情報保護學會論文誌
第20卷 第4號, 2010. 8

엔트로피를 이용한 분산 서비스 거부 공격 탐지에 효과적인 특징 생성 방법 연구*

김 태 훈,† 서 기 택, 이 영 훈, 임 종 인, 문 종 섭‡
고려대학교 정보경영공학전문대학원

An Effective Feature Generation Method for Distributed Denial of Service Attack Detection using Entropy*

Tae-hun Kim,† Ki-taek Seo, Young-hoon Lee, Jong-in Lim, Jong-sub Moon‡
Graduate School of Information Management and Security, Korea University

요 약

최근 분산 서비스 거부 공격의 근원인 악성 봇 프로그램이 널리 유포되고 있으며 보안이 유지되지 않는 PC를 통하여 악성 봇이 설치된 PC의 수가 기하급수적으로 증가하고 있다. 이를 통한 분산 서비스 거부 공격이 계속적으로 발생하고 있으며 최근 금품을 요구하는 사례도 발견되었다. 따라서 분산 서비스 거부 공격에 대응하기 위한 연구가 필요하며 본 논문에서는 네트워크 패킷 헤더의 속성에 대해 불확실성을 나타내는 척도인 엔트로피를 이용하는, 분산 서비스 거부 공격 탐지에 효과적인 특징 생성 방법을 제안한다. DARPA 2000 데이터셋과 직접 실험을 통해 구성한 분산 서비스 거부 공격 데이터셋에 대해 향상된 엔트로피 수식과 효율적인 엔트로피 계산 기법, 다양한 엔트로피 특징 값을 사용하는 제안 기법을 적용해보고 베이지안 네트워크 분류기를 이용하여 분류함으로써 제안하는 방법이 효과적인지를 검증해 본다.

ABSTRACT

Malicious bot programs, the source of distributed denial of service attack, are widespread and the number of PCs which were infected by malicious bot program are increasing geometrically thesedays. The continuous distributed denial of service attacks are happened constantly through these bot PCs and some financial incident cases have found lately. Therefore researches to response distributed denial of service attack are necessary so we propose an effective feature generation method for distributed denial of service attack detection using entropy. In this paper, we apply our method to both the DARPA 2000 datasets and also the distributed denial of service attack datasets that we composed and generated ourself in general university. And then we evaluate how the proposed method is useful through classification using bayesian network classifier.

Keywords: distributed denial of service attack, feature generation, entropy

I. 서 론

인터넷을 통한 전세계의 네트워크화는 편리함을 제공하기도 했지만 그 이면에는 많은 보안상의 위협이 존재한다. PC에 대한 개인이나 단체의 보안 의식은 많이 뒤떨어져 있으며 이런 PC들이 악성 봇을 구성하기 위한 표적이 되고 있다. 분산 서비스 거부 공격을

* 접수일(2010년 3월 22일), 수정일(2010년 5월 3일), 게재확정일(2010년 5월 24일)
* "이 연구에 참여한 연구자(의 일부)는 2단계 BK21사업의 지원비를 받았음"
† 주저자, kurie@korea.ac.kr
‡ 교신저자, jsmoon@korea.ac.kr

유발하는 악성 봇 PC는 전 세계적으로 2007년에는 약 3만대, 2008년에는 약 22만대, 2009년에는 85만대 정도 존재했으며 계속적으로 증가하고 있는 추세이다[1]. 분산 서비스 거부 공격은 공격 대상 서버의 가용성을 저해하는데 목적을 두고 있기 때문에 인터넷을 기반으로 하는 기업 및 기관의 경우, 피해가 더욱 심각할 것이라 예상된다. 최근의 분산 서비스 거부 공격은 공격 유형이 기존의 방법과는 다른 형태를 띄고 있으며 이로 인해 공격의 탐지 및 차단이 어렵다. 따라서 본 논문에서는 엔트로피를 이용해 효과적으로 분산 서비스 거부 공격을 탐지할 수 있는 특징 생성 기법을 제안한다.

본 논문은 구성은 다음과 같다. 2장에서는 분산 서비스 거부 공격을 분석하고 엔트로피와 이를 이용한 기존의 탐지 연구를 알아본다. 3장에서는 효과적으로 분산 서비스 거부 공격을 탐지할 수 있는 특징 생성 기법을 제안하고 4장에서는 실험 및 성능 평가를 기술하며 마지막으로 5장에서는 결론을 맺는다.

II. 관련 연구

2.1 분산 서비스 거부 공격

서비스 거부 공격은 의도적으로 운영 중인 시스템 또는 서비스 자원을 소모시켜 정상적인 서비스를 제공하지 못하도록 방해하는 것을 목적으로 하는 공격 행위이다[2]. 서비스 거부 공격이 소수에 의해서 이루어지는 반면에 분산 서비스 거부 공격은 불특정 다수에 의해서 동시다발적으로 공격이 이루어진다는 특징이 있다. 분산 서비스 거부 공격은 초기의 웹페이지 새로고침을 계속적으로 시도함으로써 정상적인 접속을 방해하는 단순하고 원시적인 형태에서부터 다수의 인원이 공격도구를 이용하여 공격하는 중간적인 형태, 그리고 최근에는 악성 봇에 감염된 봇넷을 이용해 공격을 하는 방법이 사용되고 있다[3]. 일반적으로 공격자는 봇넷을 관리하고 명령을 내릴 수 있는 봇 마스터 프로그램을 가지고 있으며 웹 서버를 해킹하거나 스팸 메일 전송 등을 통해 악성 봇을 유포한다. 보안에 취약한 PC를 사용하는 사용자는 공격당한 웹 서버를 방문하거나 스팸 메일을 열어보는 행위만으로도 감염되어 봇넷의 일부가 된다. 이렇게 구성된 봇넷을 이용하여 공격자는 자신이 직접 공격을 하지 않고도 봇 마스터에서 명령을 내림으로써 다수의 PC를 이용해 효과적인 공격을 할 수 있다. 최근의 공격에서는 봇 마스터를 이용하지 않고 악성 봇을 유포할 때 미리 특정 시간에 특정 서버를 공격하라는 정보를 입력해 놓는 스케쥴링 공격 사례도 발견되었다. 서비스 거부 공격은 크게 세션 고갈 공격, 대역폭 고갈 공격, 서버 자원 고갈 공격의 3가지 유형으로 나눌 수 있다. 먼저 세션 고갈 공격은 SYN Flooding을 이용한 공격으로써 TCP의 연결 과정인 Three-way Handshaking을 이용한 방법이다. 공격자가 공격 대상 서버에 SYN패킷을 스푸핑하여 특정 포트로 전송하게 되면 이 포트의 대기 큐(Backlog Queue)가 가득 차 이 포트로 들어오는 연결 요청을 무시하게 되고 정상적인 서비스 제공이 불가능하게 된다. 다음으로 대역폭 고갈 공격은 과다한 트래픽을 공격 대상 서버에 전송함으로써 네트워크를 마비시키는 공격이며 UDP 패킷을 이용한 UDP Flooding 공격과 ICMP 패킷을 이용한 ICMP Flooding 공격이 있다. 마지막으로 서버 자원 고갈 공격은 최근 많이 이용되는 공격으로 웹 부하 공격이라고도 하며 웹 서버에 과부하를 일으키는 공격이다. 이 공격에는 HTTP Get Flooding 공격이 대표적이며 대량의 HTTP Get 메세지를 공격 대상 서버에 전송해 정상적인 사용자의 웹 서버 접속 요청을 받아들이지 못하게 만드는 기법이다.

2.2 엔트로피를 이용한 분산 서비스 거부 공격 탐지 기법

엔트로피를 이용하여 분산 서비스 거부 공격을 탐지하는 여러가지 연구에 대한 현황을 분석한다. 엔트로피는 불확실성을 나타내는 척도이며 완전히 무관한 데이터 사이에서는 높은 값을 나타내고 관련이 있는 데이터 사이에는 낮은 값을 나타낸다[4]. 이러한 엔트로피의 특성을 이용하여 분산 서비스 거부 공격을 탐지할 수 있다. 먼저 엔트로피의 특징부터 살펴보도록 하겠다.

(a) 엔트로피의 특징

확률변수 x_i의 확률 p_i를 n개의 독립적인 요소로 볼 때 엔트로피는 다음과 같이 정의된다.

$$H = -\sum_{i=1}^{n} p_i \log p_i \quad (1)$$

확률변수 x_i가 모두 다를 때, 확률 $p_i = \frac{1}{n}$로,

$$H = -\sum_{i=1}^{n} \frac{1}{n} log(\frac{1}{n}) = \sum_{i=1}^{n} \frac{1}{n} \log n = \log n \quad (2)$$

확률변수 x_i가 모두 같을 때, 해당 변수의 확률 p_i =1 이고 나머지 확률변수는 모두 0이다.

$$H = -\sum_{i=1}^{n} p_i \log p_i = -\frac{1}{1} log \frac{1}{1} = 0 \quad (3)$$

따라서 엔트로피가 가질 수 있는 값의 범위는 다음과 같다.

$$0 \leq H \leq \log n \quad (4)$$

(b) 엔트로피를 이용한 기존 연구 현황

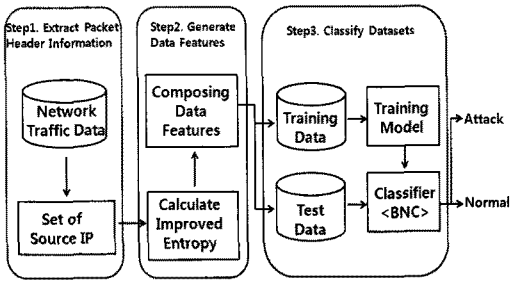
현재 엔트로피를 이용한 다양한 분산 서비스 거부 공격 탐지 연구들이 진행되고 있다. 대표적인 연구로 Laura Feinstein[5]은 엔트로피를 계산하기 위해 슬라이딩 윈도우의 크기를 패킷 10,000개로 고정하였다. 또한 헤더 정보 중 프로토콜의 플래그 분포를 비교하기 위해 카이스퀘어 검정을 사용하였는데 이것은 TCP SYN Flooding이나 ICMP Flooding과 같이 특정 프로토콜을 이용하는 공격을 효과적으로 탐지할 수 있으나 HTTP Get Flooding 공격과 같은 서버 자원 고갈 공격에는 제대로 대응하기 힘들고 적절히 용인 가능한 엔트로피에 근접하게 공격하면 탐지가 어렵다는 단점이 있다. Liyang Li[6]는 정상 엔트로피의 두배가 되면 비정상이라 판단하고 엔트로피를 누적하여 계산하는 기법을 제안하였다. Liyang Li 기법의 경우 평소에는 엔트로피를 누적하지 않고 비정상이라 판단될 때부터 누적함으로써 공격이 끝난 후 트래픽을 분석할 때 공격의 세기를 판단할 수 있는 척도로 사용할 수 있지만 소수 봇넷으로부터의 스푸핑 되지 않은 대량의 트래픽은 오히려 정상보다 엔트로피가 낮으므로 탐지가 불가능하다는 단점이 있다. Keunsoo Lee[7]는 근원지 IP주소 및 포트 번호의 엔트로피, 목적지 IP주소와 포트 번호의 엔트로피, 패킷 타입의 (ICMP, UDP, TCP, SYN) 엔트로피, 패킷 타입의 발생률 그리고 패킷의 개수를 탐지 파라메터로 이용하였으며 클러스터 분석 기법을 사용하였다. 클러스터 사이의 비유사성을 측정하기 위해 유클리디안 거리를 사용하였고 Ward의 최소 변량 방법[8]을 이용해 분류했다. 또한 클러스터의 개수를 결정하기 위해 Cubic Clustering Criterion기법[9]을 사용하였다. Keunsoo Lee가 제안하는 기법의 경우 정상 상태부터 공격까지 단계별로 엔트로피를 계산하고 클러스터링 함으로써 각 단계를 군집화하였으나 2000 DARPA Datasets에 대해서만 적용하였기 때문에 현재의 다양하고 변화된 공격에 대해 잘 맞는지 알 수 없다. Ping DU[10]는 패킷 크기 엔트로피 기반 탐지 기법을 제안하였다. 패킷에 대한 슬라이딩 윈도우의 크기를 실험을 통해 200이상으로 설정한 후 엔트로피가 높고 편차가 낮으면 서비스 거부 공격, 편차가 크고 근원지 IP주소가 목적지 IP주소보다 월등히 많다면 분산 서비스 거부 공격이며 근원지 IP주소가 스푸핑 된 것이라고 탐지하는 기법이다. Ping Du는 2000 DARPA Dataset과 같은 로컬 네트워크 뿐만 아니라 2004년도에 수집된 백본 네트워크 SINET[11] 데이터셋에도 적용함으로써 다양하게 사용할 수 있음을 보였지만 IP가 스푸핑된 공격에 한정하여 탐지하였다.

III. 제안 기법

본 논문에서 제안하는 방법은 다음과 같이 구성된다. (1)각각의 네트워크 패킷에서 근원지 IP주소를 추출한다. (2)시간적으로 들어오는 패킷의 근원지 IP주소에 대하여 반복적으로 슬라이딩 윈도우 기법을 사용해 엔트로피를 계산한다. 엔트로피를 계산하기 위해 기존의 일반 엔트로피 수식이 아닌 미세한 값도 나타낼 수 있는 향상된 엔트로피 수식을 이용한다. (3)얻어진 엔트로피를 특징 값으로 하여 데이터를 구성한다. (4)구성된 데이터를 분류기의 입력으로 사용한다. 이 데이터는 분류기의 모수를 형성하기 위한 학습 데이터와, 분류기의 성능을 검증하거나 분산 서비스 거부 공격 여부를 판단하기 위한 검증 데이터로 사용된다. 전체 시스템 구조도는 [그림 1]과 같다.

3.1 데이터 전처리

제안하는 기법에서의 데이터 전처리는 네트워크 패



[그림 1] 분산 서비스 거부 공격 탐지 시스템 구조

킷에서 정보를 추출하는 부분과 엔트로피를 계산하는 부분으로 나눠진다.

3.1.1 네트워크 패킷 전처리

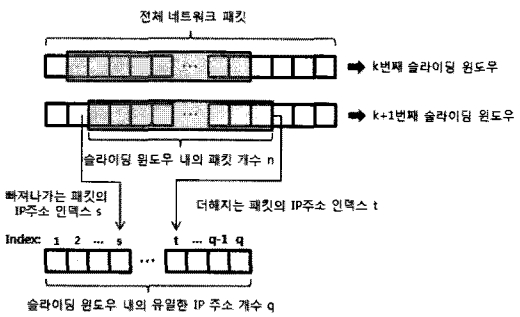
네트워크 패킷은 옵션을 제외한 20바이트의 헤더와 가변 길이 데이터로 구성된다. 헤더는 근원지 IP 주소, 목적지 IP 주소, 포트번호, TTL 등 경로 설정과 전달에 필요한 다양한 정보를 포함한다. 제안기법에서는 분산 서비스 거부 공격의 탐지에 이용하기 위해 32비트로 구성된 근원지 IP 주소를 이용한다.

3.1.2 엔트로피 계산

효과적인 분산 서비스 거부 공격의 탐지를 위하여 확률적인 접근 방식의 엔트로피를 사용한다. 향상된 엔트로피 수식을 적용하고 슬라이딩 윈도우를 이용하여 효율적으로 엔트로피를 계산한다.

(a) 효율적인 엔트로피 계산 기법

공격 탐지를 위한 특징 값을 구성할 때 엔트로피를 매번 계산해야 한다는 것은 상당히 비효율적이며 계산 시간도 오래 걸린다. 따라서 효율적으로 엔트로피를 계산하기 위해 일정 패킷 구간을 정하여 슬라이딩 윈도우 기법을 이용한다. 이것을 아래 [그림 2]에 나타내었다.



[그림 2] 엔트로피를 계산하기 위한 슬라이딩 윈도우

[표 1] 엔트로피 계산에서 사용되는 기호

기호	설 명
n	슬라이딩 윈도우 내의 패킷 갯수
q	유일한 IP주소의 개수
s	빠져나가는 패킷의 IP주소 인덱스
p_s	IP 주소 인덱스 s의 확률
t	더해지는 패킷의 IP주소 인덱스
p_t	IP 주소 인덱스 t의 확률

슬라이딩 윈도우 내의 패킷 개수를 n이라 하고 유일한 IP주소의 개수를 q라 할 때(단, $1 \leq q \leq n$), 확률 변수 x_i의 개수를 n_i라하자. 그러면, $n = \sum_{i=1}^{q} n_i$, $p_i = \frac{n_i}{n}$가 되고 k번째 슬라이딩 윈도우의 엔트로피를 다음과 같이 나타낼 수 있다.

$$H_k = -\sum_{i=1}^{q} p_i \log p_i \tag{5}$$

슬라이딩 윈도우가 이동할 때 두 가지 경우를 고려해야한다. 먼저 기존의 패킷이 빠져나가는 경우라 가정하자. 변동되는 IP주소의 인덱스가 s일 때의 확률을 p_s라 하면 k번째 슬라이딩 윈도우의 엔트로피는 다음과 같이 표현된다.

$$H_k = -\sum_{\substack{i=1,\\ i \neq s}}^{q} p_i \log p_i - p_s \log p_s \tag{6}$$

이 때 $k+1$번째 슬라이딩 윈도우의 엔트로피는 아래와 같다. 단 $p_s = \frac{n_s}{n}$이다.

$$H_{k+1} = H_k + p_s \log p_s - (p_s - \frac{1}{n}) \log (p_s - \frac{1}{n}) \tag{7}$$

다음으로 새로운 패킷이 더해지는 경우라 가정하자. 변화되는 IP주소의 인덱스가 t일 때의 확률을 p_t라 하면 k번째 슬라이딩 윈도우의 엔트로피는 다음과 같이 표현된다.

$$H_k = -\sum_{\substack{i=1,\\ i \neq t}}^{q} p_i \log p_i - p_t \log p_t \tag{8}$$

이 때 $k+1$번째 슬라이딩 윈도우의 엔트로피는 다음과 같다. 단, $p_t = \frac{n_t}{n}$이다.

$$H_{k+1} = H_k + p_t \log p_t - (p_t + \frac{1}{n}) \log (p_t + \frac{1}{n}) \tag{9}$$

이와 같이 다음 슬라이드 윈도우의 엔트로피를 계산할 때 이전 슬라이드 윈도우의 엔트로피를 이용하여 재귀적으로 계산이 가능하다. 이것을 하나의 수식으로 나타내면 다음과 같다.

$$\begin{aligned} H_{k+1} = & -\sum_{\substack{i=1,\\ i \neq s,\\ i \neq k}}^{q} p_i \log p_i - p_s \log p_s - p_t \log p_t \\ & + p_s \log p_s - (p_s - \frac{1}{n}) \log (p_s - \frac{1}{n}) \\ & + p_t \log p_t - (p_t + \frac{1}{n}) \log (p_t + \frac{1}{n}) \end{aligned} \tag{10}$$

단, 슬라이딩 윈도우에서 빠져나가는 패킷과 새로 슬라이딩 윈도우로 더해지는 패킷이 같을 경우에는 엔트로피의 차이가 없다. 즉 이 경우에는 $H_{k+1} = H_k$이다.

(b) 향상된 엔트로피 수식

최근 다수의 봇 PC를 이용해 소량의 트래픽을 지속적으로 전달하는 분산 서비스 거부 공격이 많이 발생하고 있다. 또한 컴퓨터의 성능과 네트워크 환경이 발달하여 소수의 봇 PC만으로도 충분히 공격이 가능하다. 이와 같은 경우, 일반 엔트로피 수식으로는 변화량이 크지 않아 탐지하기가 어렵다. 따라서 미세한 변화도 효과적으로 나타낼 수 있는 향상된 엔트로피 수식을 제안한다. 기존의 엔트로피를 계산하는 방법이 수식(1)과 같다면 제안하는 기법에서는 $p_i \log p_i$대신에 $\log p_i$를 사용한다. 이에 따라서 엔트로피 수식은 다음과 같이 변화된다.

$$H' = -\sum_{i=1}^{n} \log p_i \tag{11}$$

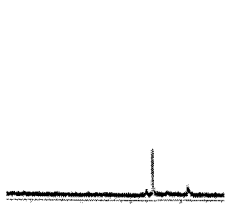
향상된 엔트로피 수식에서, 확률변수 x_i가 모두 다를 때는 확률 $p_i = \frac{1}{n}$로 다음 수식과 같이 나타낼 수 있다.

$$H' = -\sum_{i=1}^{n} \log(\frac{1}{n}) = \sum_{i=1}^{n} \log n = n\log n \tag{12}$$

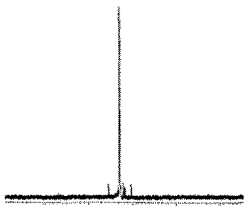
확률변수 x_i가 모두 같을 때는 수식 (3)과 마찬가지로 0의 값을 갖는다. 따라서 엔트로피의 범위는 다음과 같다.

$$0 \le H' \le n\log n \tag{13}$$

일반적인 엔트로피 수식을 이용할 때보다 값이 n배로 커지기 때문에 이 값 또한 n배로 증가하게 된다. 따라서 미세한 변화를 효과적으로 나타내고 탐지할 수 있다. [그림 3]과 [그림 4]에서 일반적인 엔트로피 수식과 향상된 수식을 이용한 엔트로피 값을 그림으로 나타내었다.



[그림 3] 일반적인 엔트로피 기법



[그림 4] 향상된 엔트로피 기법

3.2 엔트로피를 이용한 특징 생성

효율적인 엔트로피 계산 방법을 통해 얻어진 엔트로피를 데이터 x_N의 특징 값으로 이용하여 분류한다. x_N은 분산 서비스 거부 공격의 특성을 고려하여 다양한 엔트로피 집합으로 구성하며 아래와 같이 표현할 수 있다. 단, l은 차원을 의미한다.

$$x_1 = \{H_1, H_2, H_3, \cdots, H_l\} \in R^l \tag{14}$$
$$x_2 = \{H_2, H_3, H_4, \cdots, H_{l+1}\} \in R^l$$
$$\vdots$$
$$x_N = \{H_N, H_{N+1}, H_{N+2}, \cdots, H_{N+l-1}\} \in R^l$$

분산 서비스 거부 공격은 순간의 네트워크 패킷 정보를 이용해서는 탐지하기 어렵다. 따라서 시간의 연속성을 알아보기 위해 특징 값을 다양하게 적용할 필요가 있다.

3.3 데이터 분류 및 평가 방법

전처리 과정을 통해 얻어진 데이터를 베이지안 네트워크 분류기(Bayesian Network Classifier)[12]를 이용하여 분류하였다. 일반적으로 C-SVM (Support Vector Machine)[13]이나 Kernel-LDA(Linear Discriminant Analysis)[14]가 비선형 변형을 하기 때문에 베이지안 네트워크 분류기보다 훨씬 우수한 성능을 발휘한다. 하지만 본 논문에서는 우수한 성능을 보여주는 것이 목적이 아니고 데이터의 특성에 대한 연구가 주목적이기 때문에 가장 수학적인 분류기로써 데이터의 특징을 확률적으로 잘 보여주는 베이지안 네트워크 분류기를 사용하였다. 베이지안 네트워크 분류기는 이미 알고 있는 지식을 사전지식으로 사용하여 학습 목표인 조건부 확률을 계산하는 베이지안 정리에 기초를 두고 있으며 침입 탐지 분야에도 많이 사용되고 있다[15][16][17]. 본 논문에서 사용되는 베이지안 네트워크 분류기에서는 베이지안 네트워크를 통해 정상 데이터와 공격 데이터의 결합 확률 분포표를 구하고 이것을 이용하여 베이지안 정리를 통해 분류한다.베이지안 네트워크는 불확실한 상황 하에서 지식을 표현하고 결론을 추론하고자 할 때 유용하게 사용되며 광범위한 데이터를 변수간의 관계에 따라 그래프로 표시함으로써 데이터의 특성을 이

해할 수 있게 해준다. 베이지안 네트워크는 변수를 노드로 표현하며 노드와 노드들 간의 인과관계를 나타내는 비순환성의 방향성이 있는 그래프(Directed Acyclic Graph)이다. 각 노드는 조건부 확률을 나타내는 확률 테이블을 가지고 각 연결선의 강도를 모델화하며 두 노드 사이에 연결선이 없다는 것은 서로 독립이라는 의미로 해석된다. 엔트로피 집합에 대해 데이터 특징 값으로 구성한 $x=\{H_1,H_2,H_3, \cdots ,H_l\}$의 베이지안 네트워크는 x의 엔트로피 특징 값 간의 종속적 조건들을 정해주는 네트워크 구조와 엔트로피 특징 값에 대한 주변 확률 P로 이루어져 있다. 각 노드를 변수로 하고 $P_a(x)$를 변수 x의 부모 노드로 표시하면 데이터의 분포를 네트워크 구조에 따른 결합 확률 분포로 나타낼 수 있으며 다음 식과 같다.

$$P(x_1,x_2,\ldots,x_N)=\Pi_{i=1}^{n}P(x_i|p_a(x_i)) \quad (15)$$

여기서 $P_a(x_i)$는 x_i의 부모 노드를 나타내고 A→B의 그래프 구조를 가질 때 노드 A는 노드 B의 부모 노드가 되며 다음 식과 같이 표현된다.

$$P(x_1,x_2,x_3,x_4,x_5)= \quad (16)$$

$$P(x_1)P(x_2|x_1)P(x_3|x_1)P(x_4|x_2x_3)P(x_5|x_4)$$

수집된 정상 데이터를 w_1, 공격 데이터를 w_2라 하고 베이지안 네트워크를 이용한 x의 확률 분포 함수를 $P(x)$라 할 때, 특정 패턴 w_i의 발생 가능 확률은 $P(w_i)$이며 w_i에서 테스트 데이터 x가 관측될 조건부 확률은 $P(w_i|x)$이다. 이것을 이용하여 베이지안 정리를 나타내면 다음과 같다.

$$P(w_i|x)=\frac{P(x|w_i)P(w_i)}{P(x)} \quad (17)$$

이때, $P(x)$는 다음 수식과 같다.

$$P(x)=\sum_{i=1}^{n}p(x|w_i)p(w_i) \quad (18)$$

관측된 값 x가 변함에 따라 사전 확률 $P(w_i)$가 사후 확률 $P(w_i|x)$로 변화되는 것을 확인할 수 있으며 다음 식을 통하여 데이터를 분류할 수 있다.

$$P(x|w_i)P(w_i)>P(x|w_j)P(w_j),\ (i\neq j) \quad (19)$$

이와 같이 베이지안 네트워크 분류기를 이용하여 정상 트래픽과 공격 트래픽을 트레이닝 시킨 후 새로운 환경의 트래픽을 테스트함으로써 해당 트래픽이 정상인지 공격인지를 분류해낼 수 있다.

IV. 실험 및 성능 평가

4.1 데이터셋 분석

제안하는 기법에 대한 실험 및 성능 평가를 위해 두 가지 데이터셋을 이용하였다. 먼저 DARPA 2000 Datasets[18]을 이용하였다. DARPA 2000 Datasets은 2000년 3월 수집되었으며 모든 공격 패킷이 스푸핑되었다. 현재의 분산 서비스 거부 공격은 스푸핑 없이 봇 PC를 이용해 공격하는 것이 일반적이며 다양하고 복잡하게 변화되었다. 따라서 DARPA 2000 Datasets은 현실의 공격을 제대로 반영하지 못한다고 할 수 있으며 제안하는 방법이 현실에 맞는 효과적인 방법인가를 분석하기 위해 일반 대학 내에서 폐쇄망을 구축하여 분산 서비스 거부 공격 데이터셋을 수집하였다.

4.1.1 DARPA 2000 Datasets

본 논문에서 사용하고자 하는 데이터셋은 2000 DARPA Intrusion Detection Scenario Specific Data Sets이다. DARPA 2000 Datasets은 분산 서비스 거부 공격에 대한 트래픽을 담고 있으며 여러 탐지 기법들에 의해 사용되어 왔다[19][20][21]. 공격은 Mstream 소프트웨어를 통해 이루어졌다. 공격 시나리오에서 세션은 크게 다섯 단계로 나누어진다. 첫 번째 단계에서는 호스트가 네트워크에서 동작하고 있는지를 알아보기 위해 IPsweep을 이용하였다. 원격지에서 Air Force Base의 네트워크 서브넷 내에 존재하는 모든 가능한 주소에 ICMP ping 패킷을 보내어 그 응답을 기다림으로써 호스트를 탐색하였다. 두 번째 단계에서는 특정 포트에 ping 패킷이 응답하는 가를 통해 첫 번째 단계에서 탐색된 호스트 중 솔라리스 호스트에서 취약점이 존재하는 sadmind 데몬을 운영하고 있는 호스트가 존재하는가를 탐색하였다. 세 번째 단계에서는 sadmind 데몬을 운영하는 호스트에 원격 버퍼 오버플로우 공격을 통해 침입하였다. 네 번째 단계에서는 성공적으로 침입이 이루어진 호스트에 분산 서비스 거부 공격을 하기 위한 trojan mstream 소프트웨어를 설치하였다. 마지막으로 다섯 번째 단계에서 공격자는 분산 서비스 거부

공격을 수행하였다.

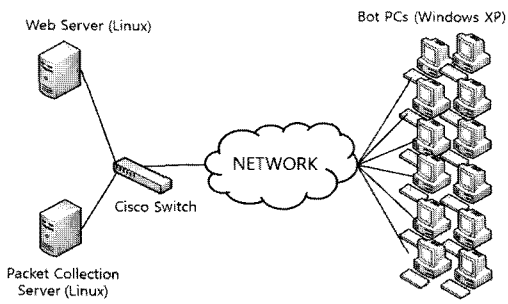
4.1.2 분산 서비스 거부 공격 데이터셋 구성

일반 대학 내의 폐쇄된 네트워크에서 봇넷을 구성하여 분산 서비스 거부 공격 트래픽을 수집하였다. 봇넷은 총 30대의 PC로 구성하였으며 공격 대상인 웹 서버와 패킷 수집용 서버를 Cisco 스위치에 연결하였고, 포트미러링 기법을 이용하여 트래픽을 수집하였다. 봇 PC의 운영체제는 Windows XP 이며 웹 서버 및 패킷 수집용 서버는 Linux 운영체제를 이용하였다. 실험 구성을 [그림 5]에 나타내었다.

공격은 Netbot Attacker 프로그램을 사용하였으며 3가지 분류로 나누고 분류에 따라 대표적인 공격 기법을 이용하였다. 세션 고갈 공격으로 SYN Flooding 공격을 사용하였고 대역폭 고갈 공격으로는 ICMP Flooding 공격, UDP Flooding 공격을 이용하였다. 마지막으로 서버자원 고갈 공격으로 HTTP Get Flooding 공격을 이용하여 총 4가지 공격에 대해 실험하였다. 또한 적은 수의 봇 PC를 이용한 공격에도 효과적으로 대응할 수 있는가를 분석하기 위해 5대의 봇 PC로 공격을 수행하였으며 30대의 봇 PC를 이용한 공격만큼 효과적이진 않지만 대량의 트래픽이 발생하며 웹서버로의 접근이 간헐적으로 불가능해짐을 확인할 수 있었다.

4.2 베이지안 네트워크 분류기를 이용한 분류

제안하는 방법이 효과적인가를 검증하기 위해 DARPA 2000 Datasets과 직접 수집한 분산 서비스 거부 공격 데이터셋에 대해 하나의 엔트로피를 계산하기 위한 슬라이딩 윈도우의 패킷 개수 n을 50부터 50개 간격으로 다양화하고, 특징 값의 수 l을 1, 2, 3, 4 등으로 다양화하여 베이지안 네트워크 분류기로 분류하였다. 분류하기 위해서 자바로 작성된 패턴 인식 프로그램인 Weka 3.6[22]을 이용하였다.



[그림 5] 분산 서비스 거부 공격 트래픽 수집 실험 구성

4.2.1 DARPA 2000 Datasets 분류

DARPA 2000 Datasets을 정상과 공격으로 분류하기 위해 첫 번째 단계부터 공격하기 전의 네 번째 단계까지를 정상 상태라 정의하고 각 단계 트래픽의 앞쪽 절반에 대한 합을 트레이닝 데이터, 뒤쪽 절반에 대한 합을 테스트 데이터로 구성하였다. 마찬가지로 공격이 발생했던 다섯 번째 단계 트래픽의 앞쪽 절반을 트레이닝 데이터, 뒤쪽 절반을 테스트 데이터로 사용하였다. 트레이닝 데이터와 테스트 데이터의 특징 값은 각각 정상 데이터의 경우 472개, 공격 데이터의 경우 37,139개이며 정상 데이터에서 공격 데이터보다 상대적으로 적은 트래픽이 발생했기 때문에 공격 데이터에서 랜덤으로 데이터를 선택해 정상 데이터와 비율을 맞추어 분류하였으며 분류 결과는 다음 [표 2]와 같다. 행은 특징 값 개수 l을 나타내며 열은 패킷 윈도우 n의 개수를 나타낸다. 테스트 결과 n을 100으로 하고, l을 2로 했을 때 99.41%로 가장 높은 분류율을 보였다.

다음 [표 3]은 99.41%에 해당하는 분류 행렬(Confusion Matrix)을 나타낸 것이고 [그림 6]에서는 이 때의 엔트로피 분포를 나타내었다.

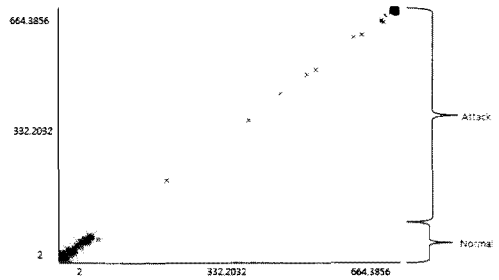
정상 상태에서는 0에 가까운 낮은 엔트로피가 측정되었으며 공격 상태에서는 모든 패킷의 근원지 IP주소가 스푸핑 되었기 때문에 엔트로피가 높게 측정되었다. 테스트 데이터셋에서는 공격이 종료되는 트래픽까지 포함되었기에 서서히 떨어지는 엔트로피가 측정되었다.

[표 2] DARPA 2000 Datasets 분류율

	50	100	150	200
1	99.11%	99.17%	86.89%	82.61%
2	99.40%	99.41%	87.14%	82.91%
3	99.33%	88.95%	86.62%	83.56%
4	93.60%	98.81%	98.75%	98.59%

[표 3] 탐지율 99.41%의 분류 행렬

TP Rate	FP Rate	
1	0.011	정상
0.989	0	공격



〔그림 6〕 분류율 99.41%의 엔트로피 분포

4.2.2 분산 서비스 거부 공격 데이터셋 분류

직접 실험을 통해 구성한 분산 서비스 거부 공격 데이터셋을 분류하기 위해 공격을 시작하기 전 10분간 수집한 트래픽을 정상 행위의 트레이닝 데이터로, 추후 10분간 수집한 트래픽을 정상 행위의 테스트 데이터로 사용하였다. 또한, 30대의 봇 PC를 이용한 4가지 공격의 앞쪽 절반 트래픽을 합하여 트레이닝 데이터로 사용하고 뒤쪽 절반의 트래픽을 합하여 테스트 데이터로 사용하고 분류하였다. 마지막으로 적은 수의 봇을 이용한 공격도 적절히 탐지되는가를 알아보기 위해 5대의 봇 PC를 이용한 공격의 트래픽에 대해서도 마찬가지로 반복하여 적용하고 분류하였다. 트레이닝 데이터와 테스트 데이터의 특징 값은 각각 정상 데이터의 수 6,692개, 30대의 봇 PC를 이용한 공격 데이터의 수는 606,791개이며 5대의 봇 PC를 이용한 공격 데이터의 수는 596,658이다. 정상 데이터와 공격 데이터의 수가 약 10배 정도 차이나기 때문에 공격 데이터에서 랜덤으로 선택하여 비율을 맞춘 후 분류하였다. 다음의 〔표 4〕는 30대의 봇 PC을 이용한 공격 데이터를 분류한 결과이다. n이 100이고 l이 2일 때 가장 분류율이 높았으며, 그 때의 분류율은 92.67%이다.

〔표 5〕에서 92.67%의 분류 행렬을 나타내었다.

아래의 〔그림 7〕에서 분류율이 92.67%일 때의 엔트로피 분포를 확인할 수 있다. 정상 행위는 일정하게 분포하였지만 공격 시 30대의 봇 PC로부터 스푸핑되지 않은 트래픽이 대량으로 발생해 공격 트래픽에 대한 엔트로피 값이 다수의 사용자가 이용하는 정상 상태 보다 낮게 측정되었다.

다음의 〔표 6〕에서는 5대의 봇 PC를 이용한 공격 데이터셋을 분류한 결과이다. 대체적으로 잘 분류되었으며 가장 높을 때 99.98%로 높은 분류율을 보였다.

n이 100이고 l이 2일 때의 분류 행렬을 〔표 7〕에 나타내었다. 공격 행위는 모두 공격으로 분류되었으며 정상 행위 6,692개 중 단 2개만 공격으로 오분류 되었다.

〔그림 8〕에서는 분류율이 99.98%일 때의 엔트로피 분포를 나타내었으며 낮은 엔트로피를 보이는 공격 상태와 고르게 분포되어 있는 정상 상태의 엔트로피를 확인할 수 있다. 정상 행위는 일정하게 분포하며 공격 시 5대의 봇 PC로부터 스푸핑되지 않은 트래픽이 대량으로 일정하게 발생하여 공격 트래픽에 5대의 봇

〔표 4〕 30대의 봇을 이용한 공격 데이터셋 분류율

	50	100	150	200
1	71.58%	90.18%	79.70%	81.25%
2	71.62%	92.67%	80.23%	84.67%
3	71.11%	89.52%	80.16%	83.99%
4	71.42%	88.37%	79.64%	82.53%

〔표 5〕 탐지율 92.67%의 분류 행렬

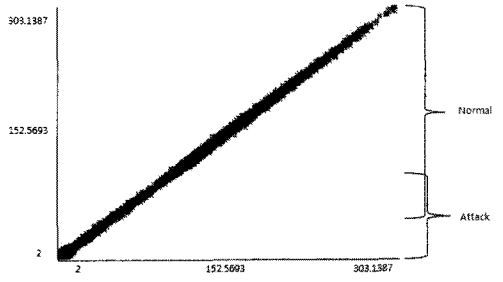
TP Rate	FP Rate	
0.977	0.123	정상
0.877	0.023	공격

〔표 6〕 5대의 봇을 이용한 공격 데이터셋 분류율

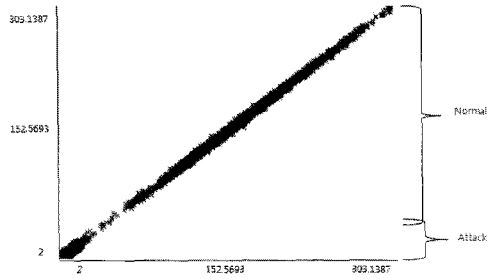
	50	100	150	200
1	98.26%	99.98%	96.32%	97.39%
2	98.42%	99.98%	99.07%	97.45%
3	98.15%	99.97%	96.47%	97.29%
4	98.37%	99.96%	96.33%	97.43%

〔표 7〕 분류율 99.98%의 분류 행렬

TP Rate	FP Rate	
1	0	정상
1	0	공격



〔그림 7〕 분류율 92.67%의 엔트로피 분포



〔그림 8〕 분류율 99.98%의 엔트로피 분포

PC에 해당하는 IP주소가 편중되어 분포하므로 공격에 대한 엔트로피가 정상 상태 보다 낮은 값을 보였다.

4.2.3 베이지안 네트워크 분류기 분류 결과

실험에서 사용했던 두 개의 데이터셋에서는 하나의 엔트로피를 계산하기 위한 패킷 윈도우 개수 n을 100으로 하고 데이터 분류를 위한 특징 값 l을 2개로 하였을 때 평균적으로 가장 높은 분류율을 보였으며 분류율은 97.35%이다. 반면에 n이 200이고 l이 1일 때 87.09%로 가장 낮은 분류율을 보였다. 결과 값을 분석해보면 n이 100일 때 분류율이 최대이며 150일 때보다 200일 때가 분류율이 높지만, 200을 초과하여 커지면서 분류율이 점차적으로 낮아졌다. 특징 값의 경우 l이 1개일 때보다 2개일 때 분류율이 높았고 3개 이상일 경우 분류율이 점차 낮아지는 것을 확인할 수 있었다. DARPA 2000 Datasets에서는 특징 값 l이 4개일 때 분류율이 높게 나타났지만 4개를 초과하여 개수가 많아질수록 점차적으로 분류율이 낮아지며 n을 100으로, l을 2개로 하였을 때 최대의 분류율을 보였다. DARPA 2000 Datasets은 IP주소가 스푸핑 되었기 때문에 공격 상태일 때 엔트로피가 높게 측정되었지만 직접 실험을 통해 얻은 분산 서비스 거부 공격 데이터셋에서는 IP주소를 스푸핑하지 않았기 때문에 공격 상태일 때 엔트로피가 낮게 측정되었다. 요즘은 익명의 봇 PC를 이용해 공격 명령을 내리기 때문에 스푸핑을 하지 않고 공격이 이루어지며, 실험을 통해 제안하는 방법이 스푸핑된 공격 뿐만아니라 스푸핑하지 않은 공격도 효과적으로 분류함을 보였다. 실험 결과는 제안하는 기법이 분산 서비스 거부 공격 탐지에 효과적으로 적용될 수 있음을 보여준다.

V. 결 론

본 논문에서는 엔트로피를 이용하여 효과적으로 분산 서비스 거부 공격을 탐지할 수 있는 특징 생성 기법을 제안하였다. Laura Feinstein[5], Keunsoo Lee[7]와 같은 기존의 연구에서는 단일 엔트로피를 이용한 탐지 및 분류 연구가 진행되었지만 제안하는 방법에서는 특징 값으로 다양한 개수의 엔트로피를 사용함으로써 효과적으로 분산 서비스 거부 공격을 탐지할 수 있음을 증명하였다. 또한 Liyang Li[6], Ping DU[10]는 누적합 기법 및 패킷 크기 엔트로피 기반 탐지 기법을 사용하였지만 점점 복잡해지고 정교해지는 분산 서비스 거부 공격을 탐지하는데 효과적이지 못하다. 따라서 본 논문에서는 확률과 시간의 연속성, IP의 랜덤성을 엔트로피를 이용하여 표현하고, 효율적인 엔트로피 계산 기법과 향상된 엔트로피 수식을 이용하여 특징 값을 생성함으로써 효과적으로 분산 서비스 거부 공격 탐지에 적용할 수 있는 기법을 제안하였다. 또한 DARPA 2000 Datasets이 현실과 맞지 않기 때문에 실제 폐쇄망을 구축하고 공격을 통한 데이터셋을 수집하여 테스트함으로써, 현실의 공격에도 효과적으로 대응할 수 있음을 보였다. 향후 분산 서비스 거부 공격 뿐만 아니라 다양한 공격에 적용하여 탐지할 수 있는 연구가 필요할 것이다.

참고문헌

[1] GJ Park, "Internet security incident trends and analysis," Korea Internet Security Center, Dec. 2009.

[2] Mindi McDowell, "Understanding denial of service attacks," US-CERT, Cyber Security ST04-015, Nov. 2004.

[3] MH Lee and CH Ryu, "Internet and security issue," National Internet Development Agency of Korea, vol. 1, Sep. 2009.

[4] C.E. Shannon and W. Weaver, "The mathematical theory of communication," University of Illinois Press, 1963.

[5] Laura Feinstein and Dan Schnackenberg, "Statistical approaches to DDoS attack detection and Response," IEEE Computer Society, 2003.

[6] Liying Li and Jianying Zhou, "DDoS

attack detection algorithms based on entropy computing," ICICS Electronic Edition, pp. 452-466, 2007.

[7] KS Lee, JH kim, and KH Kwon, "DDoS attack detection method using cluster analysis," ScienceDirect Expert Systems with Application 34, 2008.

[8] J.H. Ward Jr., "Hierarchical grouping to optimize an objective function," Journal of the American Statistical Association, vol. 58, pp. 236-244, Mar. 1963.

[9] SAS Institute Inc, "Cubic clustering criterion," SAS Technical Report A-108, 56p, Nov. 1983.

[10] Ping DU and SHunji ABE, "IP packet size entropy-based scheme for detection of DoS/DDoS attacks," IEICE Trans. INF. & SYST., vol. E91-D, no. 5, May 2008.

[11] http://www.sinet.ad.jp/what-is-the-science-information-network-sinet

[12] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Machine Learning 29, pp. 131-163, Nov. 1997.

[13] T. Joachims, "Making large-scale support vector machine learning practical, Advances in kernel methods: support vector learning," MIT Press, pp169-184, 1999.

[14] CH Park, "Efficient linear and nonlinear feature extraction and its application to fingerprint classification," University of Minnesota, 129p, 2004.

[15] MH Chung, JI Cho, SY Chae and JS Moon, "An efficient method for detecting denial of service attacks using kernel based data," Journal of the Korea Institute of Information Security and Cryptology, vol. 19, no. 1, pp. 71-79, Feb. 2009.

[16] B Cha and D Lee, "Network-based anomaly intrusion detection improvement by bayesian network and indirect relation," Lecture Notes in Computer Science, pp. 141-148, Sep. 2007.

[17] S Benferhat and K Tabia, "Novel and anomalous behavior detection using bayesian network classifiers," International Conference on Security and Cryptography, 2008.

[18] MIT/LL 2000 DARPA Intrusion Detection Scenario Specific Data Sets, http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/2000data.html

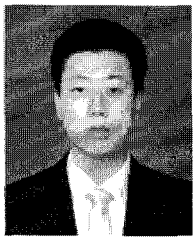
[19] CK Han and HK Choi, "An anomalous event detection system based on information theory," Korean Institute of Information Scientists and Engineers, Information Communicaion vol. 36-3, Jun. 2009.

[20] TH Kim and DS Kim, "Detecting DDoS attacks using dispersible traffic matrix and weighted moving average," Advances in Information Security and Assurance, pp. 290-300, Jun. 2009.

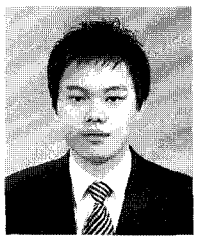
[21] S Terry Brugger and Jedadiah Chow, "An assessment of the DARPA IDS evaluation dataset using snort," UCDAVIS department of Computer Science, May. 2007.

[22] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann and H. Ian, "The WEKA data mining software," SIGKDD Explorations, vol. 11, no. 1, pp10-18, Jun. 2009.

〈著者紹介〉



김 태 훈 (Tae-hun Kim) 학생회원
2008년 2월: 고려대학교 전산학과 졸업
2008년 3월~현재: 고려대학교 정보경영공학전문대학원 정보경영공학과 석사과정
〈관심분야〉 네트워크 보안, 침입탐지, 데이터 마이닝



서 기 택 (Ki-taek Seo) 학생회원
2008년 2월: 강남대학교 컴퓨터미디어공학부 컴퓨터공학과 졸업
2008년 3월~현재: 고려대학교 정보경영공학전문대학원 정보경영공학과 석사과정
〈관심분야〉 네트워크 보안, 패턴인식, 스트리밍 암호화, 콘텐츠 보안

이 영 훈 (Young-hoon Lee) 학생회원
2009년 8월: 고려대학교 컴퓨터정보학과 졸업
2009년 9월~현재: 고려대학교 정보경영공학전문대학원 정보경영공학과 석사과정
〈관심분야〉 네트워크 보안, 시스템 보안, 데이터 마이닝

임 종 인 (Jong-in Lim) 종신회원
1986년 2월: 고려대학교 대학원 수학과 박사(암호학)
2000년 8월: 고려대학교 정보보호대학원/CIST 원장(센터장)
2004년 1월: 국가정보원 정보보호정책 자문위원
2005년 7월: 대통령 자문 전자정부 특별위원
2005년 12월: 국회 과기정위원회 정보통신 정책 자문위원
〈관심분야〉 정보보호기술, 정보보호정책, PET, 컴퓨터 포렌식

문 종 섭 (Jong-sub Moon) 종신회원
1981년~1985년: 금성 통신 연구소 연구원
1991년: Illinois Institute of technology 전산학 박사
1993년~현재: 고려대학교 전자 및 정보공학부 교수
〈관심분야〉 생체인식, 침입탐지, 네트워크 보안, 운영체제, 시스템 보안