V praxi: podmíněné PFA pro časový interval T:
    - $PFA_T^k = P(\tau < k + T | \tau \geq k, \lambda \geq k + T)$
- Run Length: střední počet detekcí v čase τ
- $FAR(\tau)$
- Zpoždění detekce ($ADD_\lambda(\tau)$)

Vyhodnocovací kritéria:

- Přesnost: jak korektně IDS pracuje (procenta detekcí a chyb)
- Kvalita dat:
  - Kvalita: legitimita zdroje, výběr vzorků
  - spolehlivost: přenost, konzistence
  - validita: platná data (dobré hodnoty v očekávaném rozsahu)
  - kompletnost: reprezentace prostoru zranitelností a útoků zachytitelných IDS

- Korektnost:
  - ROC křivka: Receiver Operating Characteristics
  - Confusion Matrix: dělení na True Positive, True Negative, False Positive, False Negative
  - Misclassification Rate: $\frac{FN+FP}{TP+FP+FN+TN}$
- Efektivita:
  - stabilita: výkon konzistentní v různých sítích
  - timeliness: zpoždění mezi časem útoku a reakcí
  - performance: využití CPU a paměťi
  - update profile: možnost přidat nové signatury
  - interoperability: schopnost korelace informací z více zdrojů
  - unknown attack: schopnost detekovat neznámé vzorce útoků