Foto: Försvarsmakten/David Gernes

AI-styrd penetrationstestning

hannes.holm@foi.se

Många IT-säkerhetstester kräver ”skarpa” angrepp

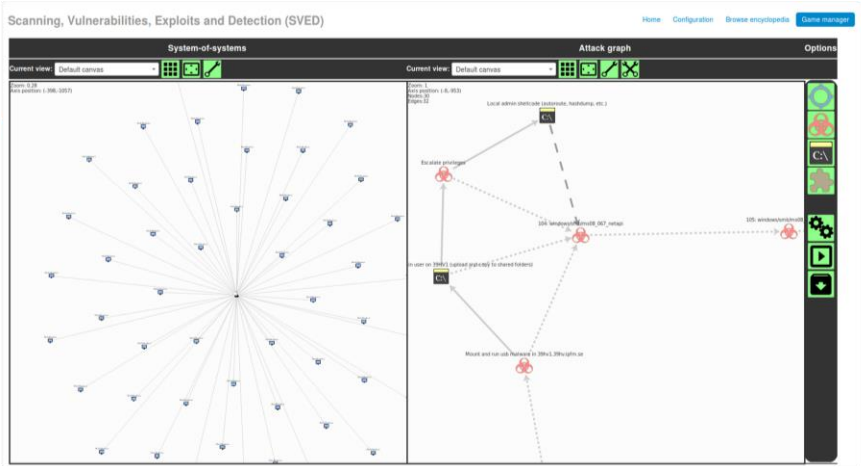
- Identifiering och validering av sårbarheter
- Övning/utbildning av logganalytiker
- Tester av tekniska säkerhetsmekanismer

- Idag görs tekniska IT-säkerhetstester främst manuellt av dyra och eftertraktade specialister
- Metodik skiljer sig mycket, loggning är begränsat

SVED – ett verktyg för planering och exekvering av IT-angrepp

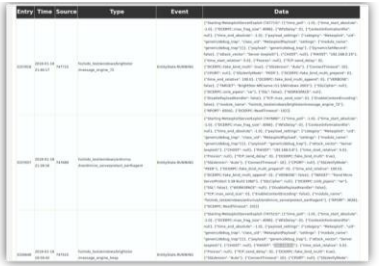
Planeringsstöd, t.ex.

- Detaljer om sårbarheter i datorer
- Information om angrepp av olika slag



Stöd för att designa komplicerade angrepp

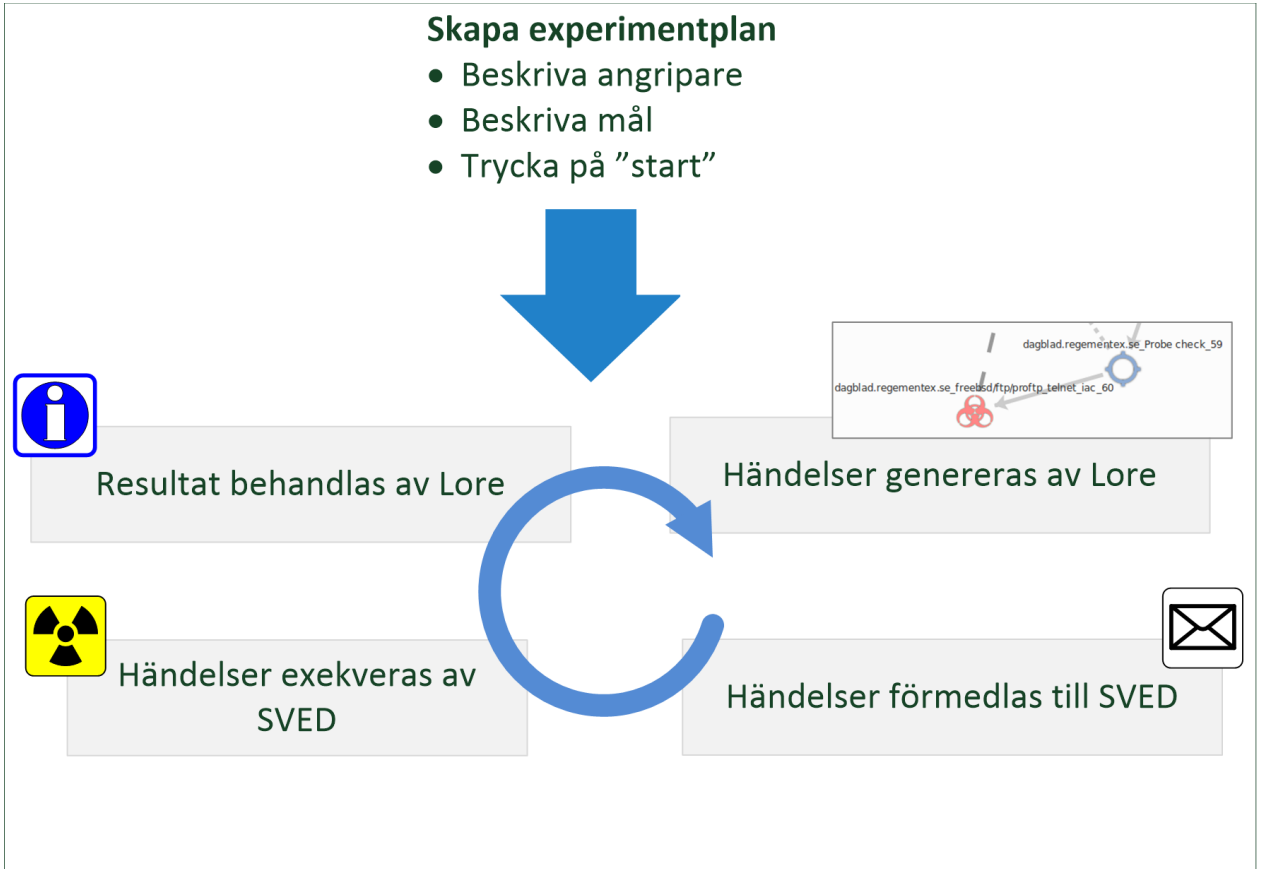
- Grafiskt gränssnitt och programmeringsgränssnitt mot angreppsverktyg
- Specifikation av händelsekedjor med dynamiska beslut och fördröjningar



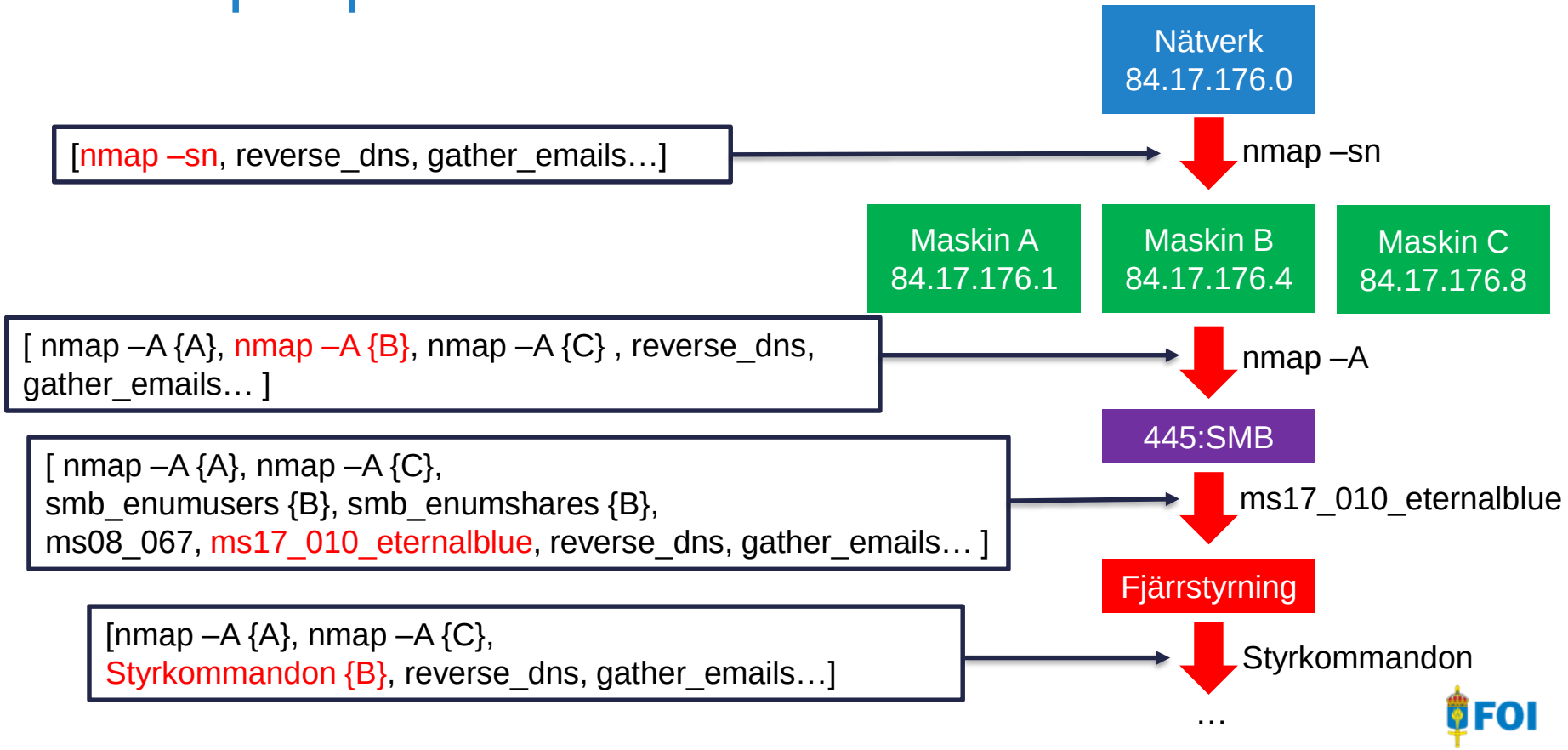
Detaljerade och precisa loggar

- Händelser loggas med millisekundsprecision.
- Integration med bland annat verktyg för att upptäcka angrepp

Lore – en artificiell intelligens för IT-angrepp



Exempel på händelsesekvens



Prioritering av aktiviteter

- **Prioritet = Sannolikhet * Kostnad * Värde**
  - **Sannolikhet** (0-1)
    - Hur sannolikt är det att en aktivitet lyckas?
  - **Kostnad** (0-1)
    - Synliga/spårbara artefakter?
    - CPU-tid?
    - Kalendertid?
  - **Värde** (0-1)
    - Hur mycket närmre målen kommer hotaktören givet att aktiviteten lyckas?

FOI

Prioritering av aktiviteter

- **Sannolikhet** (0-1)
  - Maskininlärning för angrepp. Hårdkodat för allt annat.
- **Kostnad** (0-1)
  - =1
- **Värde** (0-1)
  - =1

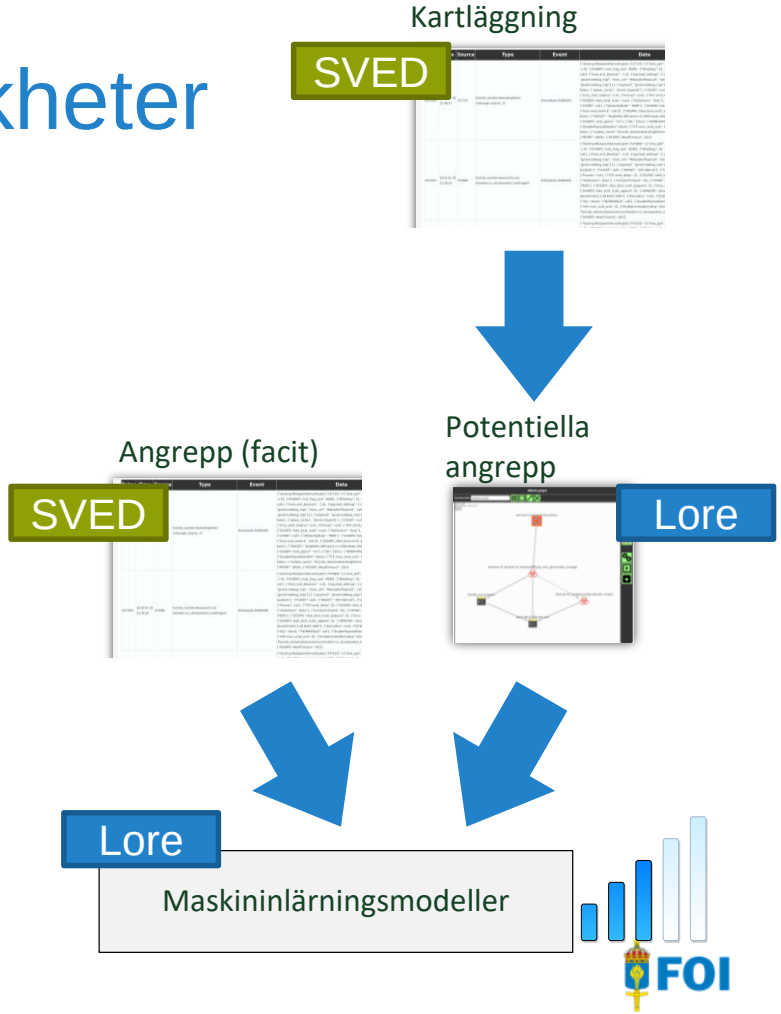
Maskininlärning av sannolikheter

1. Angrepp mot, och kartläggning av, alla datorer som var igång i datorklustret CRATE på FOI under sommaren 2017.

2. Inmatning i Lore
   – Cirka 90 000 teoretiska angrepp, varav ~400 skulle lyckas i verkligheten

3. Tester med Bayesianska nätverk, Support Vektor Maskiner, neurala nätverk och random forest
   – Ganska dålig verklig prediktion oavsett modell (cirka 9 felaktiga angrepp innan ett lyckat)

- Många förbättringsmöjligheter
  – T.ex. var 90% av alla CPE:er (mjukvarokoder) refererade i kartläggningar med nmap felaktiga
  – Fler prov från en större variation av system



Sammanfattning

- Lore är en AI som automatiserar IT-säkerhetsanalyser via planering- och exekveringsramverket SVED

- Lore fungerar idag rent tekniskt, men behöver förbättras avsevärt för att vara praktiskt användbart
  - Stöd för fler typer av moduler. T.ex. stöds ej responder.py, powershell empire eller crackmapexec.
  - Smartare prioritering av händelser