패킷 번들 기반 네트워크 행위분석을 위한 딥러닝 모델 비교 연구

김지민, 장윤성, 남승우, 유경민, 김명섭*

고려대학교

{illiard1209, brave1094, nam131119, rudals2710, tmskim* }@korea.ac.kr

A Comparative Study of Deep Learning Models for Packet Bundle-Based Network Behavior Analysis

Ji-Min Kim, Yoon-Seong Jang, Seung-Woo Nam, Gyeong-Min Yu, Myung-Sup Kim*

Korea Univ.

요 약

네트워크 트래픽 행위 분석은 암호화 트래픽의 증가와 고도화된 네트워크 서비스 환경으로 인해 점점 더 중요한 연구 주제로 부각되고 있다. 기존의 포트 기반 또는 DPI 기반 분석 기법은 암호화 환경에서 한계를 가지며, 이를 보완하기 위해 통계적 특징과 머신러닝, 나아가 딥러닝 기반의 행위 분석 기법이 활발히 연구되고 있다. 본 연구에서는 네트워크 트래픽을 패킷 단위가 아닌 패킷 번들(Packet Bundle) 단위로 표현하는 입력 표현 방식을 기반으로, 서로 다른 딥러닝 모델들이 해당 표현을 어떻게 활용하는지를 비교 분석한다. WeChat 트래픽 데이터셋을 대상으로 1D CNN, LSTM, Transformer 모델을 적용하여 이진 분류 및 5-class 네트워크 행위 분류 실험을 수행하였다. 실험 결과, 이진 분류에서는 모든 모델이 높은 정확도를 보였으며, 다중 행위 분류에서는 순차적 시간 정보를 고려하는 LSTM 모델이 가장 우수한 성능을 기록하였다. 이를 통해 패킷 번들 기반 입력 표현이 다양한 딥러닝 모델에 효과적으로 활용될 수 있음을 실험적으로 검증하였다.

Ⅰ. 서 론

현대의 네트워크 환경은 고도화된 공격 기법과 함께 지속적으로 진화하고 있으며, 이에 따라 네트워크 트래픽 분석 및 보안 시스템의 중요성은 지속적으로 증가하고 있다. 특히 고속·대규모 네트워크 환경에서는 알려진 공격뿐만 아니라 이전에 관찰되지 않은 제로데이 공격까지 탐지할 수 있는 능력이 요구된다. 이러한 요구에 대응하기 위해 다양한 네트워크 트래픽 분류 및 분석 기법이 제안되어 왔으며, 각 방식은 고유한 장단점을 가진다.

초기 네트워크 트래픽 분류는 포트 번호나 패턴 매칭에 기반한 시그니처 기반 탐지 방식이 주를 이루었다.[1] 해당 방식은 빠른 탐지 속도와 낮은 오탐율이라는 장점이 있으나, 암호화 트래픽 환경이나 새로운 공격 유형에 취약하다는 한계를 가진다. 이를 보완하기 위해 통계적 특징을 기반으로 한 머신러닝 기법이 도입되었으며, 트래픽의 패킷 길이, 인터패킷 시간(IAT) 등의 특징을 활용하여 분류 성능을 향상시켰다.[2] 그러나 이러한 접근 방식은 수작업 특징 설계에 의존하며, 다양한 트래픽 환경에 대한 일반화 성능에 한계가 존재한다.

최근에는 CNN, RNN 등 딥러닝 기반 기법이 도입되며, 원시 트래픽 또는 변환된 트래픽 표현으로부터 자동으로 특징을 학습하는 연구가 활발히 진행되고 있다.[3] 특히 암호화된 트래픽 환경에서도 행위 기반 분석을 수행할 수 있다는 점에서 딥러닝 기반 접근 방식은 큰 장점을 가진다. 그러나 기존 연구의 상당수는 특정 모델(CNN)에 국한되거나, 단순 이진 분류 문제에 초점을 맞추고 있어 입력 표현 방식이 다양한 딥러닝 모델에서도 일반적으로 활용 가능한지에 대한 검증은 충분히 이루어지지 않았다.

본 연구에서는 이러한 문제의식을 바탕으로, 네트워크 트래픽을 패킷 번들(Packet Bundle) 단위로 표현하는 입력 방식을 활용하여 서로 다른 딥러닝 모델 간 성능을 비교 분석한다. 기존 연구에서 제안된 패킷 번들 기반 표현을 확장하여, 1D CNN뿐만 아니라 LSTM과 Transformer 모델에도 동일하게 적용하고, 이진 분류 및 다중 행위 분류 문제에서의 성능을 체계적으로 분석한다.

Ⅱ. 관련연구

암호화 트래픽 환경에서 네트워크 행위를 분석하기 위한 연구로, 트래픽의 분포적 특성과 시간적 구조를 활용한 접근 방식이 제안되어 왔다. Li et al.은 암호화된 인터넷 트래픽을 대상으로 sliding window 방식을 적용하고, 윈도우 좌측과 우측의 패킷 분포 차이를 KL-Divergence로 계산하여 트래픽을 분할하는 방법을 제안하였다.[4] 해당 연구는 원시 패킷 정보 대신 클래스 간 분포 차이를 강조하는 특징을 학습함으로써 암호화 환경에서도 행위 인식이 가능함을 보였다. 그러나 File Transfer와 같이 상대적으로 작은 용량의 트래픽이 발생하는 경우에는 행위 전환 시점이 명확하게 분리되지 않아, 정확한 행위 구간 분할이 어렵다는 한계가 존재한다.

한편, WeChat 트래픽을 대상으로 세션 및 패킷 번들(Packet Bundle) 단위의 행위 분석을 수행한 연구에서는,[5] 연속된 패킷을 일정 시간 단위로 묶어 하나의 번들로 표현하는 입력 방식을 제안하였다. 해당 연구는 패킷 번들 개념을 활용하여 세션 단위의 트래픽 행위를 효과적으로 표현할 수 있음을 보였으며, 2D CNN과 1D CNN 모델을 수직적으로 결합한 행위 분석 프레임워크를 제안하였다. 이를 통해 패킷 단위 분석 대비 향상된 분류 성능을 달성하였다.

그러나 기존 패킷 번들 기반 연구는 CNN 기반 단일 모델 구조에 초점을 두고 있으며, 제안된 입력 표현이 순차적 시간 정보를 명시적으로 활용하는 LSTM이나 self-attention 기반 Transformer 모델에서도 동일하게 효과적인지에 대한 검증은 충분히 이루어지지 않았다. 또한 실험 설정이 이진 분류와 같은 비교적 단순한 행위 구분 문제에 집중되어 있어, 보다 복잡한 다중 네트워크 행위 분류 문제로의 확장 가능성에 대한 분석이 필요하다.

Ⅲ. 본론

3.1. 데이터 입력 표현

본 연구에서는 네트워크 트래픽을 1초 단위 세션으로 분할하고, 이를 다시 10ms 간격의 패킷 번들로 구성한다. 하나의 입력 샘플은 총 100개의 패킷 번들로 이루어지며, 각 번들은 총 전송 바이트 수(Bytes), 평균 인터 패킷 시간(IAT), 패킷 개수(Packet Count)로 구성된 3차원 특징 벡터로 표현된다. 따라서 입력 데이터의 형태는 (B,100,3)이며, 여기서 B는 배치 크기를 의미한다.

3.2. 모델 구조

본 연구에서는 네트워크 트래픽을 1초 단위 세션으로 분할하고, 이를 다시 10ms 간격의 패킷 번들로 구성한다. 하나의 입력 샘플은 총 100개의 패킷 번들로 이루어지며, 각 번들은 총 전송 바이트 수(Bytes), 평균 인터 패킷 시간(IAT), 패킷 개수(Packet Count)로 구성된 3차원 특징 벡터로 표현된다. 따라서 입력 데이터의 형태는 (B,100,3)이며, 여기서 B는 배치 크기를 의미한다.

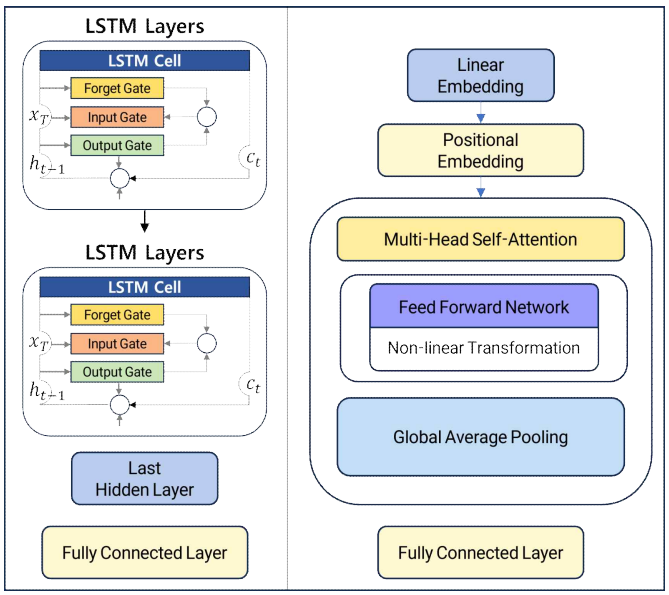

그림 1 행위분석 LSTM / Trasnformer 모델 구조도

3.3. 실험 결과

Chat과 File Transfer에 대한 이진 분류 실험에서는 1D CNN과 LSTM 모델이 약 98.68%의 정확도를 기록하였으며, Transformer 모델은 약 97.37%의 정확도를 보였다. 이는 패킷 번들 기반 입력 표현이 단순 이진 행위 구분에 효과적임을 의미한다.

Model	Accuracy (%)
1D CNN	98.68
LSTM	98.68
Transformer	97.37

표 1 모델 별 Chat/FileTransfer 이진 분류 성능

5-class 네트워크 행위 분류(Chat, File Transfer, Voice, Video, Others) 실험에서는 LSTM 모델이 가장 높은 성능을 기록하였다. LSTM은 Accuracy 92.42%, Macro-F1 72.20%를 기록하였으며, 이는 순차적 시간 정보를 효과적으로 활용한 결과로 해석된다. 반면 Transformer 모델은 상대적으로 낮은 성능을 보였으나, 전반적인 행위 분포를 반영하는 데에는 의미 있는 결과를 나타냈다.

Model	Accuracy	Precision	Recall	F1-score
1D CNN	91.62	76.81	67.84	68.73
LSTM	92.42	81.73	70.93	72.20
Transformer	90.97	69.41	66.49	64.43

표 2 모델 별 5-Class 분류 성능

Ⅳ. 결론

본 연구에서는 패킷 번들 기반 네트워크 트래픽 입력 표현을 다양한 딥러닝 모델에 적용하여 행위 분석 성능을 비교 분석하였다. 실험 결과, 제안된 입력 표현은 CNN, LSTM, Transformer 모델 모두에서 효과적으로 활용 가능함을 확인하였으며, 특히 다중 행위 분류 문제에서는 순차적 시간 정보를 고려하는 LSTM 모델이 가장 우수한 성능을 보였다.

향후 연구에서는 VPN 환경이나 더 복잡한 암호화 트래픽 환경에서의 적용 가능성을 검증하고, 바이트 레벨 정보 또는 attention 기반 pooling 기법을 결합한 입력 표현 확장을 통해 보다 정교한 행위 분석 모델을 설계할 예정이다.

ACKNOWLEDGMENT

본 논문은 과기정통부·정보통신기획평가원의 정보통신방송표준개발지원(R&D,정보화)사업(No. RS-2025-02219319, 양자컴퓨터 공격에도 안전한 양자암호 기반 제로트러스트 보안 네트워크/서비스 및 제어/관리 기술 표준개발)과 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0024177, 2023년 지역혁신클러스터 육성) (ZTA지역혁신 / 란아 : ETRI, Cross-Domain)

참 고 문 헌

[1] Guang Cheng and Song Wang, "Traffic classification based on port connection pattern," 2011 International Conference on Computer Science and Service System (CSSS), Nanjing, 2011, pp. 914-917, doi: 10.1109/CSSS.2011.5974374.

[2] Moore, Andrew & Zuev, Denis. (2005). Internet traffic classification using Bayesian analysis techniques. Sigmetrics Performance Evaluation Review - SIGMETRICS. 33. 50-60. 10.1145/1064212.1064220.

[3] Wang, W., Zhu, M., Zeng, X., Ye, X., & Sheng, Y. (2017). Malware traffic classification using convolutional neural network for representation learning. In *Proceedings of the 2017 International Conference on Information Networking (ICOIN)* (pp. 712 - 717). IEEE.

[4] Li, Ding & Li, Wenzhong & Wang, Xiaoliang & Nguyen, Cam-Tu & Lu, Sanglu. (2020). App Trajectory Recognition over Encrypted Internet Traffic based on Deep Neural Network. Computer Networks. 179. 107372. 10.1016/j.comnet.2020.107372.

[5] 김지민, MYUNG-SUP KIM, 김란아, Yoon-Seong Jang and 백의준. (2025). WeChat Behavior Analysis with Session and Packet Bundle Level. KNOM Review, 28(2), 33-40.