Självlärande System för Cybersäkerhet

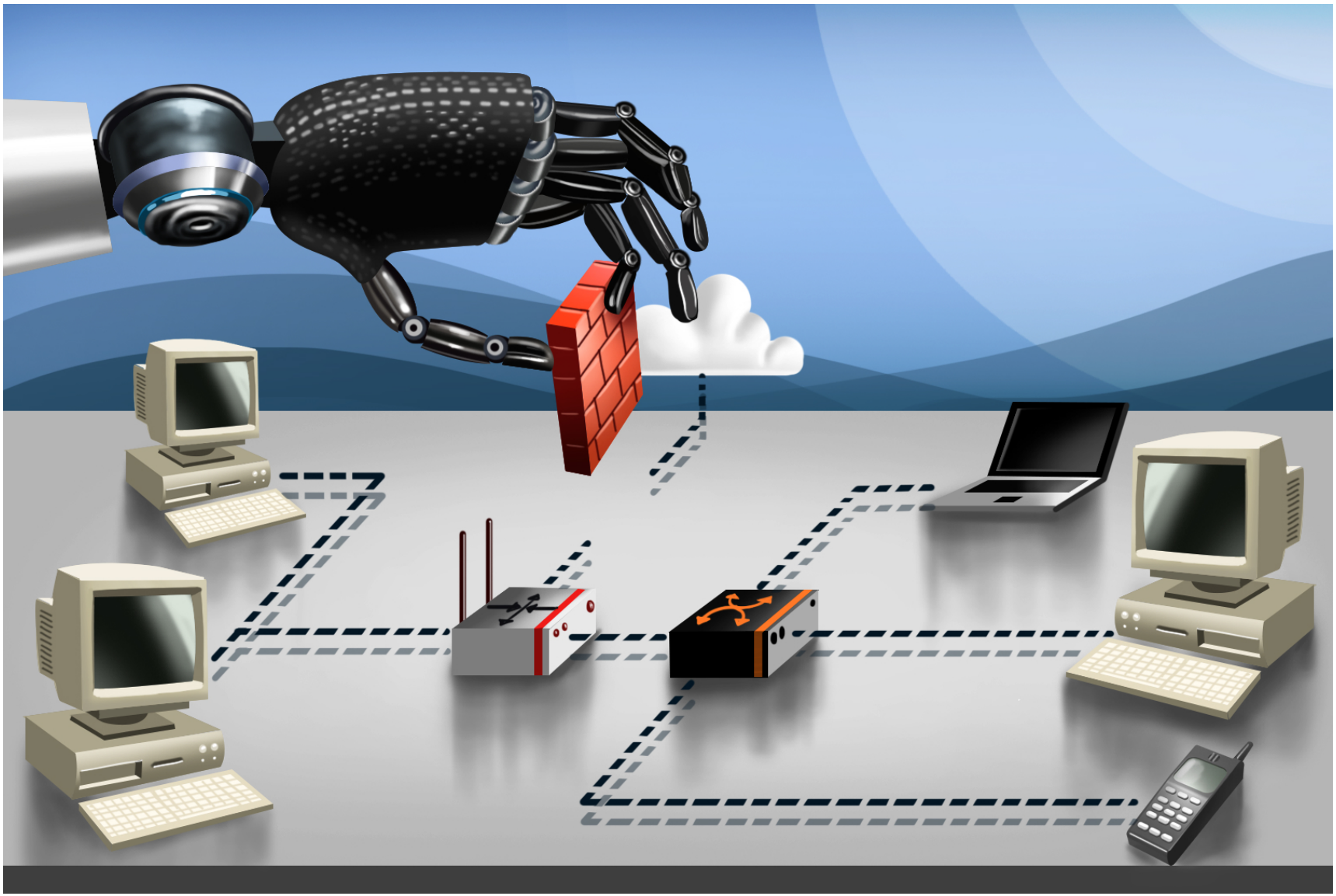
Kim Hammar

kimham@kth.se

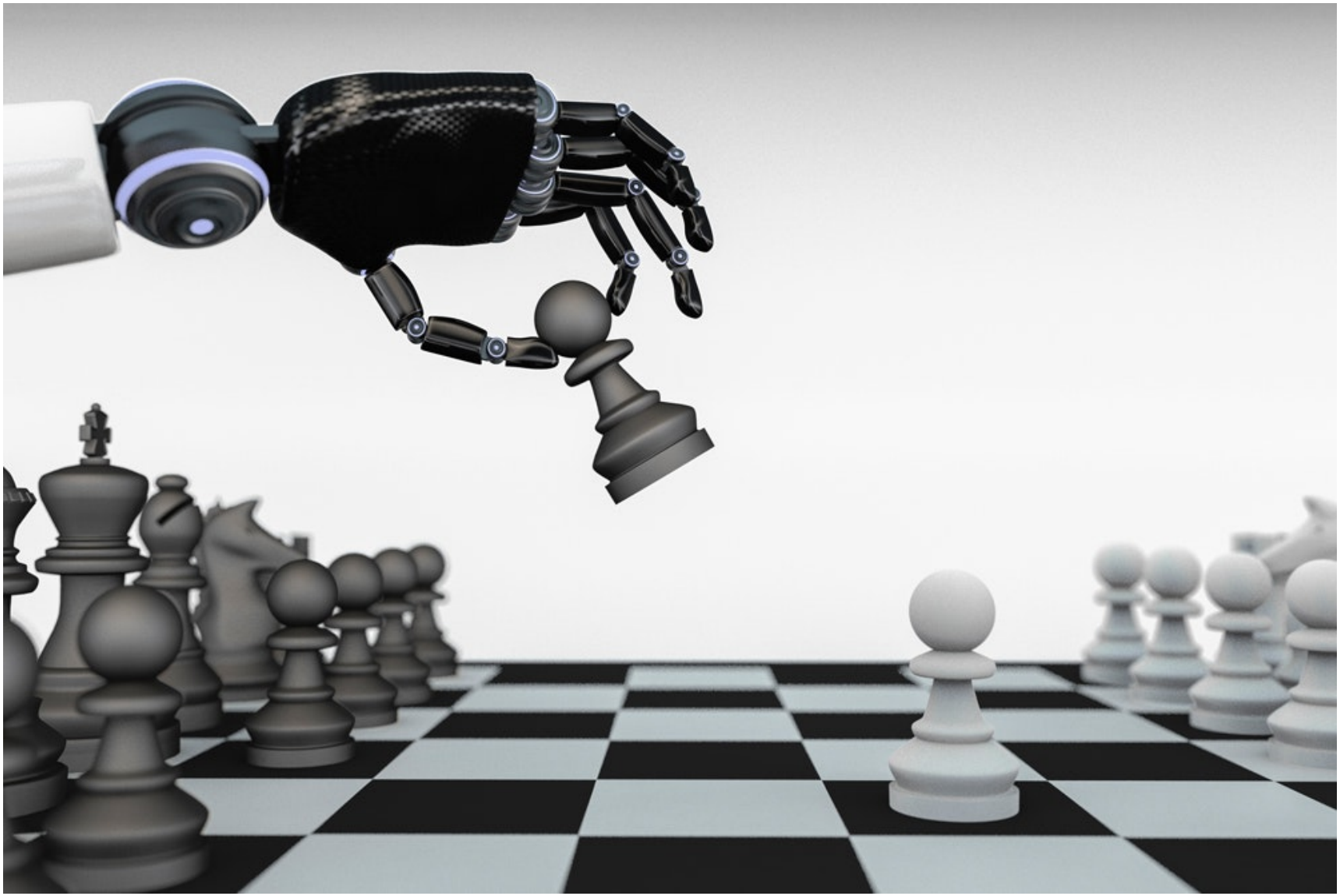
CDIS, Centrum för cyberförsvar och informationssäkerhet
NSE, Avdelningen för nätverk och systemteknik
KTH Kungliga Tekniska Högskolan

11 Jan, 2024

Hur kan vi mitigera cyberangrepp med AI?

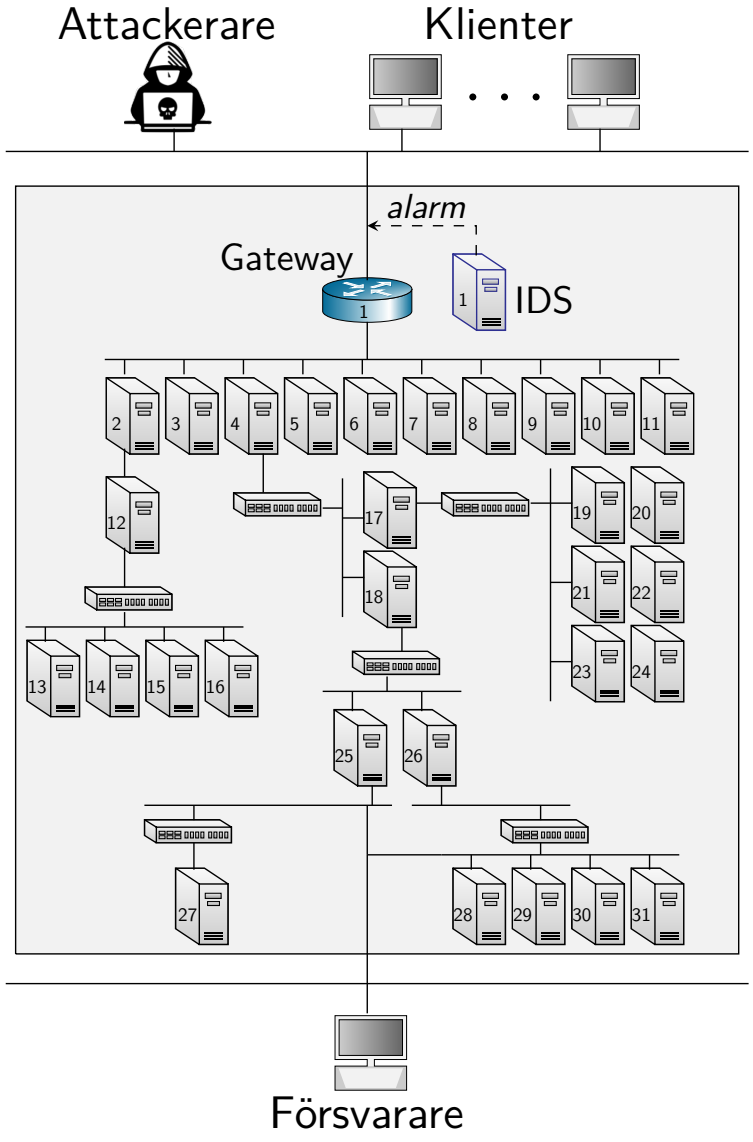


Hur kan vi mitigera cyberangrepp med AI?

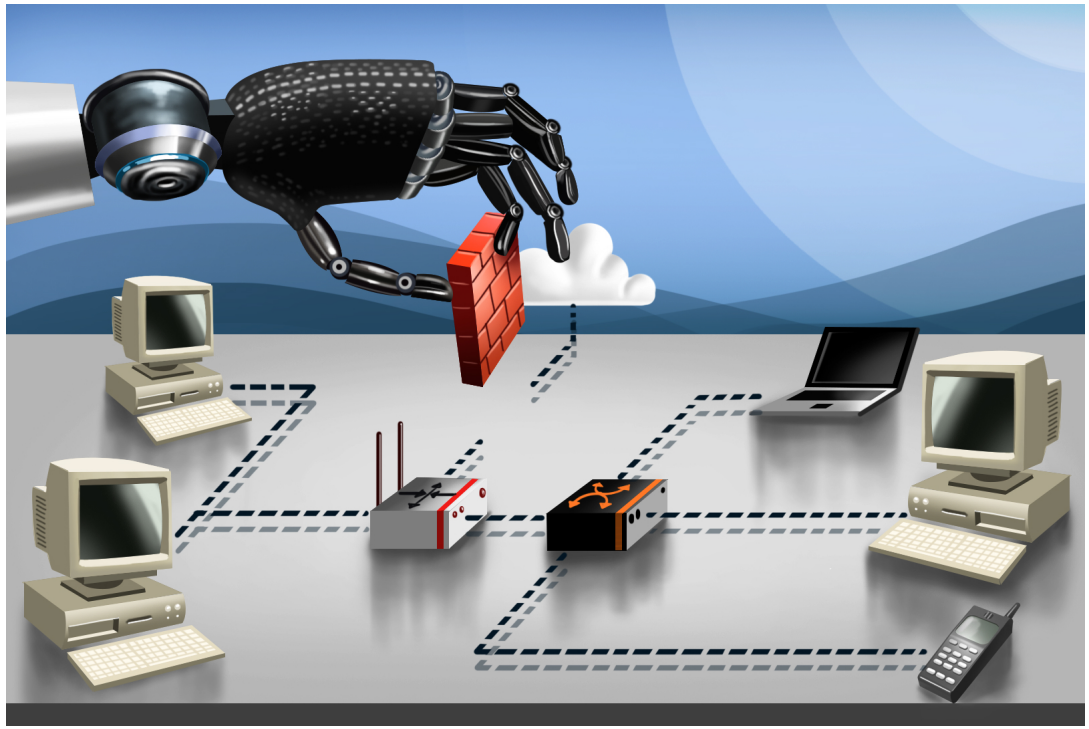


Motivation

- **Utmaningar**:
  - Attackmetoder är i en konstant förändring och utveckling
  - Komplicerade IT-infrastrukturer
- **Forskningsmål**:
  - Automatisera säkerhetsfunktioner
  - Anpassa system till föränderliga attackmetoder



Automatiserad Säkerhet: Nuvarande Forskningslandskap



Nivåer av säkerhetsautomatisering

Ingen automatisering.
Manuell detektering.
Manuell prevention.
Inga alarm.
Ingen automatiserad attack mitigering.
Brist på verktyg.

80-talet

Operatörassistans.
Manuell detektering.
Manuell prevention.
Granskningsloggar.
Säkerhetsverktyg.

90-talet

Partiell automatisering.
System har automatiserade funktioner för detektering/ prevention men kräver manuell uppdatering och konfiguration.
Intrångsdetekteringssystem.
Intrångspreventeringssystem.

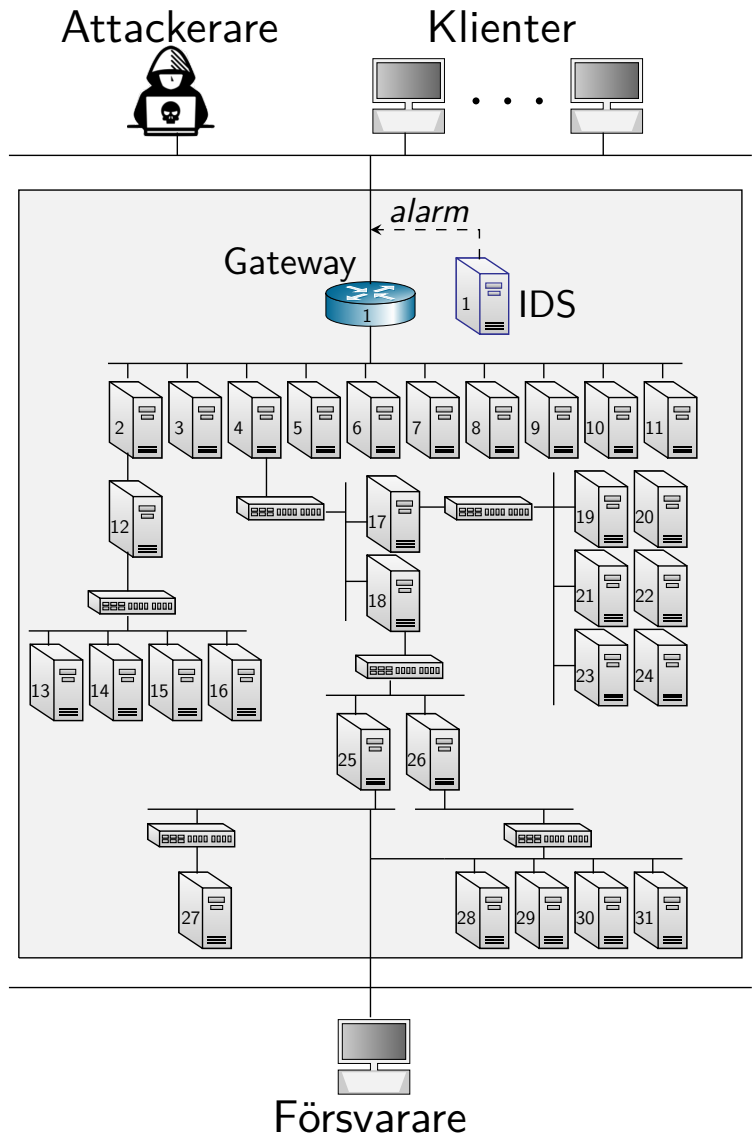
00-talet-Nu

Hög automatisering.
Systemet uppdaterar sig självt automatiskt.
Automatiserad attackdetektering.
Automatiserad attackmitigering.

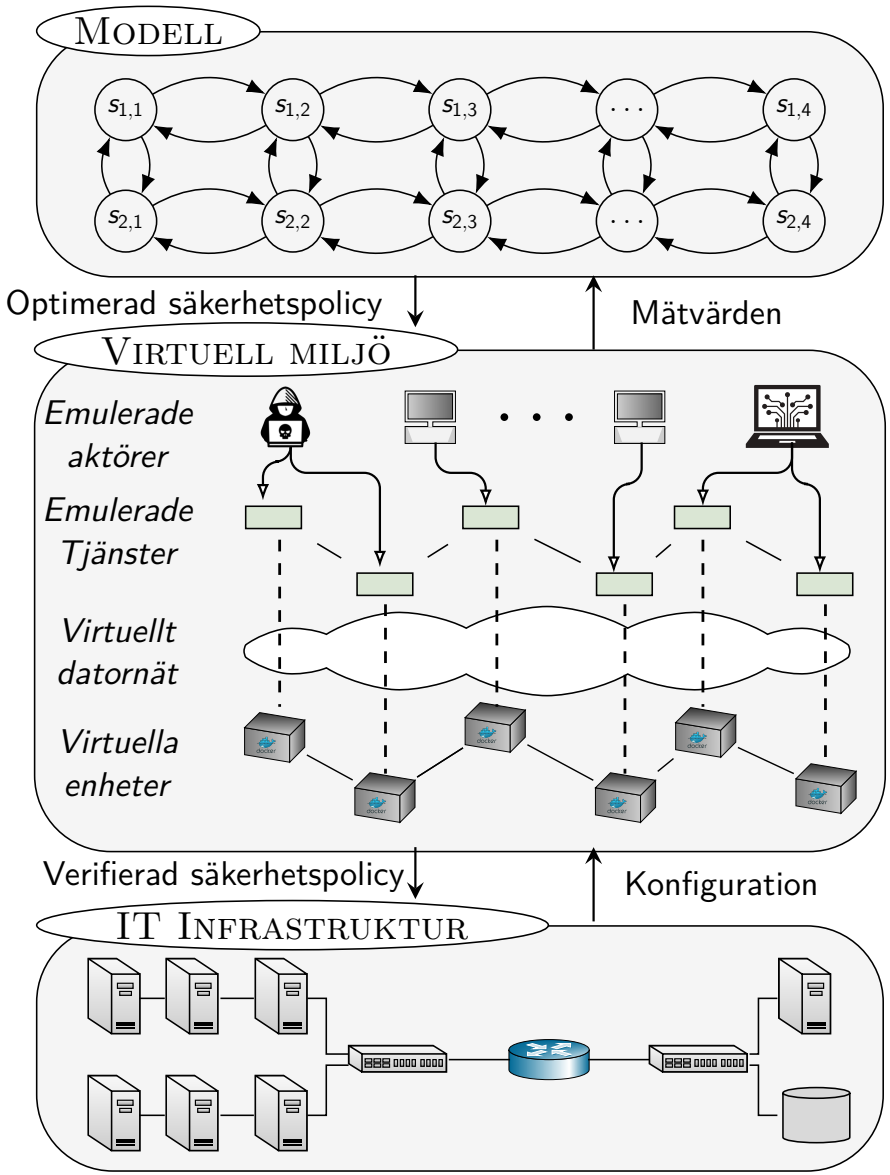
Forskning

Exempel: Intrångsmitigering

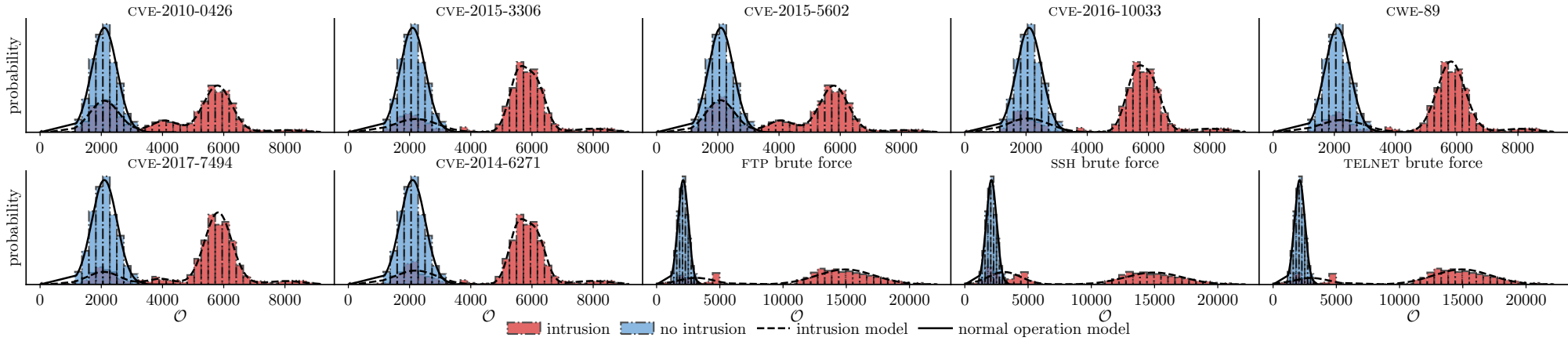
- En **försvarare** opererar en infrastruktur
  - Består av komponenter sammankopplade i ett nät
  - Komponenterna exekverar nätverkstjänster
  - Försvararen övervakar nätet och kan utföra responsaktioner
- En **attackerare** har som mål att göra ett intrång
  - Vill få tillgång till komponenter
  - Utför rekognisering samt exploatering av sårbarheter



Steg 1: Emulering



Steg 2: Samla in data

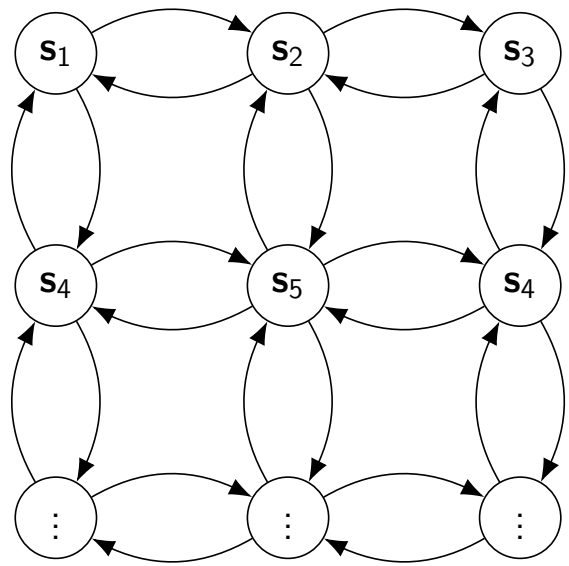


Fördelningar av intrångsdetekteringsalarm under olika typer av intrång.

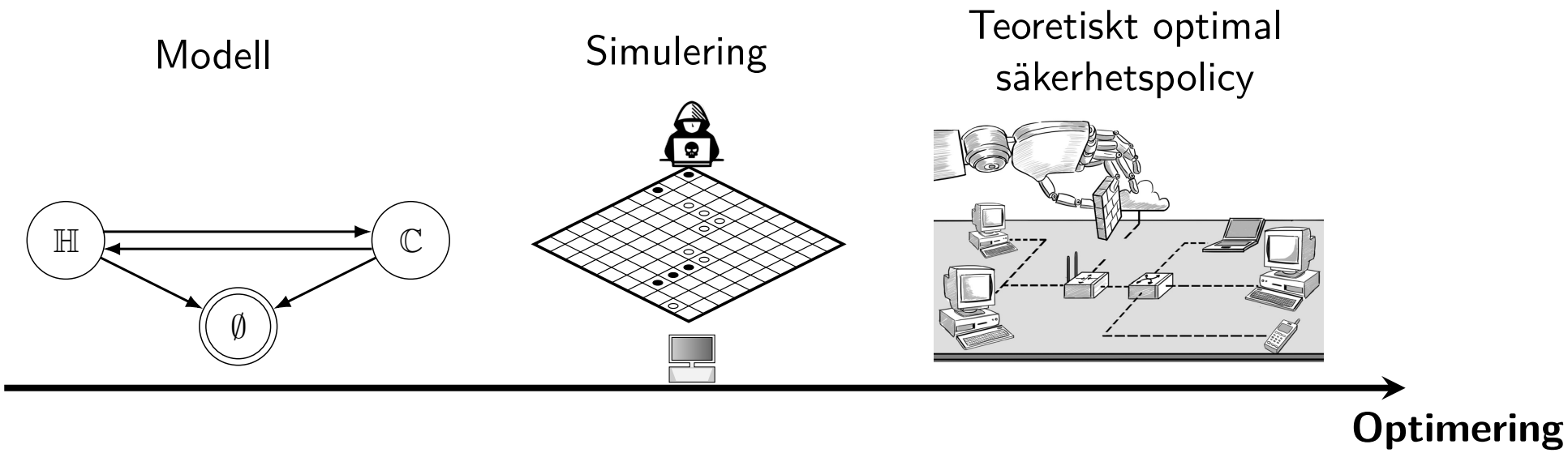
- Första steget i vår metod är att samla in stora mängder data från IT infrastrukturen.
- Vi samlar in data både under normaltillstånd samt under olika typer av intrång (emulerade intrång).

Steg 3: Modellering

- Vi modellerar systemet med hjälp av de mätvärden vi samlat in.
- Statistiska modeller baserat på
  - Hotmodellering
  - Reglerteknik
  - Beslutsteori
  - Spelteori
- Exempel på frågor vi kan svara med hjälp av modellen:
  - Vad är sannolikheten att ett intrång pågår?
  - Vilken effekt fås om vi uppdaterar vår säkerhetspolicy?

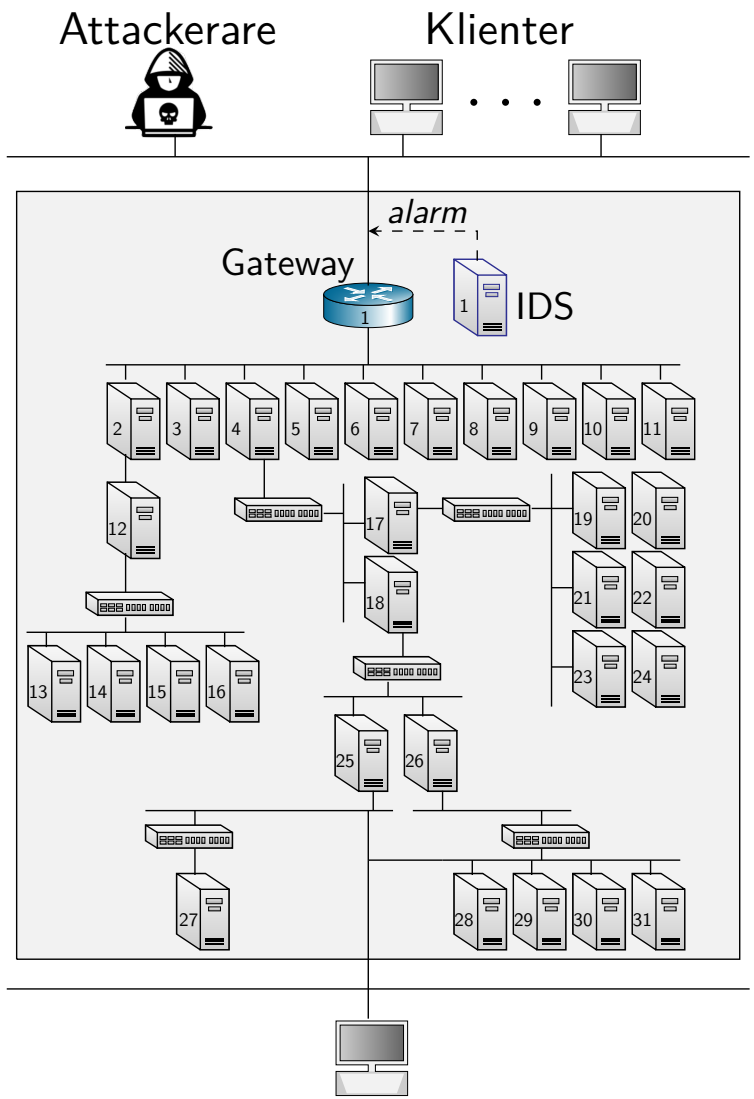


Steg 4: Optimering



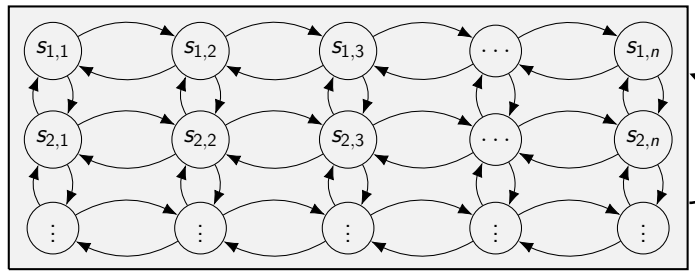
- Baserat på modellen och simuleringar så optimerar vi säkerhetspolicyn för systemet. Tekniker för optimering:
  - Artificiell intelligens
  - Förstärkningsinlärning
  - Dynamisk programmering
  - osv.

Steg 5: Implementering



Automatiserad säkerhetsstrategi

Vår metod för att automatiskt beräkna säkerhetsstrategier

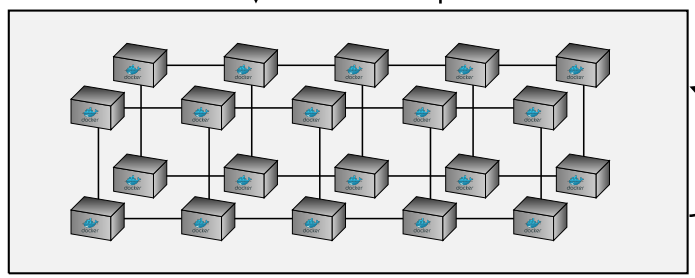


SIMULERINGSSYSTEM

Förstärkningsinlärning & optimering

Mappning av säkerhetsstrategi π

Modellskattning & Systemidentifiering

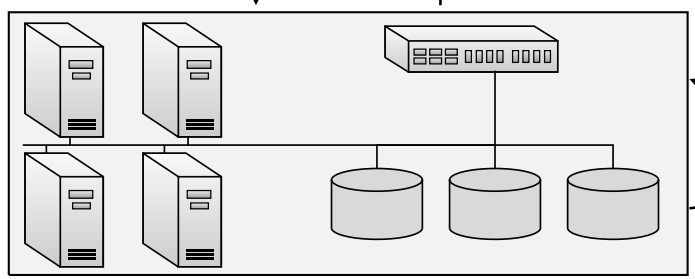


EMULERINGSSYSTEM

Utvärdering & skattning av modeller

Implementation av säkerhetsstrategi π

Selektiv replikering



IT INFRASTRUKTUR

Automatiserad säkerhet & självlärande system

Referenser

- Referenser till artiklar och videos finns tillgängligt på:

https://www.kth.se/cdis