안전한 네트워크 구축을 위한 컨볼루션 신경망 기반 악성트래픽 탐지

지세현, 박지태, 백의준, 김명섭

고려대학교

{sxzer, pjj5846, pb1069, tmskim}@korea.ac.kr

Malicious Traffic Detection Based on Convolution Neural Network for Secure Network Construction

Se-Hyun Ji, Jee-Tae Park, Eui-Jun Baek, Myung-Sup Kim

Korea Univ.

요 약

오늘날 인터넷이 발달함에 따라 네트워크 환경이 증가하고 있고, 이를 위협하는 악성 트래픽의 종류와 패턴도 갈수록 다양해지고 있다. 악성 트래픽이란 네트워크 환경에 유입되어 인터넷 망을 교란시키거나, 특정 네트워크 서버, 호스트 등에 피해를 끼칠 의도를 가지고 발생시키는 트래픽을 뜻한다. 안전한 네트워크 환경을 구축하기 위해서는 악성 트래픽을 탐지하는 것은 필수적이다. 보편적인 악성 트래픽 탐지 기법으로는 트래픽의 시그니처 기반의 탐지 방법이 있다. 그러나 동적인 포트번호 혹은 암호화된 페이로드를 갖는 트래픽이 등장함에 따라 시그니처 기반의 탐지를 어렵게 하고 있고, 시그니처를 추출하는 과정에서 많은 시간과 비용이 발생한다. 이러한 문제를 해결하기 위해 기계학습 알고리즘을 적용한 악성 트래픽 탐지 기법이 제시된다. 본 논문은 기계학습 알고리즘인 컨볼루션 신경망 기반의 악성 트래픽 탐지 방법을 제시하였다. 악성 트래픽에 대한 실험을 통해 제시한 기법의 적합성을 검증 하였다.

Ⅰ. 서론 및 관련연구

오늘날 인터넷이 발달하고 있고, 그에 따라 네트워크 환경이 급격하게 성장하고 있다. 네트워크 환경을 위협하는 악성 트래픽의 종류와 패턴도 갈수록 다양해지고 있다. 악성 트래픽은 정상적인 응용에 의해 발생한 트래픽이 아닌 인터넷 망을 교란시키거나, 특정 네트워크 서버, 호스트 등에 피해를 끼칠 의도를 가지고 발생시키는 트래픽을 뜻한다[1]. 네트워크 사용자는 고품질의 서비스를 제공받고, 네트워크 운영자는 서비스 제공의 신뢰성 확보 및 안정적인 제공을 위해 안정적인 네트워크 환경을 구축하는 것이 필요하다. 안전한 네트워크 환경을 구축하기 위해 악성 트래픽 탐지는 필수적이다.

보편적인 악성 트래픽 탐지 기법으로는 트래픽의 시그니처 기반 탐지 방법이 있다. 포트, 페이로드, 통계적 정보를 이용하여 시그니처를 정의하고 이를 바탕으로 트래픽을 식별한다[2]. 그러나 동적인 포트번호 혹은 암호화된 페이로드를 갖는 트래픽이 등장함에 따라 시그니처 기반의 탐지를 어렵게 하고 있고, 시그니처를 추출하는 과정에서 트래픽 패턴 변화를 관리자가 인지해야 할 뿐만 아니라 많은 시간과 비용이 발생한다.

표 1 시그니처 기반 트래픽 식별

Signature	Example	Properties
Port	80:HTTP 21:FTP	Using the fixed port number
Payload	“GET” “host”	Using the Unique pattern in payload
Statistic	Packet Size	Using the statistic information
Behavior	# of port # of IP	Using the pattern of behavior

시그니처 기반 트래픽 탐지 기법의 문제를 해결하기 위해 기계학습 알고리즘을 적용한 응용 트래픽 탐지 기법이 부상하고 있다. 기존의 머신러닝 알고리즘 기반의 트래픽 탐지 기법은 CNN(Convolution Neural Network), LSTM, K-Means 알고리즘을 적용한 트래픽 탐지 기법이 있다. LSTM, K-Means 알고리즘은 패턴이 불규칙적인 트래픽 데이터에 대해 낮은 탐지 성능을 보이는 반면 CNN 알고리즘은 비교적 높은 탐지 성능을 보인다[3].

표 2 머신러닝 알고리즘 기반 트래픽 식별

ML Algorithm	Properties	Disadvantages
2D CNN	2D Image Mapping	Loss of time information
	Time stamp, Feature Mapping	
1D CNN	1D Sequence Window	
LSTM	Feature Vector per minute	Difficult to detect irregular traffic
	Time Unit Vector	
	Normal Sequence Time Information	
K-Means	Dimensional Reduction Cluster	Difficult to detect with the same distribution of characteristics

본 논문에서는 서론에 이어 2장에서 기계학습 알고리즘인 컨볼루션 신경망을 기반으로 구성된 악성 트래픽 탐지 모델에 대해 언급한 뒤 3장에서 만들어진 모델의 탐지 성능을 분류 정확도를 통해 모델의 적합성을 검증한다. 마지막으로 4장에서 결론 및 향후 연구에 대해 언급한 뒤 논문을 마친다.

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No.2015R1D1A3A01018057)과 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임(No. 2017-0-00513, Security Analytics 기반의 이기종 보안솔루션 위협 분석 및 대응 기술 개발)

Ⅱ. 본론

본 장에서는 제안하는 악성 트래픽 탐지 모델에 대해 설명한다. 컨블루션 신경망은 이미지 분류에 특화된 기계학습 알고리즘이다. 악성 트래픽 탐지 모델은 그림 1과 같이 구성된다. 2개의 Convolution Layer로부터 특징 값을 추출하고, 1개의 Fully Connected Layer와의 연산과정을 거쳐 나온 값을 통해 트래픽을 식별한다. 각 Convolution Layer는 Feature Map을 추출하기 위한 Convolution 과정, 추출 된 Feature map 에 적용하기 위한 활성함수, Feature Map으로부터 가장 큰 값을 추출하기 위한 Max Pool 과정으로 구성되었다. Fully Connected Layer는 추출 된 특징 값들을 Neural Network와 Sigmoid 함수를 통해 연산을 하여 분류 작업을 거친다.

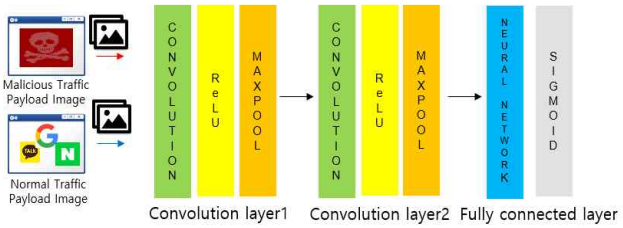

그림 1 악성 트래픽 탐지 모델 구조

컨볼루션 신경망 기반 악성트래픽 탐지 기법은 그림 2의 과정을 거친다. 정상 및 악성 트래픽에 대해서, 본 연구팀이 개발한 프로그램인 Payload Generator를 이용해 Flow with Packet 형태로 변환 한 뒤, Flow 단위의 페이로드 값을 추출 한다. Flow with Packet은 5가지 속성(Source IP, Source Port, Protocol, Destination IP, Destination Port)이 같은 패킷의 집합이다. 추출 된 Payload 값은 컨볼루션 신경망 학습에 적합한 이미지로 변환 하여 탐지 모델을 완성한다.

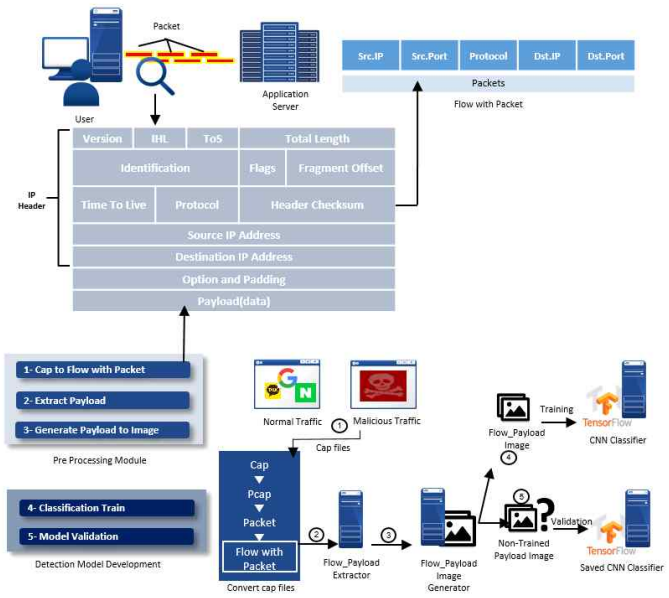

그림 2 컨볼루션 신경망 기반 악성 트래픽 탐지 기법

완성된 컨볼루션 신경망 기반 악성 트래픽 탐지 모델의 성능을 학습 데이터 분류 정확도를 통해 모델의 적합성을 검증한다.

Ⅲ. 실험 및 결과

컨볼루션 신경망 기반의 악성 트래픽 탐지 기법의 효율성을 검증하기 위해 실험을 진행한다. cap파일로 구성 된 정상 및 악성 트래픽에 대한 학습 데이터를 구성하기 위해 Flow with Packet 형태로 변환 한 뒤, 페이로드만을 추출하여 본 연구팀이 개발한 프로그램을 통해 784(28*28)의 크기를 갖는 페이로드 이미지를 그림 3과 같이 구성한다.

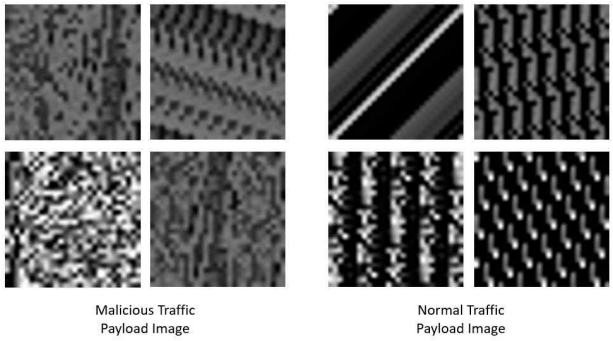

그림 3 Payload Image

각 트래픽 별 1000개의 Flow with Packet의 페이로드 이미지를 컨볼루션 신경망 분류 모델의 학습결과는 표3 과 같다. 악성 트래픽 데이터는 ransomware, malspam, android-malware로 구성하였고, 정상 트래픽 데이터는 Google, Naver, Kakaotalk으로 구성하였다. 1000번의 학습 이후 악성 트래픽의 Precision은 97.6%, Recall은 97.0%를 나타내었다. 실험 결과를 통해 컨볼루션 신경망 기반의 악성 트래픽 탐지 모델의 높은 Precision 과 Recall을 나타내는 것을 확인하였다.

표 3 악성 트래픽 탐지 성능표

Traffic	Accuracy	Precision	Recall
Malicious Traffic	97.3%	97.6%	97.0%
Normal Traffic		97.0%	97.7%

Ⅳ. 결론 및 향후 연구

본 논문은 안전한 네트워크 구축을 위한 컨볼루션 신경망 기반의 악성 트래픽 탐지 모델을 제안하였다. 정상 트래픽 및 악성트래픽에 대한 분류 실험을 통하여 제안한 모델의 학습결과가 높은 수치의 Precision과 Recall을 나타내는 것을 확인함으로써 모델의 적합성을 검증하였다.

향후 연구로는 제안한 기법을 기반으로 다양한 종류의 트래픽을 대상으로 실험을 진행하여 정교한 탐지 모델을 구상할 계획이다.

참 고 문 헌

[1] 한명지, 임지혁, 최준용, 김현준, 서정주, 유철, 김성렬, 박근수. (2014). X-means 클러스터링을 이용한 악성 트래픽 탐지 방법. 정보과학회논문지, 41(9), 617-624.

[2] 심규석, 구영훈, 이성호, Baraka D. Sija, 김명섭, "최신 네트워크 응용 분류를 위한 자동화 페이로드 시그니쳐 업데이트 시스템", 통신학회 논문지 Vol.42 No.01, Jan. 2017, pp. 1-10.

[3] T. Y. Kim and S.-B. Cho, "C-LSTM 신경망을 이용한 웹 트래픽 이상탐지," Proc. of Korea Business Intelligence Data Mining Fall Conference, 2017. 11.