

엔트로피를 이용한 이상 트래픽 측정: 실제 사례를 통한 접근

김정현, 원유진

한양대학교 전자컴퓨터통신공학과

e-mail : junghyun@ece.hanyang.ac.kr, yjwon@ece.hanyang.ac.kr

Anomalous Traffic Measurement using Entropy: An Empirical Study

Jung-Hyun Kim, Youjip Won

Department of Electronics and Computer Engineering
Hanyang University

Abstract

Entropy, one of leading metrics on anomalous traffic, attracts researcher's attention since a packet sampling and a traffic volume impact little on entropy value. In this paper, we apply the entropy metric to a domestic network traffic trace which has real anomalous traffics. We used source IP address/port and destination IP address/port that are important attributes of a packet as entropy variable. We found that entropy value of multiple-port DoS attack shows something related to a staircase fashion. Also, we show a possibility of detection of anomalous traffic on small time scale.

I. 서론

인터넷의 사용이 보편화되면서 인터넷 보안(Internet Security)에 대한 관심이 높아지고 있다. 2000년에 있었던 Yahoo나 Ebay에 대한 DDoS(distributed denial-of-service) 공격이 큰 피해를 주면서 주목을 받았다[1]. 또한 인터넷 강국인 우리나라에서도 DoS(denial-of-service) 또는 DDoS 공격으로 의심되는 사건이 종종 일어나고 있다. 최근 인터넷 보안 업계나 학계에서는 의도적인 공격에 의해 발생된 이상 트래픽(anomalous traffic)을 탐지하기 위한 많은 시도를

하고 있다. 대표적으로 트래픽량(traffic volume), 신호 처리(signal processing), 웨이브렛(wavelet), 엔트로피(entropy) 등이 이상 트래픽의 측정 규준(metric)으로 사용되고 있다. 최근 엔트로피가 트래픽 양이나 샘플링에 크게 영향을 받지 않는 것으로 알려지고 있다[2]. 이상 트래픽 탐지 연구에 있어서 엔트로피의 특성은 매우 유용하다고 할 수 있다. 본 연구는 국내에서 캡처된 네트워크 트래픽을 기반으로 효과적인 엔트로피의 사용 방향을 제시할 것이다. 연구에 사용된 트래픽에는 실제 DoS 공격에 의해 발생된 이상 트래픽이 포함되어 있다.

II. 본론

2.1 엔트로피(entropy)

엔트로피는 랜덤 변수(random variable)의 불확실성(uncertainty)을 나타내는 측정규준(metric)으로써 다음과 같이 정의 된다[3].

$$H(X) = - \sum p(x) \log p(x). \quad (1)$$

기존에는 데이터 압축이나 부호화 등에 주로 사용되었나, 최근에는 이상 트래픽 분석과 탐지에도 이용되기 시작하고 있다. 엔트로피 분석은 변수의 선택이 매우 중요하다. 본 연구에서는 트래픽의 분석에서 매우 중요한 패킷의 source IP address/port, destination IP address/port를 변수로 선택한다. 기존 연구에서는 주로 특정한 1개만을 분석하였으나[2], 본 연구에서는 4개의 변수에 대해서 어떤 변화가 있는지 확인하고 비

교 할 것이다. source port에 대한 엔트로피 분석을 위해 사용된 $p(x)$ 는 다음과 같다.

$$p(x) = \frac{\text{packet count of distinct [src port]}}{\text{total packet count}} \quad (2)$$

다른 변수의 엔트로피는 [src port] 대신 [src addr], [dst port], [dst addr]을 사용하였다.

2.2 이상 트래픽의 측정

본 연구에서 사용한 트래픽에는 DoS 공격을 여러 포트에 대해서 공격하는 UDP 기반의 ‘다중포트 DoS 공격(multiple-port DoS attack)’이 포함되어 있다. 보유 중인 전체 트래픽 데이터는 2004년 10월 30일부터 11월 4일까지 캡쳐 되었으며, 용량은 1.5TB에 이른다. 모든 트래픽은 MySQL로 데이터베이스화 되어 있다.

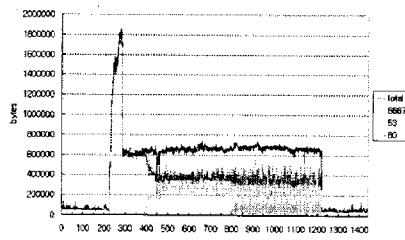
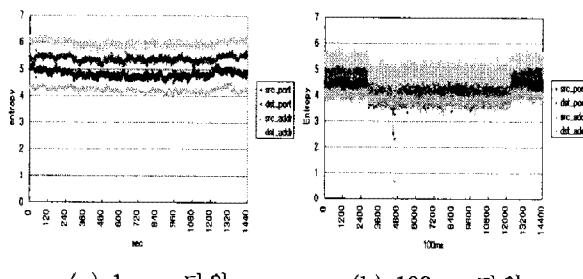
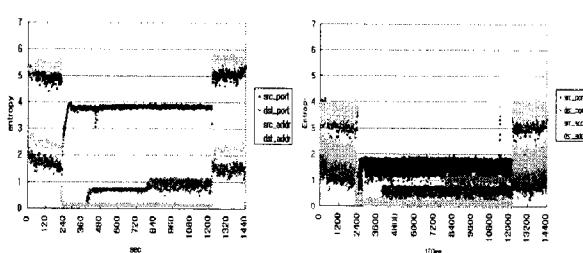


그림 1. 공격을 포함한 24분간 UDP 구간



(a) 1 sec 단위

(b) 100ms 단위



(a) 1 sec 단위

(b) 100ms 단위

기존 연구에서는 분 단위(1분-10분정도)의 긴 시간 단위를 기준으로 분석하였으나 현실적인 이상 트래픽 탐지에는 긴 시간 단위 분석은 유용하지 않다. 대부분의 DoS 공격은 수십 분정도간 진행되기 때문이다[1]. 본 연구에서는 가능한 짧은 시간 안에 이상 트래픽을 탐지하는 데 초점을 맞추어 분 단위로 분석하였다.

지할 수 있는 방안을 연구하기 위해 1초, 0.1초(100ms) 단위로 에트로피를 분석하였다.

2.3 분석

엔트로피 분석을 위한 24분간의 UDP 트래픽을 그림 1과 같이 나타내었다. 초당 수만개의 패킷이 갑자기 발생하고 있는 것을 확인할 수 있다. 이는 다중포트 DoS 공격이다. 6667(IRC), 53(DNS), 80(HTTP) 포트를 동시에 번갈아가며 공격하고 있다. 이러한 현상을 어떻게 엔트로피가 반영하고 있는지 그림 2, 3과 같이 나타내었다. 그림 2에서, UDP기반 이상 트래픽의 발생이 TCP 트래픽에 영향을 주는 것을 볼 수 있다. 흥미롭게도 0.1초 단위로 나타낸 그림 2 (b)에서 그 영향이 더 확실하게 나타난다. 공격이 발생한 UDP 트래픽을 분석한 그림 3을 보면 더욱 흥미로운 현상을 발견할 수 있다. [src_addr], [dst_addr], [src_port]에서는 단지 급격히 낮아지는 엔트로피를 보였으나, [dst_port]에서는 엔트로피의 계단 현상을 확인할 수 있다. 이는 특정한 공격자가 공격 대상의 여러 포트에 대해 공격 할 때 발생한다. 본 연구에서 발견한 다중 포트 공격은 순차적으로 6667, 53, 80의 순서로 포트 공격이 시작되기 때문에, 이 현상이 계단 형태(staircase fashion)의 엔트로피로 나타나기 때문이다. 본 연구에서 발견한 DoS 공격은 일반적인 IP spoofing[1]을 사용하지 않았다. 단지 source IP address를 임의로 변경하였다. 그렇기 때문에 [src_addr]의 엔트로피가 낮아지는 것을 볼 수 있다. 일반적인 IP spoofing을 사용할 경우 [src_addr]의 엔트로피는 급격히 증가하게 된다.

IV. 결론 및 향후 연구 방향

본 연구에서는 엔트로피가 이상 트래픽의 발생을 어떻게 반영하고 있는지 실제 트래픽을 대상으로 분석하였다. 본 연구에서 발견한 엔트로피의 계단현상을 찾았던 것과, 짧은 시간 단위(0.1초)에서도 이상 트래픽 탐지가 가능함을 보였다. 앞으로는 엔트로피를 이용한 인터넷 웹 분석과 탐지에 대해서 연구할 계획이다.

참고문헌

- [1] D. Moore, G. M. Voelker, and S. Savage, "Inferring Internet Denial-of-Service Activity", The 2001 USENIX Security Symposium, 2001.
 - [2] D. Brauckhoff, B. Tellenbach, A. Wagner, M. May, "Impact of Traffic Sampling on Anomaly Detection Metrics", IMC 2006.
 - [3] Thomas M. Cover, Joy A. Thomas, "Elements of Information Theory", Wiley Interscience, 1991