國立台北大學資訊工程學系專題報告

基於對比學習的入侵偵測模型

專題組員: 曾慶哲、王致越、陳育昇、鄭宇謙

摘要

基於深度學習的入侵偵測模型已成為應對網路攻擊的重要手段。然而，通常需要大量標籤資料來訓練有效的模型，而這些資料往往難以取得，相較之下，無標籤資料卻十分豐富。自監督式學習特別適用於利用無標籤資料，其中的對比學習分支更符合入侵偵測模型的需求。本研究嘗試在探索對比學習在此領域中的應用，試圖僅使用部分標籤資料，並結合大量無標籤資料以提高分類準確率。

相關研究文獻探討

Momentum Contrast

Momentum Contrast（MoCo）的特色為其使用 momentum encoder，momentum encoder 使用了普遍的兩種方法，EndToEnd 跟 MemoryBank 的優點，並改善了其缺點。EndToEnd 無法支撐過大的 batch size，而 MemoryBank 雖能解決此問題，但卻會導致一致性不佳的問題。momentum encoder 不使用反向傳播而是以動量的方式更新。動量式的更新可以解決一致性的問題，也因為不須使用反向傳播，所以可用較大量的負樣本。

Barlow Twins

Barlow Twins 概念簡單易實現，不需使用負樣本訓練，適合小批量資料訓練。它在低數據狀態下的半監督分類效果優於當時其他方法，且在 ImageNet 分類及轉移任務中表現優異。

UNSW-NB15 資料集

資料集包含 2015 年 1 月 22 日和 2015 年 2 月 17 日的網路流量，共計 49 個特徵和 2,540,044 筆紀錄。第 48 個特徵標示 9 種網路攻擊行為類別，第 49 個特徵標示資料的正常與否（0 代表正常，1 代表攻擊）。此資料集的多樣性使其成為研究網路安全問題的理想實驗資料集。

目標

本研究利用 Momentum Contrast（MoCo）和 Barlow Twins 兩種對比學習模型來建立入侵偵測模型，期望通過對比學習從無標籤數據中提取有用特徵，減少對有標籤數據的依賴，從而降低數據標註的成本和時間，為網路安全領域提供一個降低入侵偵測系統構建和維護成本的方案。

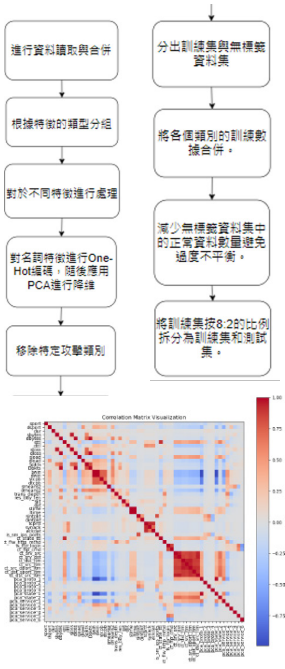
專題進行方式

資料前處理

1. 進行資料讀取與合併。
2. 根據特徵的類型 (名詞、整數、二進位和浮點數) 分組，以便於後續的資料轉換和處理。
3. 對於名詞特徵，去除前後空白；對於數字特徵，將無效值設置為 NaN，並進一步替換為 0。
4. 對名詞特徵進行 One-Hot 編碼，隨後應用 PCA 進行降維。
5. 因為特定的攻擊類別資料在原始資料集內的數量過少，故移除掉這些攻擊類別 ('Analysis'，'Backdoor', 'Shellcode'，'Worms')。
6. 對每個類別進行處理。抽取 10000 筆作為訓練集，並將剩餘的數據用作無標籤資料集。否則就將該類別的所有數據都用作訓練。
7. 將各個類別的訓練數據合併。
8. 減少無標籤資料集中的正常資料數量至 220000，避免過度的不平衡。
9. 將訓練集按 8:2 的比例拆分為訓練集和測試集。

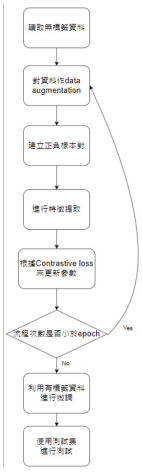
將資料轉換成圖片

由於 CNN 主要處理圖像數據，因此需將分類資料集轉換成圖片。



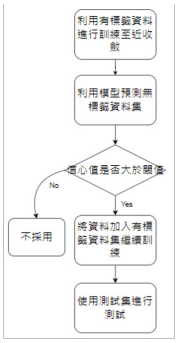
利用對比學習進行實驗

1. 讀取無標籤資料。
2. 使用 data augmentation 建立正負樣本對。方法包括旋轉、局部放大、變色、去雜訊、高斯模糊等，以提升模型的泛化性。
3. 使用 ResNet-18 作為模型主幹進行特徵提取，根據 Contrastive loss 更新參數。
4. 將全連接層改為輸出為 6 的線性層架構，進行微調以預測資料集內的 6 個類別。
5. 使用測試集進行測試。



利用半監督學習進行實驗對照

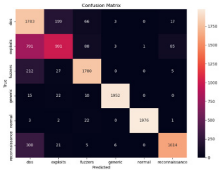
1. 利用有標籤資料進行訓練至近收斂
2. 利用模型預測無標籤資料集
3. 若信心值大於閾值，將資料加入有標籤資料集繼續訓練
4. 使用測試集進行測試



主要成果與評估

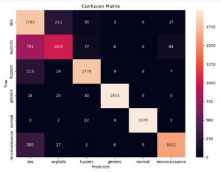
MoCo

Fold	Accuracy	Precision	F1
Fold1	84.28%	86.46%	84.38%
Fold2	83.89%	86.26%	84.03%
Fold3	83.53%	85.81%	83.63%
Fold4	84.51%	86.70%	84.68%
Fold5	83.95%	85.96%	84.15%



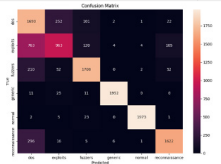
Barlow Twins

Fold	Accuracy	Precision	F1
Fold1	82.57%	84.20%	82.58%
Fold2	82.41%	84.61%	82.49%
Fold3	82.27%	84.23%	82.31%
Fold4	82.80%	84.79%	82.89%
Fold5	82.25%	84.06%	82.34%



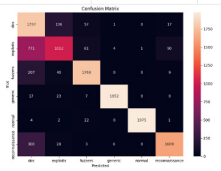
Semi-supervised CNN

Fold	Accuracy	Precision	F1
Fold1	84.13%	86.40%	84.22%
Fold2	83.88%	86.36%	84.07%
Fold3	83.51%	85.93%	83.59%
Fold4	84.39%	86.74%	84.59%
Fold5	84.02%	86.36%	84.21%



CNN

Fold	Accuracy	Precision	F1
Fold1	84.17%	86.42%	84.29%
Fold2	84.01%	86.57%	84.19%
Fold3	83.77%	86.32%	83.95%
Fold4	84.31%	86.74%	84.49%
Fold5	83.96%	86.47%	84.23%



結語與展望

在本研究中，我們探討了利用 MoCo 和 Barlow Twins 算法對網路流量進行預測，以區分正常流量和惡意攻擊。雖然對比學習在其他領域已顯示出優越性，但在網路流量入侵偵測中，其表現與半監督和 CNN 相比並未明顯優勢。未來的研究可尋找更合適的 transform 方法以提升結果。

參考文獻

[1]Barlow Twins: Self-Supervised Learning via Redundancy Reduction Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, Stéphane Deny arXiv:2103.03230
[2]Momentum Contrast for Unsupervised Visual Representation Learning Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, Ross Girshick arXiv:1911.05722
[3]Exploring Simple Siamese Representation Learning Xinlei Chen, Kaiming He arXiv:2011.10566
[4]A Simple Framework for Contrastive Learning of Visual Representations Ting Chen, Simon Kornblith, Mohammad Norouzi, Geoffrey Hinton arXiv:2002.05709
[5]How Does SimSiam Avoid Collapse Without Negative Samples? A Unified Understanding with Self-supervised Contrastive Learning Chaoning Zhang, Kang Zhang, Chenshuang Zhang, Trung X. Pham, Chang D. Yoo, In So Kweon arXiv:2203.16262
[6]https://research.unsw.edu.au/projects/unsw-nb15-data
[7]https://blog.csdn.net/qq_38308388/article/details/129862346