自適應隱私保護的網路流量異常偵測框架

專題編號：114-2-CSIE-S002
執行期限：114年第1學期至114年第2學期
指導教授：陳昱圻
專題參與人員：112590011 林維信
112590044 邱翊豪
112590054 婁開翔

1、 摘要

本專題實作自適應隱私保護異常偵測系統（APPAD），針對傳統明文偵測易造成資料外洩，以及全同態加密偵測延遲過高之問題，提出以資料敏感度為依據的動態處理機制。系統先透過敏感度分類器判定資料欄位性質，並於自適應模組中對敏感資料採用CKKS同態加密、非敏感資料維持明文處理，形成混合資料流，最後以Logistic Regression模型進行異常偵測。本研究不進行不同模型比較，而是聚焦於APPAD系統中明文與密文推論結果之比對，以驗證加密處理對模型輸出的影響與可行性。實驗結果顯示，在維持資料隱私的前提下，系統能有效完成異常偵測，且密文推論結果與明文結果具有高度一致性，證明APPAD架構具備實務應用潛力。

關鍵詞：
APPAD、自適應隱私保護、同態加密、異常偵測、Logistic Regression、明文密文比對

2、 緣由與目的

隨著網路服務與資安威脅日益增加，異常偵測系統已成為識別攻擊行為的重要技術。然而傳統以明文資料進行分析的方式，雖具備較低延遲與較高準確率，卻存在敏感資訊外洩風險；相對地，採用同態加密之隱私保護方法雖可提升資料安全性，但計算成本高、延遲明顯增加，影響實務應用可行性。

本專題以此為背景，實作自適應隱私保護異常偵測系統（APPAD），透過敏感度分類機制判斷資料特徵，動態決定採用明文或加密方式進行處理，建立明文與密文並存之混合推論流程。本研究之目的在於完成APPAD系統之重現與整合，並透過明文與密文推論結果之比對，驗證同態加密應用於異常偵測之可行性與對模型輸出的影響，作為兼顧隱私與效能之解決方案基礎。

3、 研究報告內容

一、研究範圍

本研究以網路流量異常偵測為應用場景，針對具有隱私風險之資料進行保護與分析。研究範圍涵蓋資料前處理、敏感度判定、混合資料保護機制（明文與密文並存）、同態加密運算，以及基於Logistic Regression之異常偵測模型。資料來源以Web Authentication相關資料集為主，並模擬實際流量進行測試。期中階段主要聚焦於APPAD系統核心模組之建構與整合，不涉及與其他方法之效能比較，而以系統內部明文與密文推論結果之比對為主要驗證方向。

二、使用技術與方法

本研究採用以下技術進行系統實作：

1. 資料前處理：進行資料清洗（NaN/Inf處理）與標準化（Standard Scaler）。
2. 敏感度分類（Classifier）：依據特徵

語意建立規則，將資料標記為敏感或非敏感（flag ∈ {0,1}）。

3. 同態加密（CKKS）：使用TenSEAL實現向量加密與同態運算，支援加法與內積。

4. 自適應保護機制（Adaptive Module）：依敏感度結果決定資料以明文或密文方式處理。

5. 異常偵測模型：採用 Logistic Regression進行分類與機率預測。

異常偵測之核心計算為

Logistic Regression模型，其輸出為異常機率，表示如下：

$$P\left(y=1 \middle| \ x\right) = \frac{1}{1+e^{-(w^T x+b)}}$$

其中 x 為輸入特徵向量，w 為模型權重，b 為偏差項。

此外，本研究參考以下效能指標進行分析：

$$L_t = N \times p \times l$$

(1)

其中 L_t 為資訊洩漏量，N 為總流量數量，p 為敏感資料比例，l 為單筆資料大小。

三、系統架構與流程

本系統採用APPAD架構，主要流程如下：

1. 資料輸入與前處理
將原始流量資料進行清洗與正規化處理。

2. 敏感度判定（Classifier）
根據特徵語意（如IP、登入狀態等）判定是否為敏感資料。

3. 自適應資料分流（Adaptive Module）
   a. 敏感資料 → 同態加密（CKKS）
   b. 非敏感資料 → 保持明文

4. 模型推論（Server端）
對混合資料進行Logistic Regression推論。

5. 結果回傳與解密（Client端）
若為密文結果則進行解密，並計算最終異常判斷。

四、工具說明

本研究使用以下工具與環境進行開發：

1. Python：主要開發語言
2. TenSEAL：實作CKKS同態加密
3. Scikit-learn：Logistic Regression模型訓練
4. NumPy / Pandas：資料處理與運算
5. Figma：結果視覺化（後續）

五、實驗結果

目前已完成以下實驗與驗證：

1. 明文模型訓練：完成Logistic Regression模型訓練，並達到穩定準確率（約0.88以上）。
2. 資料前處理流程：建立可重現之標準化機制。
3. 敏感度分類機制：完成初步規則設計，能區分敏感與非敏感特徵。
4. 混合資料流（Mixed Protection）：成功實現明文與密文並存之資料處理流程。
5. 同態加密驗證：完成CKKS加密、運算與解密測試，誤差在可接受範圍內。

已完成系統核心功能:

1. 已成功完成明文與密文混合資料流（Mixed Protection）之實作。
2. 已完成CKKS加密、密文推論、解密與sigmoid計算流程整合。
3. 明文與密文推論結果具有高度一致性，驗證APPAD架構可行性。
4. 已完成延遲（Latency）與資訊洩漏（Information Leakage）等指標分析。
5. 系統能依敏感資料比例動態切換處理模式，在隱私保護與運算效率間取得平衡。

六、結論

1. 本研究已完成APPAD（Adaptive Privacy-Preserving Anomaly Detection）系統之核心架構與主要模組實作，成功建立結合敏感度分類、自適應路由、CKKS同態加密與Logistic Regression異常偵測之混合式隱私保護流程。

2. 研究結果顯示，系統可根據資料敏感度動態決定明文或密文處理方式，並完成端對端異常偵測流程。透過明文與密文推論結果之比對，可確認加密推論結果與明文結果具有高度一致性，同時兼顧資料隱私與系統可用性。

3. 此外，本研究亦完成延遲、資訊洩漏與模型效能等指標分析，驗證APPAD架構於隱私保護異常偵測應用中的可行性。未來可進一步優化同態加密運算效率、改善敏感度判定機制，並擴展至更多真實網路流量場景進行測試。

參考文獻

[1] Y.-r. Jeon, S.-h. Jee, S.-K. Kim, and I.-G. Lee, “Adaptive Privacy-Preserving Framework for Network Traffic Anomaly Detection,” Proc. MobiSec, 2025.

[2] A. Acar, H. Aksu, A. S. Uluagac, and M. Conti, “A Survey on Homomorphic Encryption Schemes,” ACM Computing Surveys, vol. 51, no. 4, 2018.

[3] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward Generating a New Intrusion Detection Dataset,” ICISSP, 2018.