國立臺北大學資訊工程學系113學年度專題

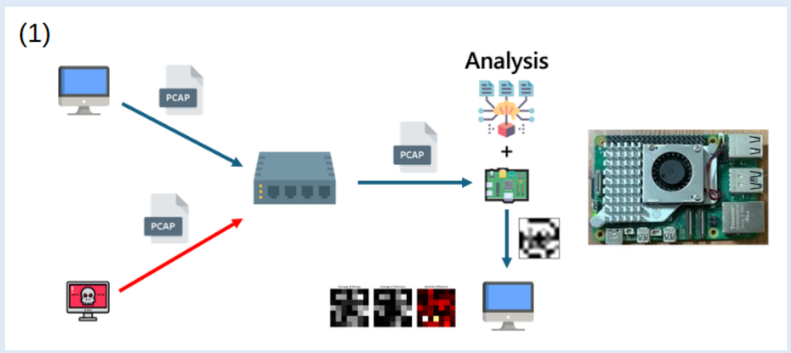
基於對比學習之入侵偵測系統

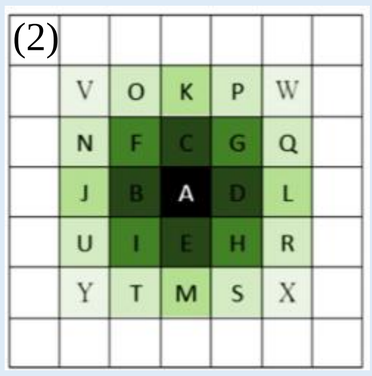
成員：歐陽平、陳孟舜、謝佳恩、林致諺

1. 動機

隨著網路攻擊頻率與複雜度日益增加，現有的深度學習入侵偵測系統多依賴高算力與大量標註資料，難部署於運算資源有限的裝置上。對此本研究運用自行開發的架構、資料處理方法、製圖方法、資料篩選方法。透過架構與資料優化的方式，使其能夠以低成本的方式運行在嵌入式裝置上。並透過即時的數據傳輸分析與模型偵測，將流量進行分類，以檢測訪問數據是否為安全流量。

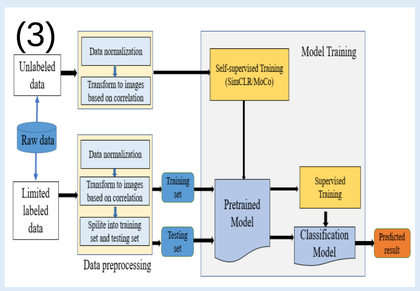
2. 系統流程介紹

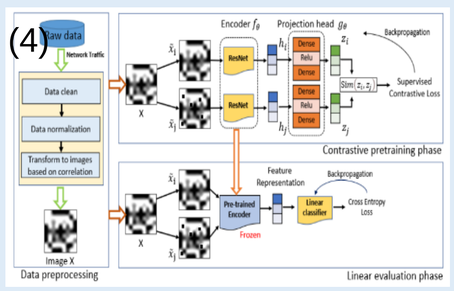


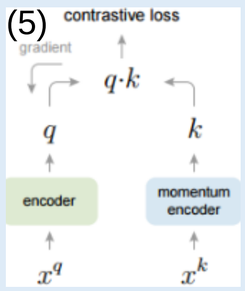


圖(1)示意整體系統流程：原始流量經過正規化與相關性轉換為影像後，先藉自監督對比學習（MoCo）預訓練出深層特徵模型，再以少量標註資料對分類微調，最終於嵌入式裝置上即時輸出入侵偵測結果。

圖(2)示意圖像如何生成，將PCAP流量特徵轉換為「圖像化資料」，先計算各欄位之Pearson相關係數，透過相關性排序生成易於CNN處理的二維圖像。







上圖展示模型架構，分為半監督訓練(圖3)與有監督訓練(圖4)。在半監督流程中，我們先將訓練集拆成「大量無標籤」與「少量有標籤」兩部分，利用 MoCo 結合 ResNet-18 作為encoder(圖5)進行自監督對比學習：同一張影像經兩種增強後視為正樣本，批次中其餘影像為負樣本，並以 Momentum encoder 穩定負樣本佇列，學得通用表徵後再用少量標籤進行 finetune，得到最終分類器；而在有監督流程中，資料不再拆分，而是直接以完整有標籤集進行「有監督對比學習」，將同類標籤視為正樣本、異類標籤視為負樣本，先學語意表徵後再次 finetune，產出決策模型。

3. 實驗結果

半監督訓練

Model	Dataset	Training set	Testing set	ACC
BYOL	UNSW-NB15	175341 Unlabeled: unknow Labeled: unknow	82332	89.97%
TCG-IDS	UNSW-NB15	1674243 Unlabeled: unknow Labeled: unknow	717532	91.48% (**BAC**)
Ours	UNSW-NB15	82332 Unlabeled:78332 Labeled:5000	175341	91.94%

監督訓練

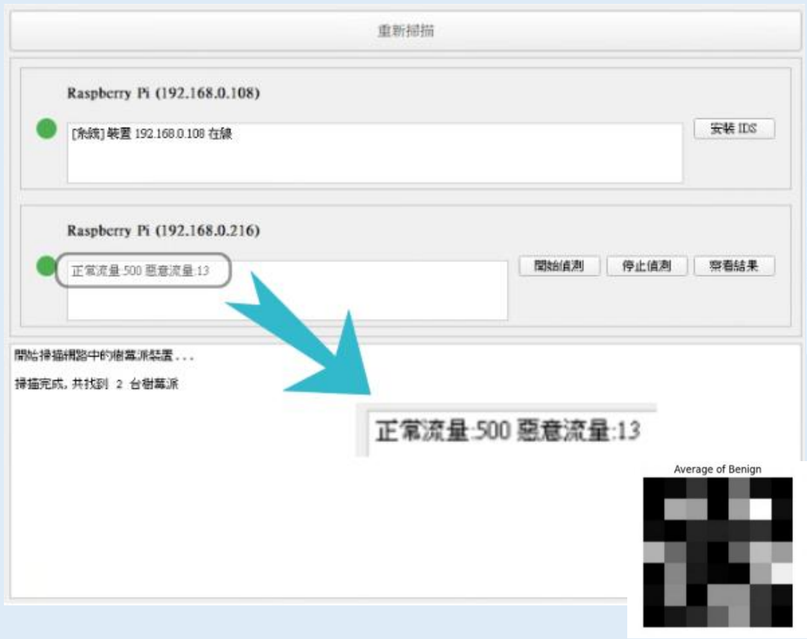
Model	Dataset	Training set	Testing set	ACC
Yue et al.	UNSW-NB15	82332	175341	92.99%
Lopez-Martin et al.	UNSW-NB15	2540044	82332	88.80%
Ours	UNSW-NB15	82332	175341	**93.18%**

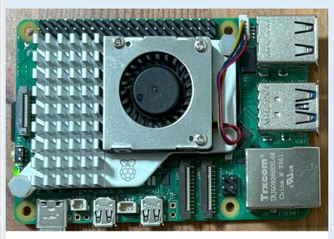
資料增強效果

Augmentation Techniques	ACC	Effectiveness
Feature Masking	**88.61%**	**High**
Random Horizontal Flip	**88.24%**	**High**
Gaussian Blur	**88.15%**	**High**
Random Vertical Flip	87.75%	Low
Color Jitter	87.07%	Low
Random Resized Crop	86.71%	Low

4. 偵測系統

系統實作部分我們使用半監督對比學習訓練完的決策模型進行實作，並使用Raspberry Pi 5b作為我們系統的載體。





5. 結論

我們提出「相關係數流量製圖＋ 對比學習」方法，證實對比學習在 IDS 可用；半監督與監督並行，標籤多寡皆能適應。流量圖像化後，正常與攻擊差異一眼可辨。未來擬結合聯邦學習，讓各裝置共享經驗，進一步提升偵測率與實用性。