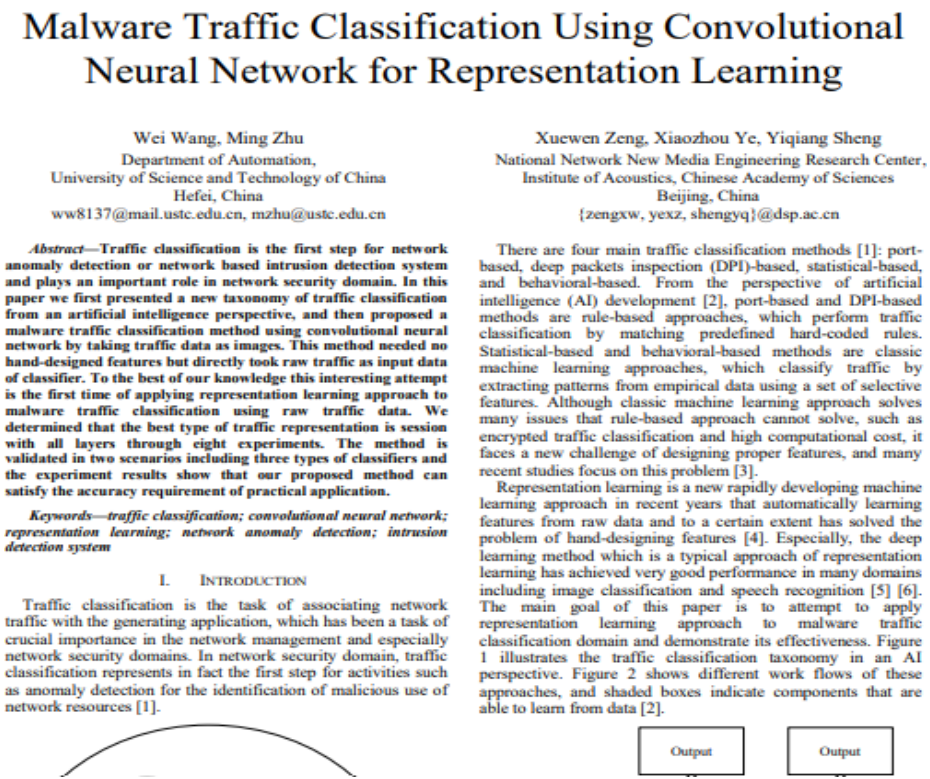
세션 슬라이싱 기법을 활용한 CNN 기반 악성 트래픽 분류 연구

지도교수 : 김명섭 교수님
2022270649 양지윤 · 2022271329 정고은

Index

01 프로젝트 소개



Malware Traffic Classification Using Convolutional Neural Network for Representation Learning

Wei Wang, Ming Zhu
Department of Automation,
University of Science and Technology of China
Hefei, China
ww8137@mail.ustc.edu.cn, mzhu@ustc.edu.cn

Xuewen Zeng, Xiaozhou Ye, Yiqiang Sheng
National Network New Media Engineering Research Center,
Institute of Acoustics, Chinese Academy of Sciences
Beijing, China
{zengxw, yexz, shengyq}@dsp.ac.cn

***Abstract*—Traffic classification is the first step for network anomaly detection or network based intrusion detection system and plays an important role in network security domain. In this paper we first presented a new taxonomy of traffic classification from an artificial intelligence perspective, and then proposed a malware traffic classification method using convolutional neural network by taking traffic data as images. This method needed no hand-designed features but directly took raw traffic as input data of classifier. To the best of our knowledge this interesting attempt is the first time of applying representation learning approach to malware traffic classification using raw traffic data. We determined that the best type of traffic representation is session with all layers through eight experiments. The method is validated in two scenarios including three types of classifiers and the experiment results show that our proposed method can satisfy the accuracy requirement of practical application.**

***Keywords*—traffic classification; convolutional neural network; representation learning; network anomaly detection; intrusion detection system**

I. INTRODUCTION

Traffic classification is the task of associating network traffic with the generating application, which has been a task of crucial importance in the network management and especially network security domains. In network security domain, traffic classification represents in fact the first step for activities such as anomaly detection for the identification of malicious use of network resources [1].

There are four main traffic classification methods [1]: port-based, deep packets inspection (DPI)-based, statistical-based, and behavioral-based. From the perspective of artificial intelligence (AI) development [2], port-based and DPI-based methods are rule-based approaches, which perform traffic classification by matching predefined hard-coded rules. Statistical-based and behavioral-based methods are classic machine learning approaches, which classify traffic by extracting patterns from empirical data using a set of selective features. Although classic machine learning approach solves many issues that rule-based approach cannot solve, such as encrypted traffic classification and high computational cost, it faces a new challenge of designing proper features, and many recent studies focus on this problem [3].

Representation learning is a new rapidly developing machine learning approach in recent years that automatically learning features from raw data and to a certain extent has solved the problem of hand-designing features [4]. Especially, the deep learning method which is a typical approach of representation learning has achieved very good performance in many domains including image classification and speech recognition [5] [6]. The main goal of this paper is to attempt to apply representation learning approach to malware traffic classification domain and demonstrate its effectiveness. Figure 1 illustrates the traffic classification taxonomy in an AI perspective. Figure 2 shows different work flows of these approaches, and shaded boxes indicate components that are able to learn from data [2].

Output Output

- **논문 재현**

「Malware Traffic Classification using Convolutional Neural Network for Representation Learning」논문 재현

- **기존 논문의 한계 보완**

세션 슬라이싱 기법으로 기존 방식의 한계를 보완하여 분류 성능 향상

02 과제수행 목적 및 필요성

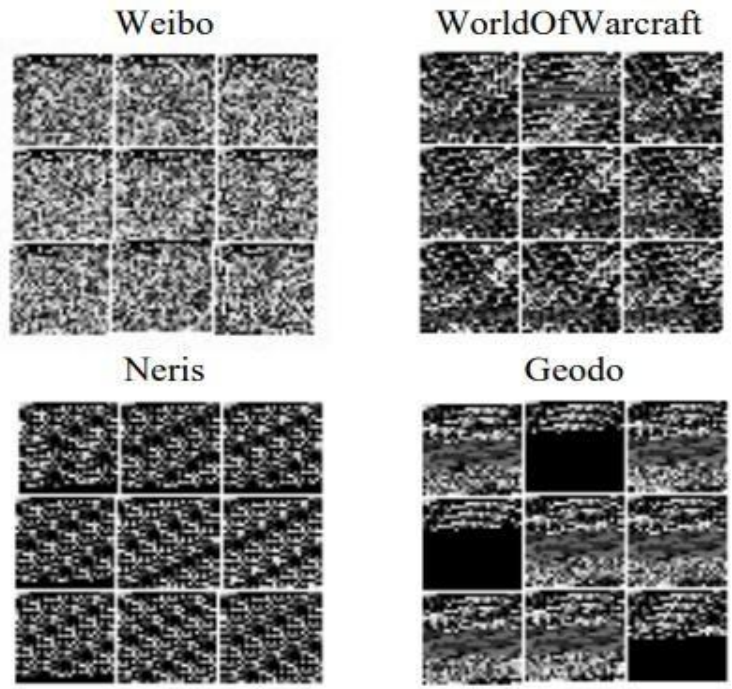


Fig. 5. Consistency in the same traffic class

[기존 논문 방식]

- 세션 앞부분 784바이트를 CNN 입력으로 사용
  - 짧은 세션에서는 패딩 발생 → 학습 왜곡
  - 세션 중·후반부 특징 반영 불가

[본 연구의 개선 방법]

- **세션 전체 구간에서 3구간 X 256B 추출**
  - Front : [0, 256)
  - Mid : [mid−128, mid+128]
  - Back : [L−256, L)

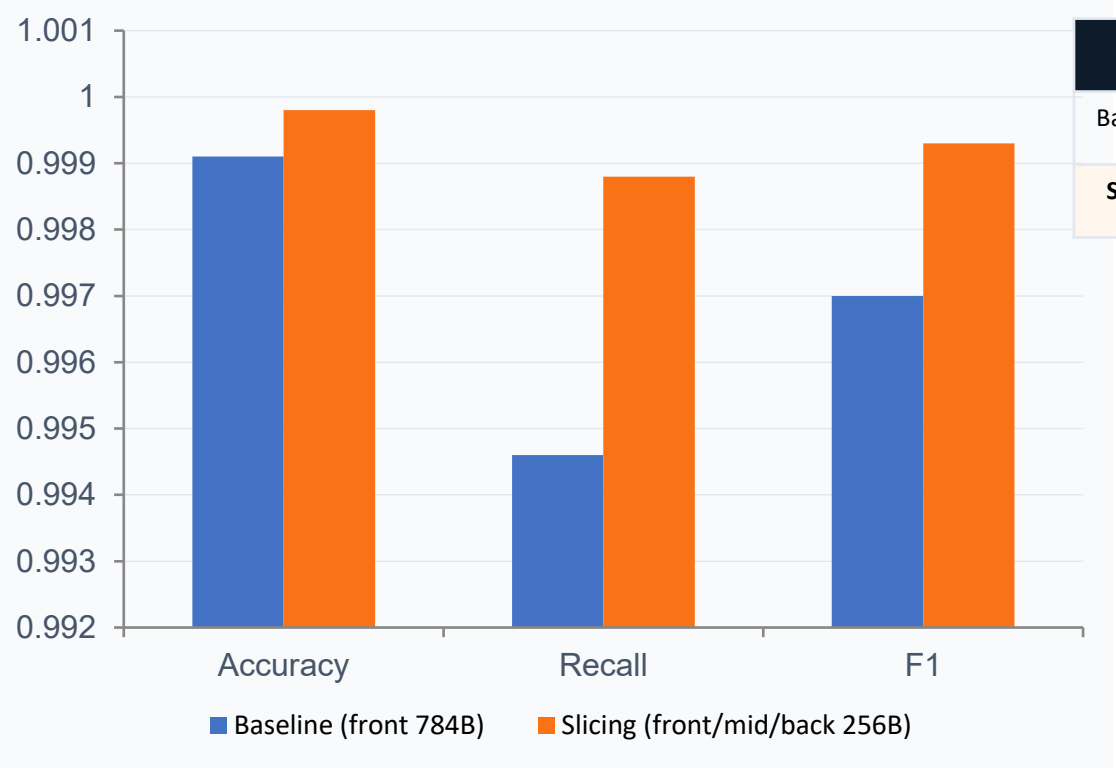
→ 추출된 3구간을 RGB 3채널로 변환하여 CNN 입력

03 과제수행 방향

① 기존 논문 재현	② Payload 길이 재정의	③ 세션 슬라이싱 적용	④ RGB 3채널 입력 구성	⑤ CNN 학습 및 성능 검증
USTC-TFC2016 데이터셋, 논문의 CNN 모델 구현	TCP 3-way handshake & ACK-only 패킷 제외	앞·중·뒤 각각 256B 추출	3구간을 채널로 쌓아 28×28×3 이미지 생성	Accuracy, Precision, Recall, F1-score 비교

입력 구성 비교

04 실험 결과 — 성능 비교



모델	Accuracy	Precision	Recall	F1
Baseline	0.9991	0.9995	0.9946	0.9970
Slicing	**0.9998 ↑**	**0.9999 ↑**	**0.9988 ↑**	**0.9993 ↑**

Malware Miss

31 → 7

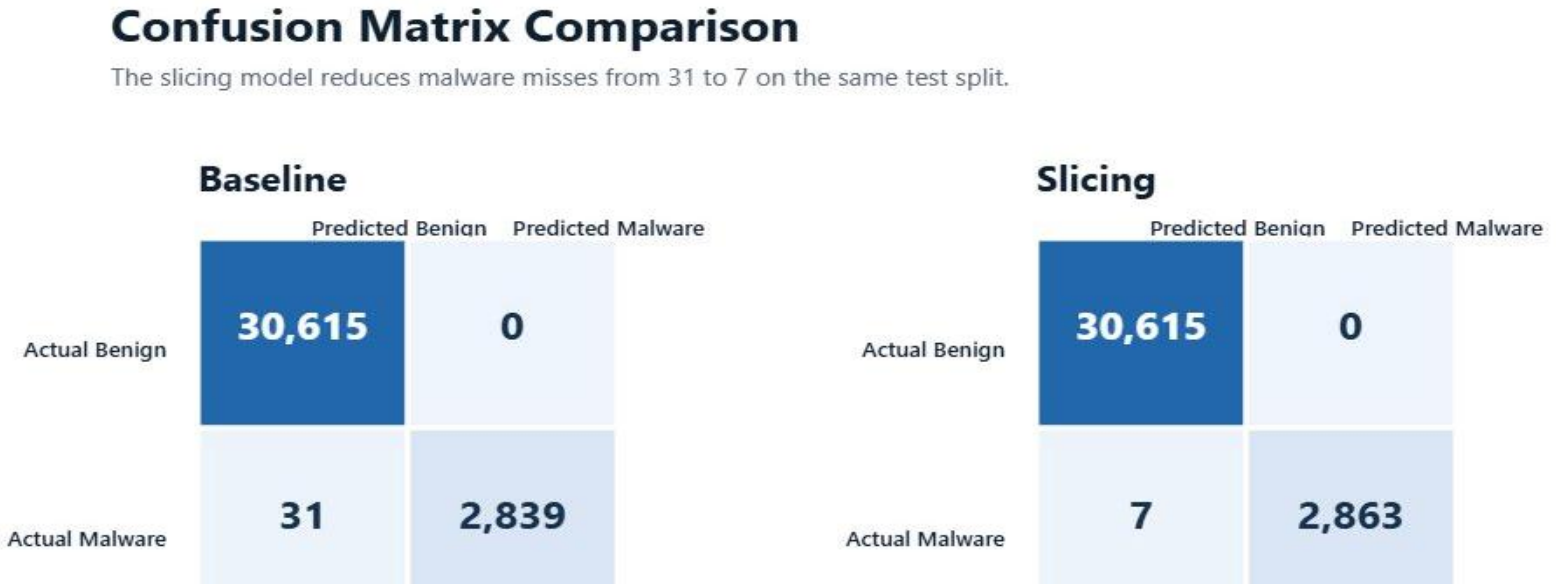
−77.4%

Recall 향상

+0.42%p

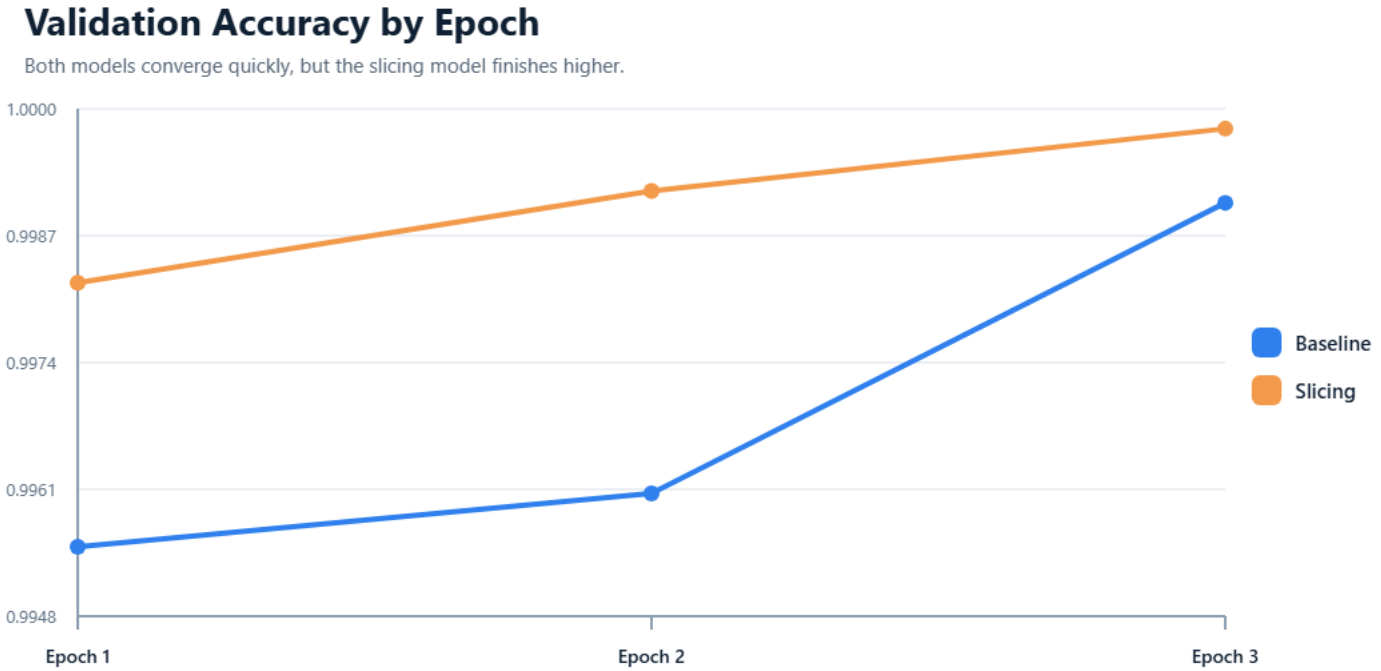
가장 큰 개선

04 실험 결과 — 혼동행렬 비교



Slicing 모델: 악성 트래픽 미탐지(False Negative) 31건 -> 7건 감소 (-77.4%) | 정상 트래픽 오탐(False Positive) 0건 동일 유지

04 실험 결과 — Epoch별 검증 정확도



Slicing 모델은 Epoch 1부터 높은 검증 정확도로 시작하여 수렴 속도도 우수 | 두 모델 모두 빠르게 수렴하나 Slicing이 최종 정확도에서 우위

05 결론 및 기대효과

01

연구 목표 달성

기존 논문 재현 후 세션 슬라이싱 기법을 성공적으로 적용하여 모든 지표에서 성능 향상 확인

02

핵심 성과

악성 트래픽 미탐지 31건 -> 7건 (-77.4%)

Recall +0.42%p, F1 +0.23%p 향상

03

향후 적용 가능성

입력 크기 단순화로 연산 효율 향상, 실시간 악성 트래픽 탐지 시스템 적용 검토

감사합니다.

캡스톤디자인 II | 지도교수 : 김명섭 교수님 | 2022270649 양지윤 · 2022271329 정고은