암호화된 VPN 프로토콜 탐지를 위한 오토인코더 기반 이미지 분류 기법

홍석현 [1], 박예진 [1], 엄서정 [1], 김정훈 [1], 김태욱 [2], 조영필 [3]
[1] 한양대학교 미래자동차-SW 융합전공 석박통합과정
[2] 한양대학교 미래자동차-SW 융합전공 석사과정
[3] 한양대학교 컴퓨터소프트웨어학과 교수

{ghazard8572, pkyj09029, tjwjdus, qkenr7895, rlawjdgns527, ypcho}@hanyang.ac.kr

Autoencoder based image classification technique for detecting encrypted VPN protocols

[1]Dept. of Computer and Software (Automotive-Computer Convergence), Han-Yang University
[2]Dept. of Computer and Software (Automotive-Computer Convergence), Han-Yang University
[3]Dept. of Computer Science, Han-Yang University

요 약

최근 COVID-19 팬데믹으로 전 세계적으로 원격 근무로의 전환 속도가 가속화되면서 VPN 을 사용하는 기업이 증가하면서 VPN 을 통한 국내 개인정보 및 기술 유출이 빈번하게 일어나고 있다. 기존 전통적인 네트워크 프로토콜 분석 방법은 다양한 우회 방법과 패킷의 암호화로 인해서 VPN 프로토콜 탐지가 불가능하다. 하지만 AI 기반 모델을 사용하면 암호화된 패턴을 학습을 하여 분류가 가능하다. 따라서 본 논문에서는 오토인코더 기반 이미지 분류 기법으로 전통적인 방법으로 탐지하기 불가능하다고 생각했던 암호화된 VPN 패킷 중의 VPN 프로토콜을 직접 수집 및 탐지했고 성능이 0.99 가 나왔다.

1. 서론

최근 COVID-19 팬데믹으로 인해 전 세계적으로 원격 근무로의 전환 속도가 가속화되면서, 조직들이 네트워크 트래픽을 보호하기 위해 VPN(Virtual Private Network) 사용을 크게 늘리고 있다. 그러나 VPN 사용 증가는 네트워크 트래픽의 보안과 무결성을 유지하는 데 있어 도전과제를 만들어 냈다. 또한 급성장한 클라우드 및 IoT 시장과 함께 VPN 프로토콜들이 난립하고 개인정보 보호의 중요성으로 인해 암호화가 필수적으로 요구되는 오늘날의 네트워크 환경에서는 기존의 VPN 프로토콜 분류 기술의 적용이 어렵다.[1] 국가 정보원에서 최근 5 년간 국내 첨단 기술유출 사건은 총 93 건으로, 피해액은 25 조원에 이르는 것으로 추산된다.[2] 기술 유출 경로 중 하나인 네트워크를 감시하여 이를 방지할 수 있지만, 송수신 데이터를 암호화하고 송수신 IP 주소를 숨겨주는 VPN 기술을 통해 우회가 가능하다. 기업 내 무단 VPN 사용뿐만 아니라 VPN 프로토콜을 탐지하면 무단 액세스 또는 데이터 유출 시도를 감지하고 완화하는데 도움이 된다. 하지만 VPN 의 사용을 차단하려면 VPN 을 식별해야 하는 데, 전통적인 방법으로 포트 접근법으로는 다양한 우회 방법들이 있고 패킷 분석 기반 접근법은 VPN 패킷이 암호화가 되면 적용이 불가능하다. 따라서 본 논문은 AI 로 기존에 식별하기 불가능하다고 생각했던 암호화된 VPN 프로토콜을 탐지를 목표로 한다.

2. 관련연구

2.1. 전통적인 네트워크 프로토콜 분류 방법

네트워크 분석으로 사용하는 와이어샤크는 전통적으로 사용하는 포트 기반 접근법은 네트워크 포트를 기반으로 프로토콜을 분류한다. 간단하면서도 빠른 방법이지만 임의 포트 할당, 포트 포워딩 등이 존재해 정확성이 높지 않다.[3] 패킷 분석 기반 접근법은 네트워크 패킷에 존재하는 프로토콜의 시그니처를 이용해 분류한다. 하지만 현대의 VPN 프로토콜에서 사용하는 암호화 및 난독화 기술은 VPN 트래픽의 암호화된 프로토콜들을 효과적으로 분류하는 것을 크게

방해한다.[4]

2.2. AI 기반 네트워크 프로토콜 분류 방법

통계 기반 접근법은 프로토콜 페이로드의 크기, 트랜잭션 시간, 세션 정보 등을 기반으로 하여 분류하는 방법이다.[5] 지도학습 기반 접근법은 네트워크 패킷의 처음 512 바이트를 가지고 CNN 모델을 활용하여 VPN 을 분류했다. 비지도 학습 기반 접근법으로 네트워크의 특징 정보를 추출하여 클러스터링을 통해 VPN 을 분류했다.[6] 강화 학습 기반 접근법은 네트워크 프로토콜의 필드를 분류하는 방법을 제안했다. 본 논문은 CNN 모델을 활용하여 VPN 패킷을 대상으로 암호화된 VPN 프로토콜을 분류 및 탐지하는 것을 목표로 한다.

3. 실험 설계

3.1. 암호화된 VPN 프로토콜 탐지를 위한 오토인코더 기반 이미지 분류 모델

암호화된 VPN 프로토콜의 이미지 기반 분류를 위한 오토인코더 기반 머신러닝 모델을 사용하는 접근법을 제안한다. 네트워크 트래픽의 패킷을 암호화된 VPN 프로토콜의 페이로드 이미지로 만들고 대중적으로 많이 사용한다고 알려진 5 가지 VPN 프로토콜을 탐지 및 분류한다. 이 방법은 전통적인 기술보다 훨씬 발전된 것으로, 점점 더 암호화된 온라인 세계에서 네트워크 보안을 보장하기 위한 방법론이다. 다음 5 가지 VPN 프로토콜은 IPSEC 과 마이크로소프트의 SSTP 는 TCP 의 페이로드가 인증서로 전부 암호화되어 인증서가 없으면 페이로드를 볼 수 없다. 대중적으로 많이 쓰이는 프로토콜인 OPENVPN 과 Wireguard 의 경우 TCP 또는 UDP 페이로드에 VPN 프로토콜을 구분할 수 있는 프로토콜 헤더 식별코드가 있다. 하지만 오토인코더 기반으로 학습할 경우 비암호화된 VPN 프로토콜 및 암호화된 VPN 및 난독화된 VPN 프로토콜이더라도 특정 패턴을 학습하여 분류 모델을 식별할 수 있다.

3.2. 데이터셋

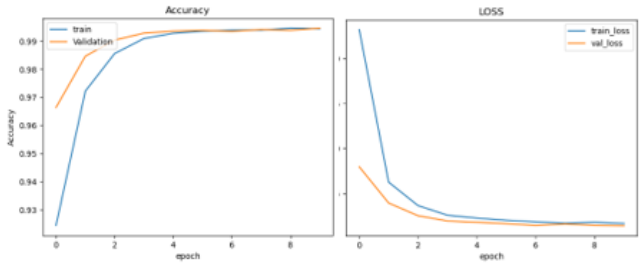
VPN 프로토콜을 탐지하는데 암호화된 VPN 과 암호화되지 않은 VPN 프로토콜을 직접 수집했다. OPENVPN 과 Wireguard 의 경우 NordVPN 으로 수집했고 나머지 VPN 의 경우는 사 VPN Gate 로 수집했다.[8,9] OPENVPN, SSTP, IPSEC, Wireguard 를 분류했다. OPENVPN 의 경우 실험을 통해서 TCP 와 UDP 의 프로토콜이 다른 패턴으로 분류되어 추가적으로 분류를 했다. VPN 패킷은 MTU(Maximum Transmission Unit)이 1,400 바이트로 포트 및 송수신 IP 의 학습 편향을 막기 위해 데이터셋의 경우 TCP 와 UDP 의 페이로드로만 학습을 했다. 고정된 32 X 32 이미지로 만들기 위하여 패킷의 남은 바이트는 0 으로 채웠다. 다음 표 1 은 데이터셋 학습을 위한 VPN 프로토콜이며 8:2 로 구분했다.

<표 1> VPN 프로토콜 데이터셋

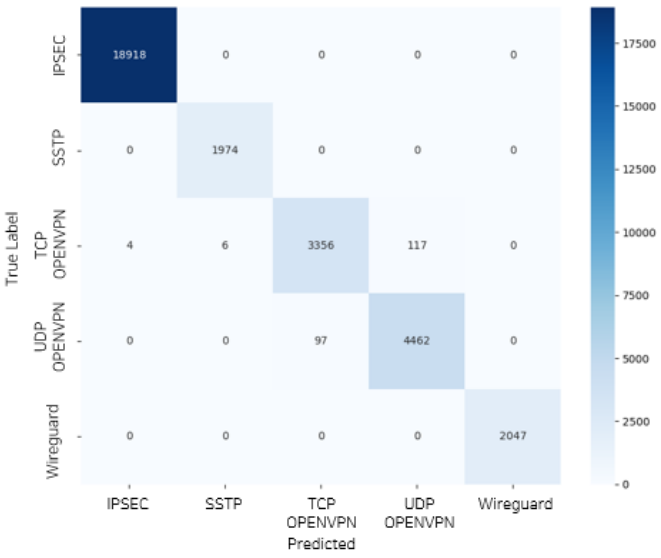
Protocol	Train	Test
IPSEC	75,672	18,918
SSTP	7,892	1,974
TCP OPENVPN	13,935	3,356
UDP OPENVPN	18,232	4,462
Wireguard	46,270	2,076

4. 실험 결과

성능 평가는 훈련과 검증 그리고 테스트 데이터셋으로 7:1:2 로 나누었다. 다음 그림 1 은 암호화 및 비암호화 VPN 프로토콜 패킷을 10 번 학습 후 분류한 측정 결과이다. 10 번 진행 시 loss 가 0.015 이고 accuracy 가 0.99 그리고 AUC 가 0.99 로 높은 성능 결과가 나왔다. 테스트 데이터셋으로 예측 결과 Accuracy 와 F1-Score 가 모두 0.99 가 나왔다. Wireguard 를 커스텀한 VPN 프로토콜인 Nordlynx 를 탐지한 결과 Wirueguard 로 탐지되는 것을 보았을 때 커스텀한 VPN 프로토콜도 식별을 할 수 있을 것으로 예상된다. 하지만 CNN 의 모델의 한계점으로 인해서 분류하지 않은 VPN 프로토콜과 HTTPS 나 TLS 로 암호화된 이메일 프로토콜(SMTP, IMAP, POP3) 등은 식별이 되지 않는다.



(그림 1) Accuracy 및 Loss 측정 결과



(그림 2) VPN 프로토콜 분류 결과

5. 결론

본 연구는 COVID-19 팬데믹으로 인해 VPN 을 통한 원격 근무에 따른 정보 및 개인정보 유출을 막고자 VPN 트래픽 중 암호화된 VPN 프로토콜을 탐지 및 분류하는 모델을 제안한다. 무단으로 VPN 을 사용하거나 VPN 을 통해 들어온 네트워크 패킷을 확인을 오토인코더 기반 이미지 분류 기법으로 암호화된 VPN 프로토콜로 탐지한다. VPN 프로토콜이 식별되지 않은 VPN 패킷을 직접 수집했다. 5 가지 대중적으로 사용되는 VPN 프로토콜을 분류하는 모델을 만들었고 성능이 0.99 가 나왔다. 향후 연구로 VPN 프로토콜이 아닌 TLS 로 암호화된 프로토콜도 탐지하는 모델 연구하여 안전한 네트워크 관리를 할 것이다.

이 논문은 과학기술정보통신부의 재원으로 정보통신기획평가원(No. 2020-0-01840, 스마트폰의 내부데이터 접근 및 보호 기술 분석)과 한국연구재단(No. NRF-2022R1A4A1032361, Processing-in-Memory 보안 기술 개발)의 지원을 받아 수행된 연구임

참고문헌

[1] Velan P, Cˇ ermák M, Cˇ eleda P, Drašar M (2015) A survey of methods for encrypted traffic classification and analysis. Int J Netw Manag 25(5):355–374

[2] 국정원 제 4 차 산업기술의 유출방지 및 보호에 관한 종합계획

[3] Madhukar A, Williamson C (2006) A longitudinal study of p2p traffic classification. In: Modeling, analysis, and simulation of computer and telecommunication systems, 2006. MASCOTS 2006. 14th IEEE international symposium on, IEEE, pp 179–188

[4] Khalife J, Hajjar A, Diaz-Verdejo J (2014) A multilevel taxonomy and requirements for an optimal traffic-classification model. Int J Netw Manag 24(2):101–120

[5] OTFOLLAHI, Mohammad, et al. Deep packet: A novel approach for encrypted traffic classification using deep learning. Soft Computing, 2020, 24.3: 1999-2012

[6] PIET, Julien; NWOJI, Dubem; PAXSON, Vern. Ggfast: Automating generation of flexible network traffic classifiers. In: Proceedings of the ACM SIGCOMM 2023 Conference. 2023. p. 850-866

[7] JAMIL, Hasibul; WENG, Ning. Multibit tries packet classification with deep reinforcement learning. In: 2020 IEEE 21st International Conference on High Performance Switching and Routing (HPSR). IEEE, 2020. p. 1-6.

[8] NordVPN, https://nordvpn.com/

[9] VPN Gate, https://www.vpngate.net/