以人工智慧與虛擬化技術驅動之物聯網安全評估與攻防演練

- 分別透過非技術背景主管與技術背景主管的角度，以資安風險分析流程為標的，
- 分享非監督型學習與虛擬化技術於物聯網裝置安全攻防演練中，可能扮演的角色與範疇

陳彥宏 pplong.ch@gmail.com

臺北護理健康大學資訊管理系 助理教授

2019-08-26
14:00-17:00

國立臺北護理健康大學
National Taipei University of Nursing and Health Sciences

網路動態分析實驗室
Net Telescope Lab

無形資產資本化-折舊年限來衡量軟體資安標準的問題

現在: 被攻破的損失只有倒閉，不是機率跟期望值，我們要的數據是新方案可以增加多少時間做[疏散資訊服務]跟[回復資訊服務]

核心價值應修改成

增加被攻破的時間，增加疏散服務的可用時間

讓專家能在被攻破的時間內來救你

1. 資訊系統「耐資安攻擊」標準是什麼？
2. 資安防災避難基準與參數是什麼？

2.1. 資訊系統「耐資安攻擊」標準是什麼？

內政部營建署《建築物耐震設計規範》，作為全國建築物耐震設計與興建的準則。明定新建的建築物，要能達到「小震不壞、中震可修、大震不倒」的耐震標準

小震不壞

根據歷史地震統計，當地平均每30年就會發生一次的最大地震，其強度不會使建築物受損，在地震過後能夠維持其正常機能。

中震可修

根據歷史地震統計，當地平均每475年就會發生一次的最大地震，其強度只會使建築物局部受損，但經過修繕後仍然可以居住。

大震不倒

根據歷史地震統計，當地平均每2500年就會發生一次的最大地震，其強度可能使建築物全面受損，但不會倒塌，大樓裡的人仍可逃離大樓。

- Annualized Loss Expectancy (ALE) and Annualized Rate of Occurrence (ARO)
- ALE = SLE x ARO
- Single Loss Expectancy (SLE)
- SLE = Asset Value × Exposure Factor (loss due to successful threat exploit, as a %)

2.2. 資安防災避難基準與參數是什麼？

- 北市府參考日本巨蛋防救災的模擬標準，歸納出以下7項公安基準。
  - 安全避難原則，以**8分鐘**內全員離開觀眾席至室內疏散空間，並須於**15分鐘內**達成全館人員避難至戶外避難空間為設計基準。

https://taipeicity.github.io/tpe_dome/security.html

- 內政部「消防機關火場指揮及搶救作業要點」
  - 於出動警鈴響起至消防人車離隊，白天八〇秒內，夜間一二〇秒內為原則

核心價值

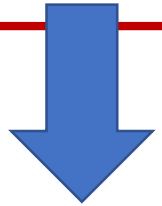
增加被攻破的時間，增加疏散服務的可用時間

讓專家能在被攻破的時間內來救你

核心價值

增加被攻破的時間，增加疏散服務的可用時間

讓專家能在被攻破的時間內來救你



【這個網路結構可以在資安團隊來救火前不被攻破】來衡量公司的資安品質，該怎麼做?

> 使用Fuzzy Test模擬可承受多久的攻擊?

=> Fuzzy Test的攻擊意圖是人工制定，會有倖存者偏差或其他政策考量

Fuzzers are often limited to simple errors because they won't handle business logic or attacks that require knowledge from the application user.

> 若用非監督型學習人工智慧的【學習時間】來量測相對強度的可行性?

e.g.

【舊的網路結構】需要5 hr的學習才能攻破

【新的網路結構】需要24 hr的學習才能攻破 => good

Risk analysis flowchart & what do we attempt to improve?

1. System Characterization
2. Threat Identification
3. Control Assessment
4. Vulnerability Identification
5. Likelihood Determination: Prob? Time?
6. Impact Analysis: High or Low
7. Risk Determination: Score?
8. Recommended Controls: Reduce it? Take it?
9. Results Documentation

Q: 花了這麼多錢，資訊系統比以前安全多少?

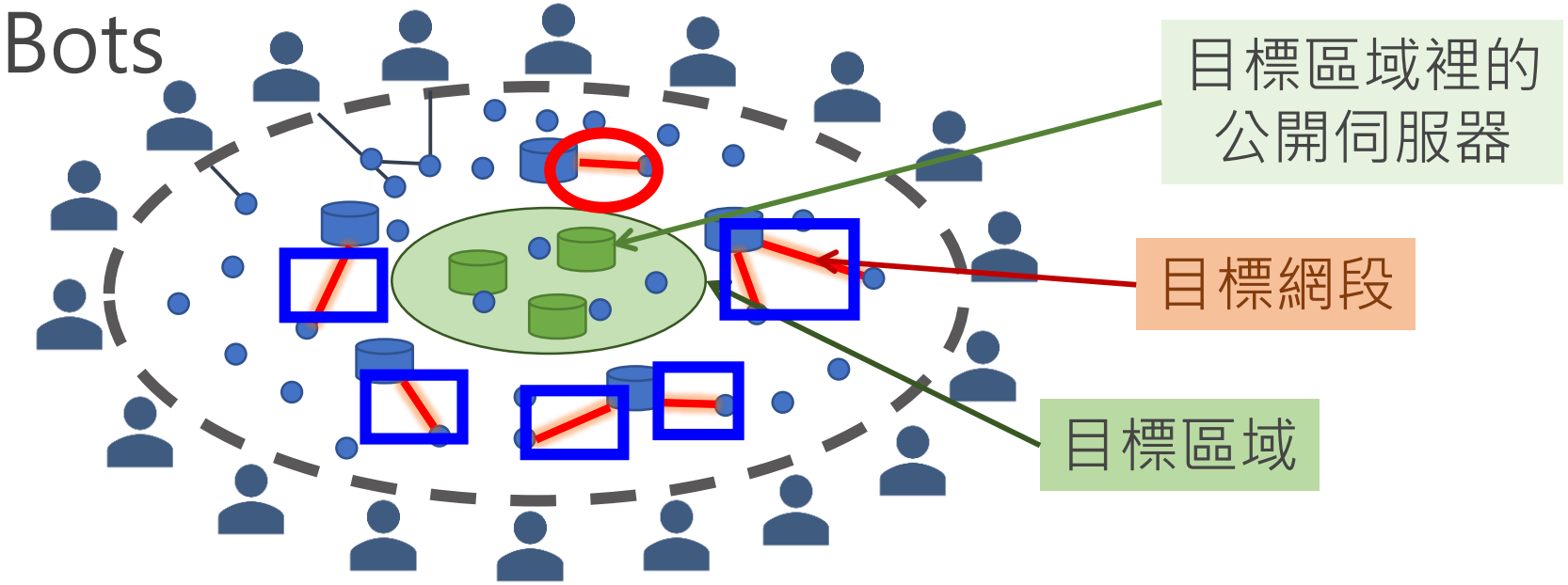
A:
(1) 小攻不壞、中攻可修、大攻不倒
(2) 增加被攻破的時間

Q: 衡量依據什麼？

A: AI(據統計特徵)的intentional attacking，以避免fuzzy test的人為介入

案例A

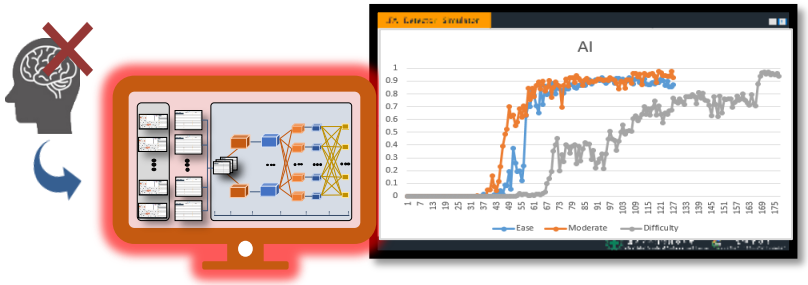
LFA以鏈路為攻擊目標，為防禦LFA攻擊
針對原來物聯網網路結構(Easy)，提出兩方案
(1)折衷方案 (Moderate: ⬭)，需要預算4M
(2)全面改善方案 (Hard: ▭)，需要預算10M



BEFORE: Prob.

- Single Loss Expectancy (SLE)
  SLE = Asset Value (in $) × Exposure Factor (loss due to successful threat exploit, as a %)
- Annualized Loss Expectancy (ALE)
- Annualized Rate of Occurrence (ARO)
  ALE = SLE x ARO

AFTER: Time

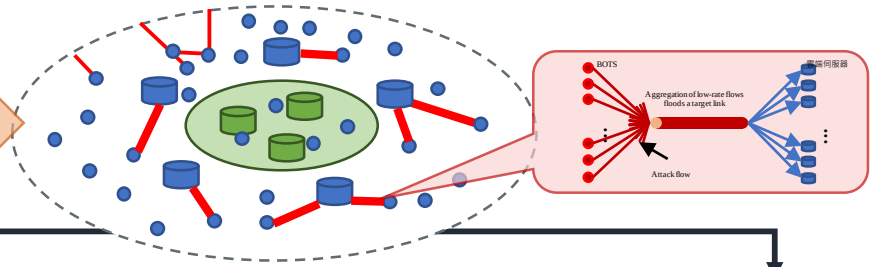


可否全自動化防禦與攻擊以減少人工成本?

BEFORE

以往是以人工方式來檢測LFA攻擊並設立規則，花費相當多的時間與人力成本

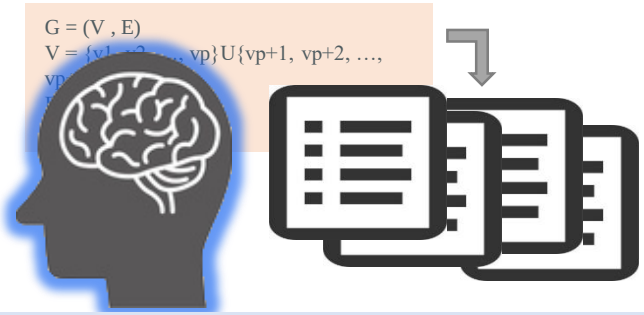
網路遭受攻擊後



人工觀察流量變化

流量變化利用人工觀察，無法即時分析複雜的流量特徵與網路結構，也需要花費相當多的時間與人力成本

人工設立攻擊規則

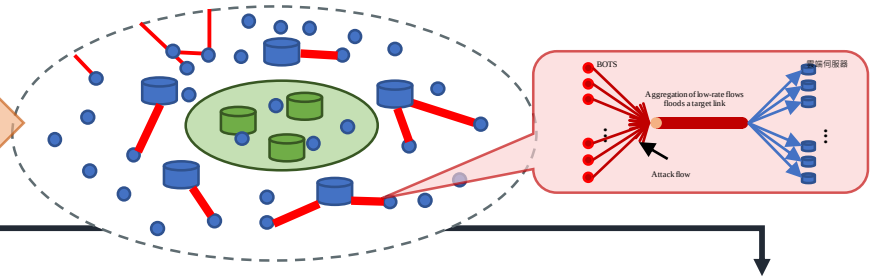


每當遭受攻擊，檢測人員就必須針對此一攻擊設立新規則，但是病毒型態、攻擊方式有千百種，如僅透過人力，則會耗費許多時間

AFTER

可否讓AI學習LFA攻擊特徵，**自動建立LFA檢測規則**，節省大量人力與時間成本

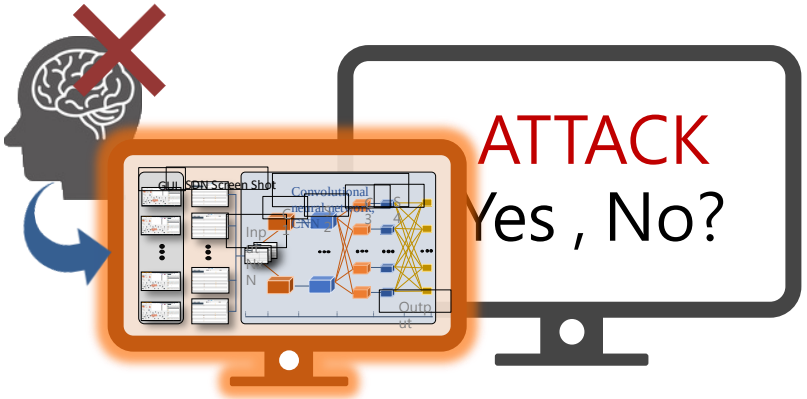
網路遭受攻擊後



不再使用人工，利用AI觀察流量變化

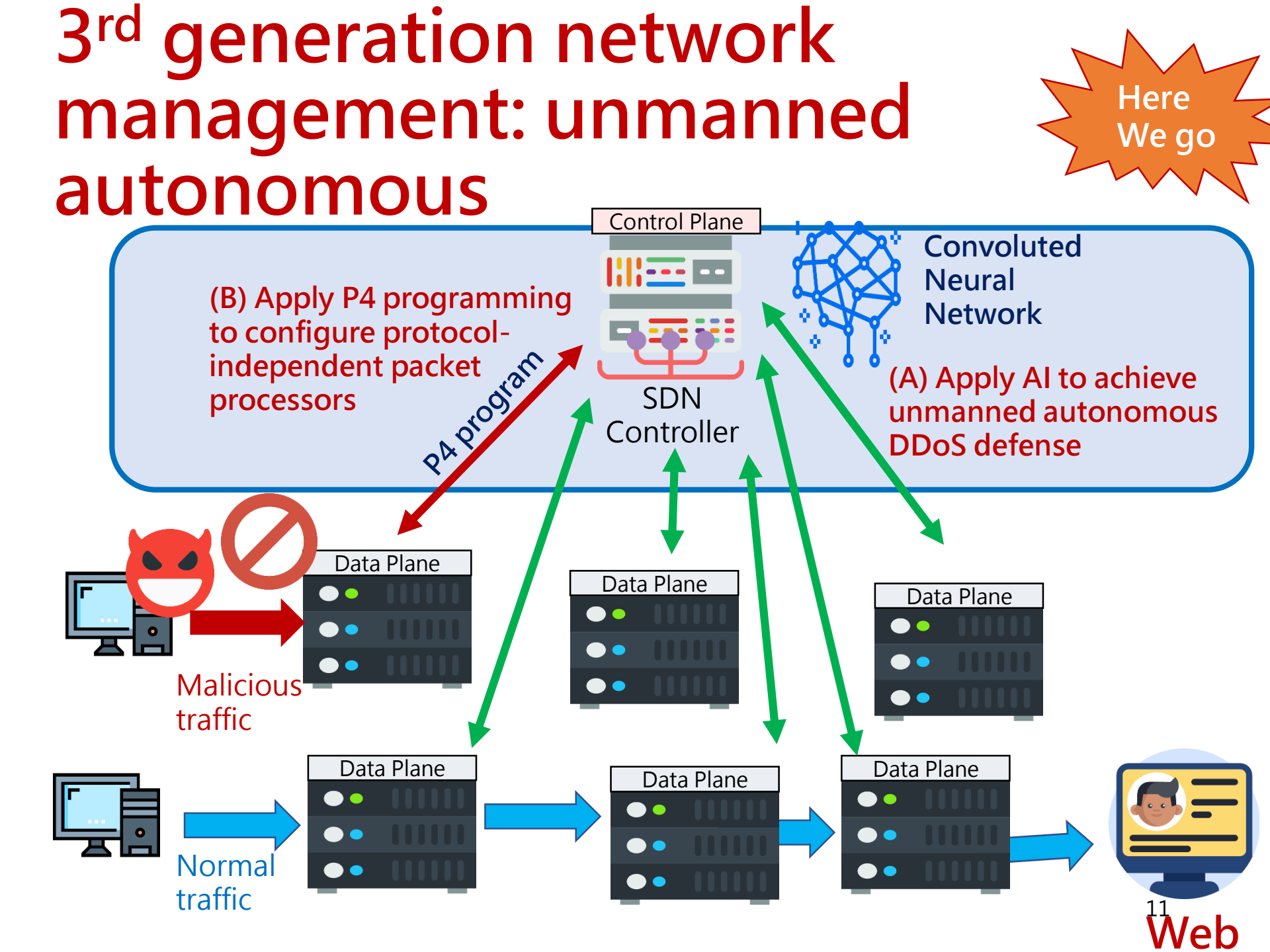
利用AI來自動觀測流量變化，AI檢測流量的同時也產生GUI介面以便檢測人員進行觀察與決策

不再使用人工，利用AI偵測攻擊



利用卷積神經網路訓練AI，讓AI能成功判斷是否為攻擊的準確度接近百分百，這樣不但能省下許多人力與時間出錯率也大幅降低

3rd generation network management: unmanned autonomous



Risk analysis flowchart & what do we attempt to improve?
(若以業者的角度來思考)

1. System Characterization
2. Threat Identification
3. Control Assessment
4. Vulnerability Identification
5. Likelihood Determination: Prob? Time?
6. Impact Analysis: High or Low
7. Risk Determination: Score?
8. Recommended Controls: Reduce it? Take it?
9. Results Documentation

Q2: 若業主願意承擔風險不做任何改善，總是要提供點建議吧?

a) 資安演練背後的問題?
b) 外包服務可以幫忙嗎?

BEFORE

往是自行設計的實驗室環境，建置成本太大，不適合於企業內部訓練環境

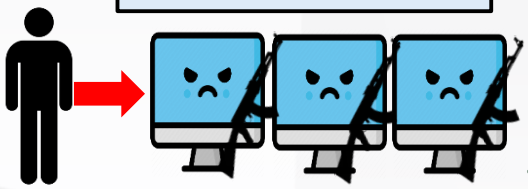
將自行設計的環境給資安人員進行攻擊時的防禦練習
藉由各式攻擊模擬提升其因應網路攻擊的防護知識

SCENARIO DESIGHER

① 需要多台電腦才可模擬出大規模的網路攻擊

Self-design Scenario

② 人工灌入電腦



ATTACK

JUDGE→HUMAN

DEFEND

TRAINERS

③

資安人員完成訓練後，成果由其他人以人工方式作檢測並評分

解決方法:

1. 用docker製作出類真實環境
2. 加入GUI，雙方零時差溝通

PROBLEM 1.
自行設計的環境往往與現實中有很多的不同且建置成本太高 ✗

PROBLEM 2.
訓練人員無法立即修正錯誤，造成訓練不同步的情況 ✗

AFTER

利用Docker建立類真實的攻擊環境，增強資安人員面對實際惡意攻擊時的應變能力

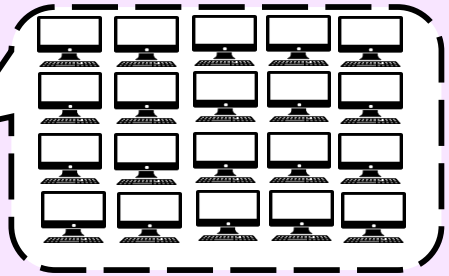
1.利用docker建立類真實環境
2.利用人工智慧設計評分標準(AI)、正確答案
3.加入GUI圖示

SCENARIO DESIGHER

預期效益:
1.縮短資安人員的訓練時間，變得更有效率。
2.訓練企業資安人員具備充裕的資安戰之應變與防護能力。

① docker

特點1: 設計虛擬化docker靶場環境，不需建立真的靶場



② 特點2.利用docker自動產生多台電腦

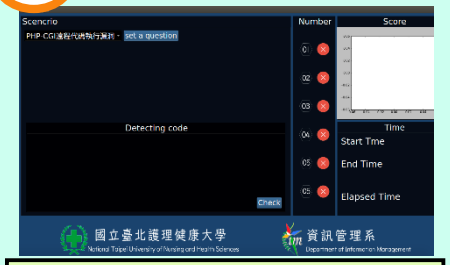
ATTACK

JUDGE→COMPUTER

DEFEND

TRAINERS

③



設計好GUI圖示
加入一套評分標準
給電腦檢測並給予評分機制，評估受訓人然的學習曲線，檢查是否合乎預期

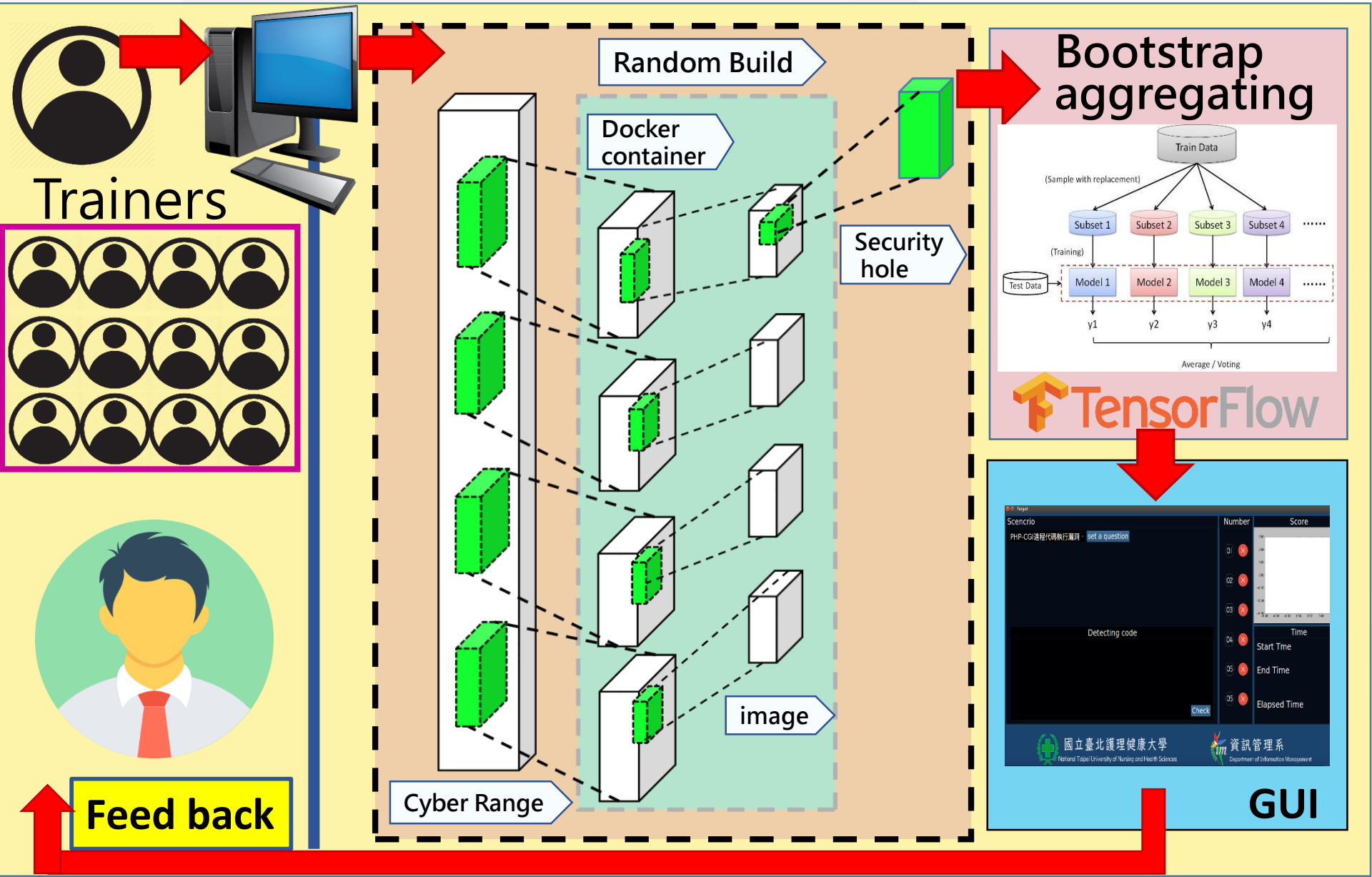
期望的學習標準曲線

特點3: 透過學習曲線，量化進度落後程度

目前學習曲線

SYSTEM MODEL

一台抵五十台，團進訓練、團出結業
以最小成本的大班制訓練靶場

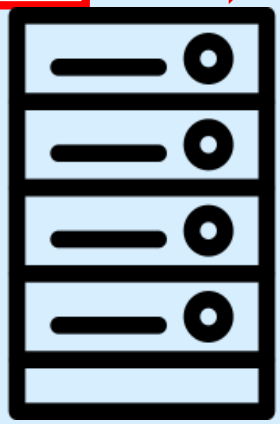


SCORING SYSTEM

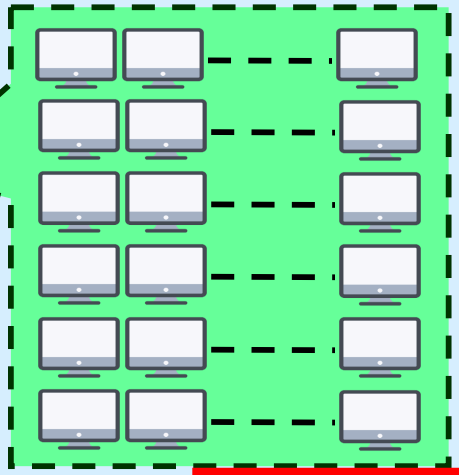
藉由電腦給予評分人員評分標準，
檢視員工資安演練成果
掌握員工資安訓練程度。

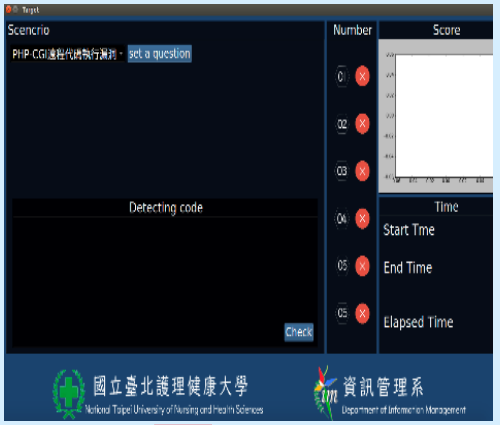
1.進入靶場訓練

Trainer



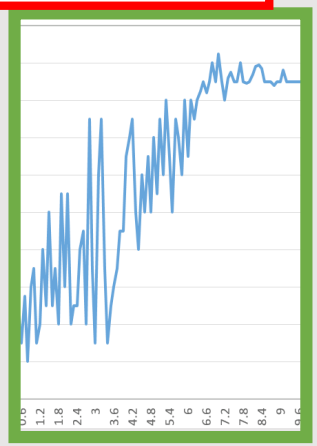
docker



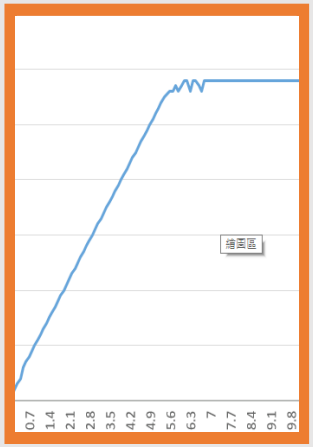


2.產生LOG

訓練後產生的資料



Trainer data

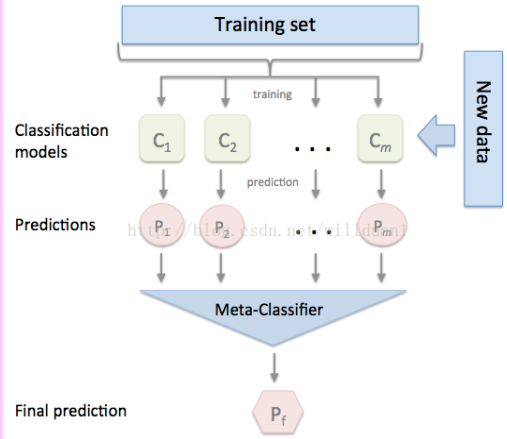


Stander data

Judge

P Value

檢測是否合乎預期



Bootstrap aggregating

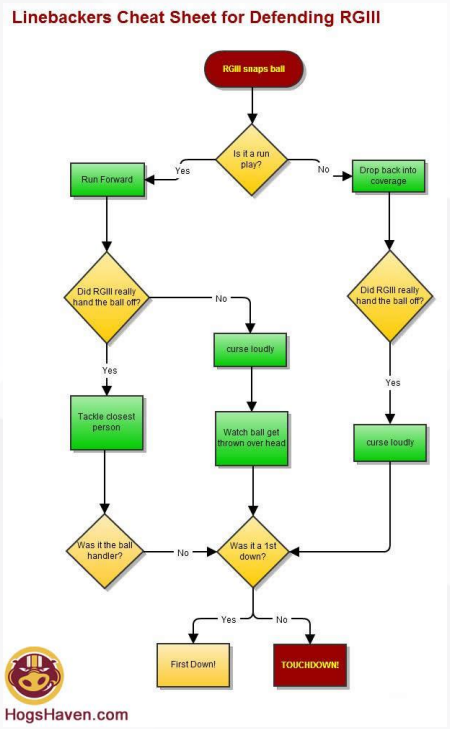
4.結業

3.產生訓練報表

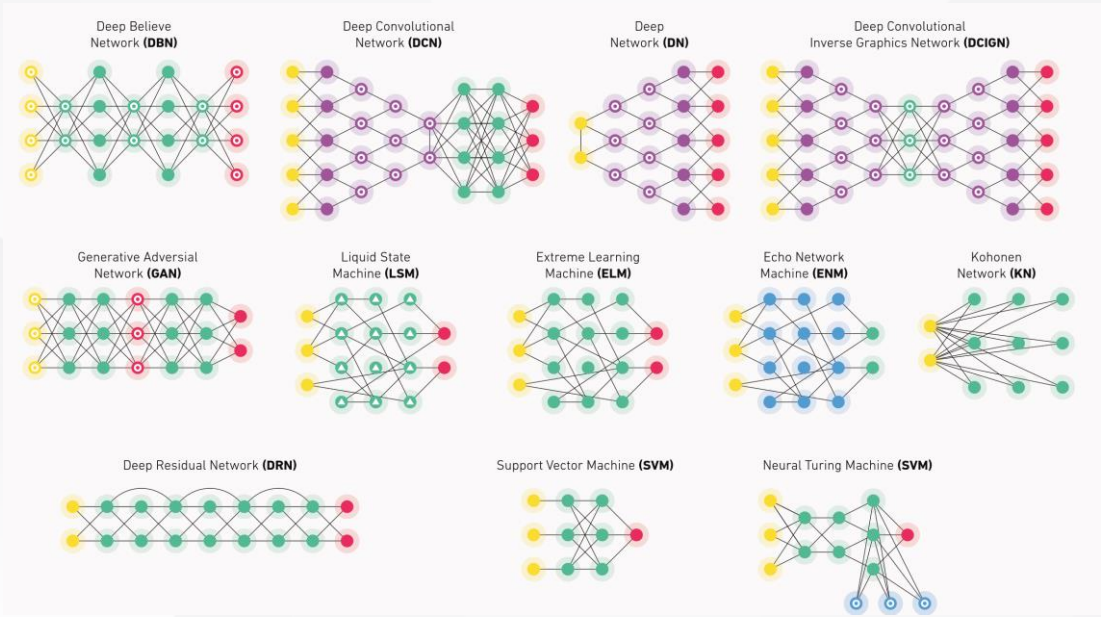
我們現在遇到的難關是甚麼?

1. 現在學生思維結構是左邊
但是政府與甲方需要右邊的人
(同時有甲方思維與乙方技術)

SOP流程腦
技術潔癖喜歡找漏洞



機率統計腦
成本化任何資安風險



2. 非監督型學習: 政府不易檢驗模型產生的過程
(1)每個程序否有效且合法?
(2)組織基於無效預測會做出違法行為?

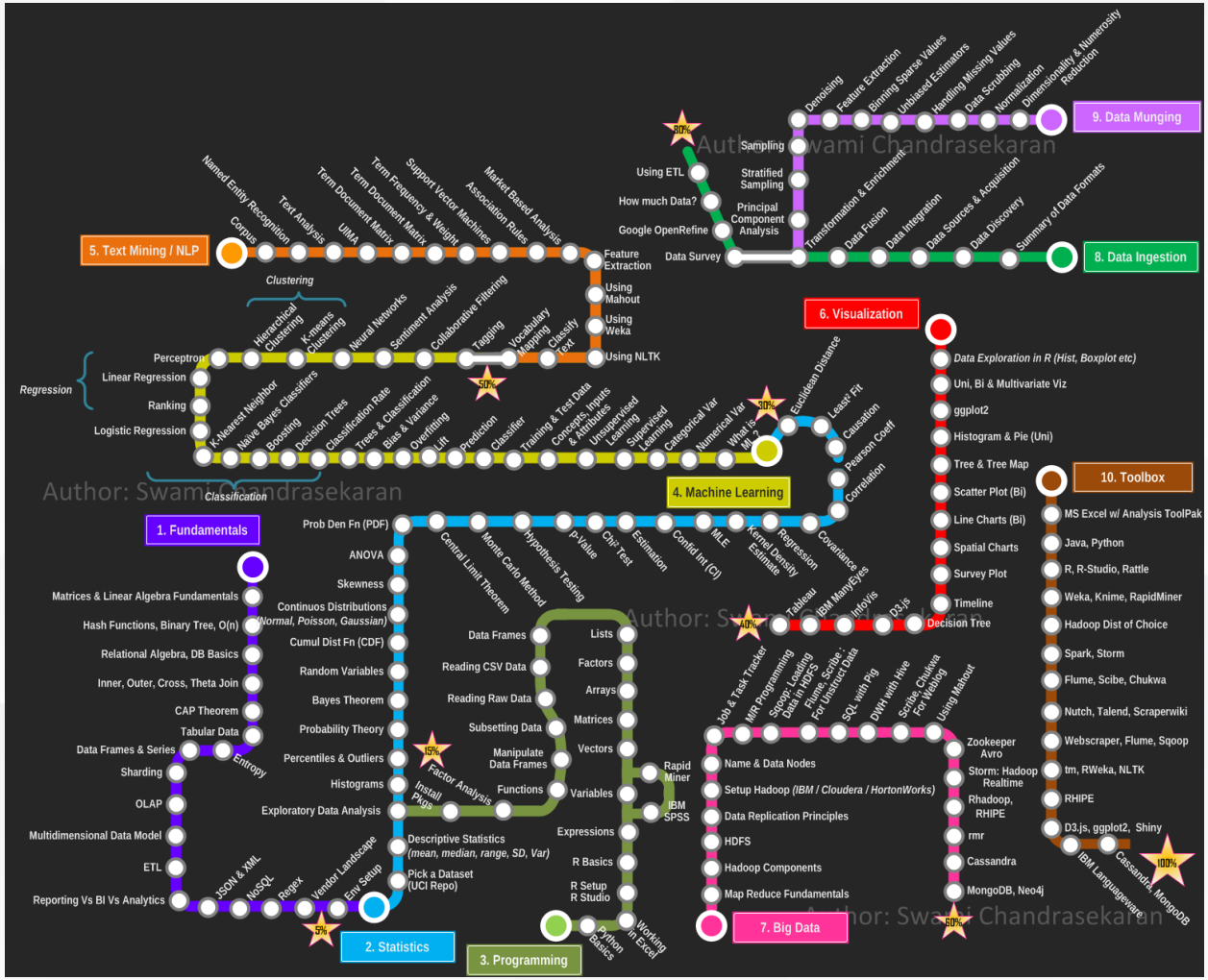
紐約時報2018/2/9 AI 應用人臉辨識，不同種族的準確率差異甚大。其中，黑人女性的錯誤率高達 21%~35%，而白人男性的錯誤率則低於 1%。

▲ 一組 385 張照片中，白人男性的辨識誤差最高只有 1%。（Source：Joy Buolamwini / M.I.T. Media

▲ 一組 271 張照片中，膚色較黑的女性辨識誤差率高達 35%。（Source：Joy Buolamwini / M.I.T. Media Lab）

https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

3. 採用監督型學習的學習門檻高，無法在現今大學教育制度中培育出所需人才 (教學評鑑/人本主義/每個小孩都不能放棄)



Thank you