제 26 회 한국정보처리학회 추계학술발표대회 논문집 제 13 권 제 2 호 (2006. 11)

협력적인 통계기반 탐지기법을 이용한 DDoS 공격 방어 시스템

송병학*, 홍충선**
경희대학교 컴퓨터공학과
e-mail : bhsong@networking.khu.ac.kr*, cshong@khu.ac.kr**

Cooperative defense mechanism using statistical detection method against DDoS attacks

Byung Hak Song, Choong Seon Hong
Dept. of Computer Engineering, Kyung Hee University

요 약

DDoS(Distributed Denial-of-Service) 공격은 인터넷 침해가운데 가장 위협적인 공격들 중 하나이며 이러한 공격을 실시간으로 방어하기 위한 연구는 활발히 이루어져 왔다. 그러나 분산된 공격 형태를 가지는 DDoS 공격을 특정 침입탐지대응시스템(IDRS)에서 탐지하고 대응하기에는 정확성과 속도 측면에서 한계를 가진다. 뿐만 아니라 기존의 탐지 방법을 피하거나 탐지를 지연시키는 지능화된 DDoS 공격에 대해서 높은 오탐지율을 가진다. 따라서 본 논문에서는 공격자(attacker)와 희생자(victim)의 네트워크에서 가장 가까운 탐지지점에 위치한 각각의 침입탐지대응시스템간의 상호 협력을 통한 효과적인 방어 시스템을 제안한다. 또한 탐지 임계값을 서서히 증가시키는 다양한 공격에 대해 조기에 탐지할 수 있는 메커니즘을 제안한다.

1. 서론

초고속 인터넷 인프라가 구축되면서 해킹 및 인터넷 침해에 대한 사고 사례가 매년 증가하고 있다. 특히 DDoS 공격은 단순한 공격기법과 어디서나 구할 수 있는 툴로 인해 초급해커(Script kiddie)도 얼마든지 공격할 수 있다. 2000 년 2 월에 야후, 아마존과 같은 인터넷 포털 사이트가 심각한 피해를 입었으며, 전 세계 인터넷 트래픽을 관장하는 미국 내 13 개의 루트 서버가 DDoS 공격을 받아 그 중 9 대가 일시적으로 정상 작동이 불가능 해지는 사례도 있었다[1]. 그리고 이러한 공격으로 인해 우리나라는 2003 년 1.25 인터넷대란을 겪기도 했다. 또한 최근 봇(Bot)[2]의 증가는 이러한 DDoS 공격과 같은 네트워크 자체에 위협을 주는 요소를 증가시키고 있다. 앞으로 유무선 통합 환경에서의 이러한 사고는 더욱 증가할 것으로 예상된다.

본 논문은 한국정보사회진흥원(NIA)의 지원으로 수행되었음.

분산된 공격 형태를 가지는 DDoS 공격을 효과적으로 방어하기 위해서는 기존의 독립적이고 고립된 방법으로는 탐지 및 대응이 어려우며 침입탐지대응시스템간의 상호협력이 필요하다[3]. 기존의 분산 침입탐지대응시스템은 특정 공격 형태나 시나리오에 있어서 좋은 성능을 보이지만, 다양하게 변형된 공격 형태[4]에 있어서 높은 오탐지율을 가지고 있다.

따라서 본 논문에서는 이러한 문제를 해결하기 위해 협력적인 침입방어 시스템의 탐지의 정확도를 향상시키고 수집된 정보를 바탕으로 효과적으로 대응할 수 있는 메커니즘을 제안한다. 뿐만 아니라 탐지 임계값을 서서히 증가시키는 다양한 공격에 대해 조기에 탐지할 수 있는 메커니즘을 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서 기존의 침입탐지 및 대응 기법과 세 가지 비정상 행위 탐지 알고리즘에 대해서 알아보고 3 장에서는 Source-end 와 Victim-end 간 협력을 통한 침입탐지 및 대응 기법을 제안한다. 마지막으로 4 장에서는 본 연구가 가지는 의의와 향후 연구 계획으로 마무리한다.

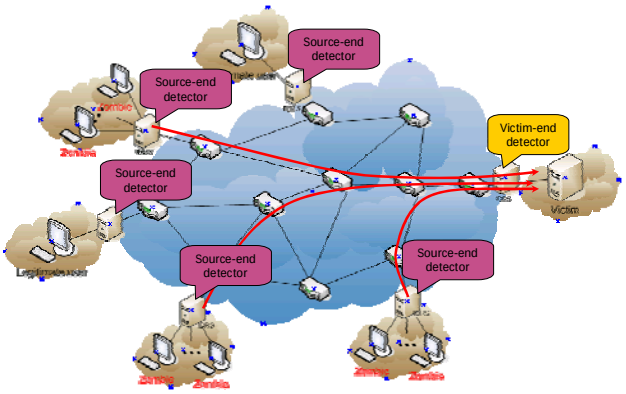
2. 관련 연구

2.1 침입탐지 및 대응 기법

침입탐지대응시스템은 탐지 위치에 따라서 Source-end 와 Victim-end 침입탐지대응시스템으로 구분된다.

일반적으로 침입탐지대응시스템은 희생자의 네트워크에 가장 가까운 에지(edge) 라우터에 위치한다[5]. 분산된 네트워크에서 들어오는 공격 트래픽을 통합적으로 분석해서 공격의 징후를 탐지하기 가장 좋은 위치이기 때문이다. 그러나 Victim-end 에서의 탐지대응시스템은 공격에 대응하기 전에 네트워크의 리소스를 고갈시키는 매우 높은 트래픽 양의 DDoS 공격 형태에 대해서 대응 시간이 오래 걸리고 대응 방법의 계산 복잡도가 높다는 단점이 있다.

반면에 Source-end 침입탐지대응시스템[6]은 패킷 필터링과 같이 공격에 대응하기에는 가장 효율적인 위치에 있지만 전체 DDoS 공격 트래픽의 일부만을 탐지의 판단 근거로 사용하기 때문에 Victim-end 침입탐지대응시스템보다 상대적으로 false positive 와 false negative 가 높다는 단점이 있다. 그림 1 은 Source-end 와 Victim-end 침입탐지대응시스템의 위치를 나타낸다.



(그림 1) Source-end 와 Victim-end 침입탐지대응시스템

2.2 통계기반 탐지 알고리즘

DDoS 공격은 일반적인 패킷으로 공격이 이루어지므로 합법적인 패킷과 구분하기 어려우며 탐지를 위해서는 탐지의 정확성과 복잡도가 동시에 고려되어야 한다[7]. 따라서 이를 탐지하기 위해서는 통계적인 방법을 사용하는 것이 가장 효율적이다.

통계적인 탐지 알고리즘에는 트래픽 볼륨(traffic volume), 패킷 속성값의 엔트로피(entropy) [8], 카이 제곱 (Chi-Square) 검증법[8]등이 사용되고 있다.

이 가운데 트래픽 볼륨 측정은 패킷이 이더넷 카드에 도착하는 시간을 계산하는 것이다.

$$T = \sum_{i=1}^{n} (PAT[i-1] - PAT[i])$$

위 공식은 각각의 패킷이 도착하는 시간을 측정하여 그룹단위로 패킷이 도착한 시간을 계산한다. 즉 100 개의 패킷을 한 그룹으로 설정을 했다면 100 개의 패킷이 이더넷 카드에 도착하는 시간을 측정하는 것이다. 트래픽 볼륨 값이 낮을 수록 이더넷 카드에 도착하는 패킷의 수가 증가했다는 것을 의미한다. 이것은 현재의 트래픽이 갑자기 증가했다는 것으로 이상 트래픽 발생 가능성을 암시한다.

다음 엔트로피 연산법은 어떠한 네트워크 속성값에 대한 임의성(randomness)을 계산한 뒤, 그 값의 평균의 변화량을 탐지하는 방법이다.

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i$$

위의 공식은 n 개의 속성 값에 대한 엔트로피 H 를 구하는 공식이다. 여기서 p_i 는 i 번째의 속성값이 선택될 확률을 나타낸다.

카이 제곱 검증법은 속성값에 대한 분산도를 측정하는 방법이다. 여기서는 기대값에 대한 분산도를 계산하여 그 값에 따라 비정상적인 속성값을 탐지할 수 있다. 이 방법의 구체적인 식은 다음과 같다.

$$x^2 = \sum_{i=1}^{B} \frac{(N_i - n_i)^2}{n_i} \quad , n_i = \frac{n}{B} \quad , n = total\ sample\ size$$

여기서 B 는 샘플 패킷들이 가질 수 있는 값들을 묶어놓은 binning 값이다.(ex. 패킷 길이는 0-64, 65-128, 129-255 로 binning 될 수 있다) N_i 는 N 개의 샘플 패킷에서 각각의 binning 범위에 속하는 패킷의 개수이고, n_i 는 일반적인 분포에서 binning 에 속하는 기대값이다.

3. 제안 사항

Source-end 탐지 기법과 Victim-end 탐지 기법 간 협력을 통해 기존의 Source-end 탐지 방법과 Victim-end 탐지 방법의 단점을 보완하고 탐지 및 대응 성능 향상 시키는 방법에 대해서 제안하다. 뿐만 아니라 지능화된 DDoS 공격 형태에 대한 탐지 메커니즘을 제안한다.

3.1 Source-end 탐지

Source-end 탐지는 크게 두 가지 모듈로 구성된다. 하나는 IP spoofing 검사 모듈이고 다른 하나는 DDoS 공격 탐지 모듈이다.

IP spoofing 검사 모듈은 outgoing 트래픽에 대해서 자신이 속한 서브넷이 아닌 소스 IP 주소를 가진 패킷을 필터링한다. 그래서 좀비(zombie) 에이전트는 공격 시 자신의 소스 IP 주소를 외부 네트워크의 IP 주소로 스푸핑(spoofing)한다면 차단된다.

DDoS 공격 탐지 모듈은 트래픽 볼륨, 소스 주소 엔트로피, 카이 제곱, 그리고 목적지 주소 엔트로피 값은 이용하여 계산한다. DDoS 공격 시 전체 트래픽에 대한 T 값은 감소하고 Victim 의 응답 패킷이 증가하므로 소스 주소 엔트로피 값은 감소하게 된다. 그리고 좀비 에이전트가 보내는 트래픽의 증가로 송신하는 소스주소 엔트로피와 카이 제곱 값은 증가한다. 따라서 $H_s(s)/H_s(r)$ 과 $x^2(s)$ 값은 급격히 증가한다. 그리고 목적지 주소 엔트로피는 소스 주소 엔트로피와 반대의 특성을 가지기 때문에 $H_d(s)/H_d(r)$ 값은 급격히 감소한다. 그리고 각각의 임계값은 최근 측정값에 가중치를 부여한 평균값과 분산값을 이용하여 동적으로

변화하는 트래픽에 대해 일정한 간격마다 임계값을 산출한다.

- 트래픽 볼륨

$$T < T_t$$

T : *traffic volume*

T_t : *threshold of traffic volume*

- 소스 주소 엔트로피

$$\frac{H_s(s)}{H_s(r)} > T_{H_s}$$

$H_s(s)$: *source address entropy of sent packets*

$H_s(r)$: *source address entropy of received packets*

T_{H_s} : *threshold of source address entropy*

- 카이 제곱

$$x^2(s) > T_{x^2}$$

$x^2(s)$: *chi – square of sent packets*

T_{x^2} :*threshold of chi–square*

- 목적지 주소 엔트로피

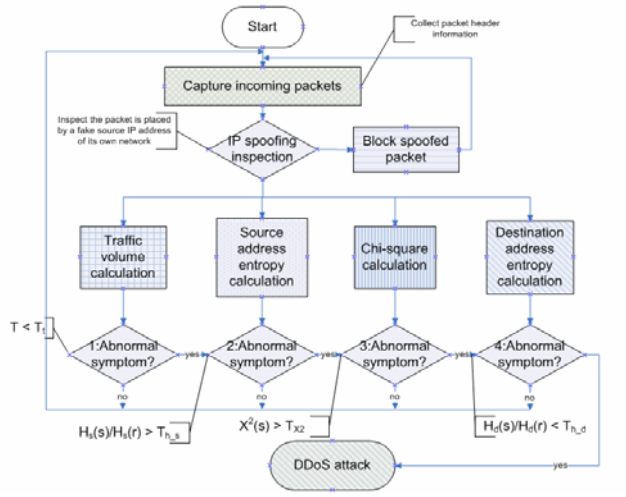
$$\frac{H_d(s)}{H_d(r)} < T_{H_d}$$

$H_s(s)$: *destination address entropy of sent packets*

$H_s(r)$: *destination address entropy of received packets*

T_{H_d} : *threshold of destination address entropy*

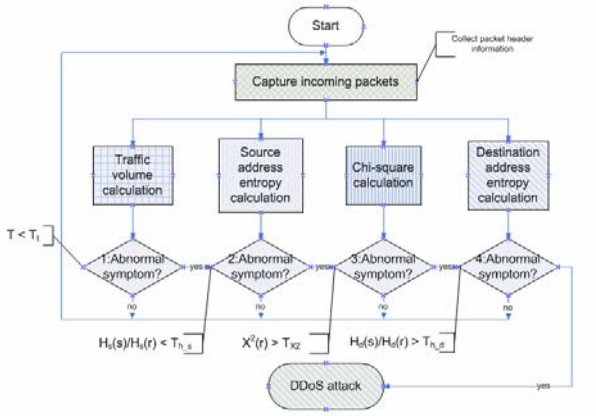
아래 그림은 source-end 탐지 기법의 과정을 나타낸 순서도이다.



(그림 2)Source-end 탐지 순서도

3.2 Victim-end 탐지

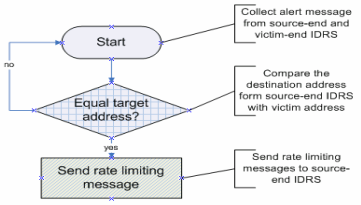
Victim-end 탐지 기법은 Source-end 탐지 기법의 DDoS 탐지 모듈과 같은 구성을 가진다. 그러나 DDoS 공격 시 소스 주소 엔트로피, 카이 제곱, 그리고 목적지 주소 엔트로피 각각의 비율 값은 Source-end 탐지의 수식과 반대의 특성을 가진다. 즉 $H_s(s)/H_s(r)$ 값은 정상범위보다 작아지고 T, $H_d(s)/H_d(r)$과 $x^2(r)$ 값은 정상범위보다 커진다. 그림 3 은 source-end 탐지 기법의 과정을 나타낸 순서도이다.



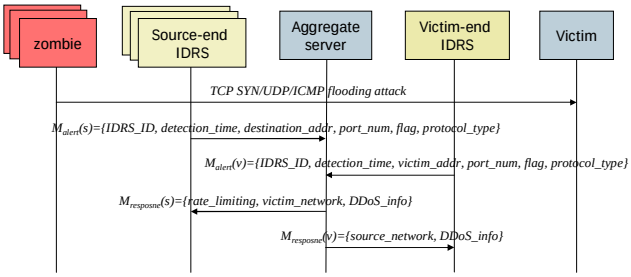
(그림 3)Victim-end 탐지 순서도

3.3 통합 서버(Aggregate server)

공격 발생 시 Victim-end 와 Source-end 에서 탐지된 Alert 은 통합 서버로 모아진다. Victim-end 의 탐지 정확도가 Source-end 의 탐지 정확도보다 높기 때문에 통합 서버는 Victim-end 의 Alert 을 기준으로 Source-end 에서 보낸 Alert 과 비교한다. 그리고 일치하는 침입탐지대응시스템에 Rate limiting 메시지를 보내 적정 범위를 초과하는 이상 트래픽에 대해 제한하도록 한다. 뿐만 아니라 Victim-end 와 Source-end 에서 보낸 포트번호, 프로토콜 타입, 플래그 값 등의 통계 정보를 각 침입탐지대응시스템 관리자에게 제공해 공격에 대한 통합적인 분석이 가능하도록 돕니다. 그림 4 는 통합 서버의 대응 순서도를 나타내고 그림 5 는 협력적인 방어 시스템의 구성 컴포넌트간 교환하는 메시지와 그 순서를 나타낸다.



(그림 4)통합 서버의 대응 순서도



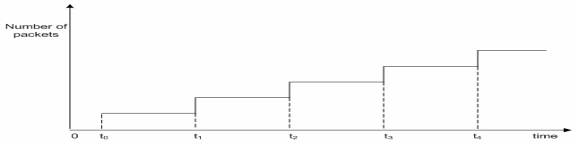
(그림 5)협력적인 방어 시스템의 시퀀스 다이어그램

3.4 지능화된 DDoS 공격 탐지

기존의 가중치를 가지는 평균과 분산을 이용하여 변화하는 트래픽에 대한 임계값 계산 방법의 문제는 마스터(master)가 좀비 에이전트의 개수를 서서히 증가시키면서 플러딩(flooding) 공격을 할 경우 탐지를

피하면서 Victim 의 대역폭(Bandwidth)을 고갈시킬 수 있다. 그림 6(a)는 t_0 에서 일시적으로 트래픽 양을 증가시키는 전형적인 공격 형태를 나타내고 그림 6(b)는 t_0 에서 t_4 에 이르기까지 장시간 점차적으로 임계값을 증가시켜 공격을 피하면서 ISP(Internet service provider)의 네트워크에 장애를 일으키는 형태를 나타낸다.

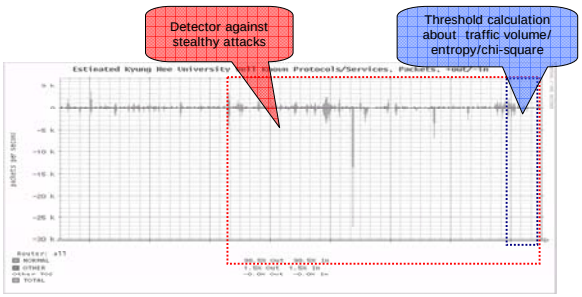
(a) Constant rate 공격



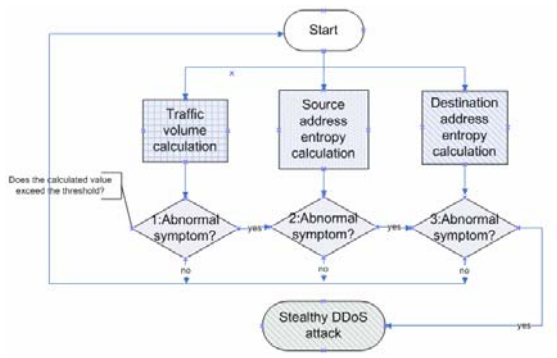
(b) Increasing threshold 공격

(그림 6) DDoS 공격 형태

이러한 공격은 트래픽 볼륨과 소스 주소 엔트로피 값은 서서히 증가하고 목적지 주소에 대한 엔트로피 값은 서서히 감소하는 특징이 있다. 따라서 트래픽 볼륨, 소스 주소 엔트로피 그리고 목적지 주소 엔트로피의 임계값에 대한 평균과 분산을 이용해 다양한 형태로 탐지 임계값을 피하는 공격 형태에 대한 탐지가 가능하다. 일반적으로 네트워크 관리 시스템은 정적인 임계값을 가지고 있지만 제안한 메커니즘을 이용하면 탐지 임계값을 증가시키는 공격을 조기에 탐지하고 대응할 수 있는 판단 근거를 제공할 수 있다. 그림 7 은 이러한 이중 탐지 윈도우를 이용한 탐지 방법을 나타내며 그림 8 은 임계값에 대한 트래픽 볼륨, 소스 주소와 목적지 주소 엔트로피 값을 이용해 지능적인 공격 형태를 탐지하는 방법의 순서도를 나타낸다.



(그림 7)이중 탐지 윈도우를 이용한 방법



(그림 8) Increasing threshold 공격 탐지 방법의 순서도

4. 결론

본 논문은 기존의 협력적인 DDoS 방어 시스템이 가지고 있는 문제점을 언급하고 탐지에 초점을 맞춰 해결 방안을 제시하였다. Victim-end 와 Source-end 침입탐지대응시스템에서 트래픽 볼륨, 소스 주소 엔트로피, 카이 제곱, 그리고 목적지 주소 엔트로피를 이용하여 실시간으로 신뢰성 있는 탐지가 가능하며 각 구성 컴포넌트간 협력을 통해 신속하고 정확한 대응이 가능하다. 또한 탐지 임계값을 피하면서 Victim 의 대역폭을 고갈시키는 Increasing threshold 공격에 대해 이중 탐지 윈도우를 이용한 방법을 이용해 탐지할 수 있다. 따라서 Source-end 와 Victim-end 에 이러한 방법을 적용하여 협력적인 방법을 통해 다양한 DDoS 공격 형태에 대해 효과적인 탐지 및 대응이 가능하다.

향후 연구 과제로는 본 논문에서 제안한 탐지 메커니즘을 테스트베드로 구축하여 구현하고 다양한 공격 형태[4][9]에 대한 성능 평가를 하는 것이다.

참고문헌

[1] R. Power, “2000 CSI/FBI Computer Crime and Security Survey”, Computer Security, vol. 16, no 2, 2000, pp33-49

[2] Bill McCarty, “Botnets: Big and Bigger”, Security & Privacy Magazine, IEEE, Vol. 1 Issue 4, pp.87-90, July-Aug. 2003

[3] Mirkovic J., Robinson M., Reiher P., "Alliance formation for DDoS defense", Proceedings of the 2003 workshop on New security paradigms, 2003

[4] Mirkovic J., Reiher P., "A taxonomy of DDoS attack and DDoS defense mechanisms", ACM SIGCOMM Computer Communication Review, April 2004

[5] Ratul M., Steven M., Sally F, John I., Vern P., Scott S., "Controlling high bandwidth aggregates in the network", ACM SIGCOMM Computer Communication Review, July 2002

[6] Mirkovic J., Prier G., Reiher P., "Source-end DDoS defense", Network Computing and Applications, NCA 2003. Second IEEE International Symposium on, 16-18 April 2003

[7] Carl G., Kesidis G., Brooks R.R., Suresh Rai, "Denial-of-service attack-detection techniques", Internet Computing, IEEE, Jan.-Feb. 2006

[8] Feinstein L., Schnackenberg D, Balupari R., Kindred D., "Statistical Approaches to DDoS Attack Detection and Response", DARPA Information Survivability Conference and Exposition(DISCEX 2003), April 22-24, 2003

[9] The "stacheldraht" distributed denial of service attack tool, http://staff.washington.edu/dittrich/misc/stacheldraht.analysis