군집분석을 사용한 효율적인 DDoS 공격 탐지 방법
An Effective Detection Method of DDoS attack using Cluster Analysis

김주현 , 김세헌
KAIST 산업공학과 통신시스템 및 인터넷 보안 연구실 jhkim@tmlab.kaist.ac.kr
KAIST 산업공학과 통신시스템 및 인터넷 보안 연구실 shkim@tmlab.kaist.ac.kr

Abstract

DDoS(Distributed Denial of Service) attack can easily exhaust the computing and communication resources of its victim within short period of time and it deteriorates performance of whole network as well as interrupts communication of an specific host. This paper analyzes network traffic using statistical method and presents a method of effective detection of DDoS attack by observing change of source IP address, destination IP address, source port, destination port, the type of packets, the number of packets.

1. 서론

2000년 11월 CNN과 eBay, Yahoo 등 주요 인터넷 호스트들이 서비스 거부 공격 (Denial of Service: DoS)에 의해 피해를 입었다. 네트워크 기술이 빠르게 발전하고 있는 이 시점에서 네트워크 보안이 현재 가장 중요한 이슈 중 하나로 떠오르고 있다 [1], [2].

DoS 공격은 서비스 자체를 파괴하는 대신 호스트나 서비스에 접속하는 것을 제한하여 서비스를 방해하는 것으로, 네트워크가 정상적인 서비스를 제공하지 못하도록 한다. 이는 일반 사용자들의 접속을 거부할 만한 처리 용량이나, 네트워크의 대역폭을 넘어서는 패킷을 대상 호스트에 보냄으로써 이루어진다. 분산 서비스 거부 공격 (Distributed Denial of Service: DDoS)의 경우는 하나의 호스트가 아니라 분산된 여러 개의 호스트가 특정 호스트로 패킷을 집중시킴으로써 대상 호스트가 서비스를 제공하지 못하도록 방해한다. 공격자들은 추적을 피하기 위해 작은 크기의 패킷으로 필드를 변화시켜 공격을 시도한다.

이 논문에서는 DDoS 공격이 일어나는 동안 네트워크의 트래픽 경향을 분석하고, 군집분석의 방법을 이용해 공격과 비공격의 군집으로 나누어 효과적으로 DDoS 공격을 탐지할 수 있는 방법을 제안하려 한다.

본 연구는 대학 IT연구센터 육성지원사업의 결과로 수행되었음.

2. DDoS 공격

앞에서 언급한 바와 같이 DDoS 공격의 목표는 컴퓨터나 네트워크가 정상적인 서비스를 제공하지 못하도록 하는 공격이다. DDoS 공격은 대상 호스트 자체의 취약점을 이용하는 것이 아니라 주로 접속 자원이 노출되어 있는 인터넷의 구조상의 약점을 이용하고 있기 때문에 전통적인 보안 메커니즘을 무력하게 만들며, 또한 이는 어떤 시스템도 공격할 수 있음을 뜻한다.

2.1 DDoS 공격의 4가지 구성 요소

DDoS 공격은 실제 공격자와 공격자가 호스트 또는 네트워크의 취약점을 이용해 침입하여, 공격 프로그램을 설치해 중간 단계의 노드로 사용하게 될 핸들러 혹은 마스터, 핸들러 혹은 마스터가 실제 공격 노드로 사용하기 위해 프로그램을 설치한 좀비 호스트 또는 에이전트, 그리고 DDoS 공격의 대상 호스트 이렇게 4가지로 구성되어있다.

2.2 DDoS 공격의 단계

① 에이전트의 선택: 공격자가 접속을 할 수 있도록 취약점을 가지고 있고, 강한 공격 스트림을 만들기에 충분한 자원을 가지고 있는 에이전트를 선택한다.
② 침입과 통신: 에이전트의 보안의 결점과 기계의 취약점을 이용하여 공격코드를 설치한다. 이런 침입이 일어난 후 공격자는 공격을 시행할 시기를 결정하고 에이전트의 업그레이드를 위해 수많은 핸들러와 교신하는 데, 이 때 교신은 TCP, UDP, 또는 ICMP 패킷을 통해 이루어진다.
③ 공격: 패킷의 유형, TTL, 포트 같은 패킷의 필드를 수정하여 다양한 패킷을 사용함으로써 탐지를 어렵게 할 수 있다.

2.3 DDoS 공격의 분류

DDoS 공격은 공격에 사용된 약점에 따라 다음과 같이 분류될 수 있다.

① Flood 공격: 대상 호스트의 대역폭을 넘치게 하기 위해 많은 양의 트래픽을 대상 호스트로 보내는 공격.

1) UDP flood 공격- 대상 호스트의 임의의 포트로 UDP 패킷을 보내, 대상 호스트의 포트에 해당 애플리케이션이 없다는 것을 알아차리면 대상 호스트는 destination unreachable 이라는 ICMP 패킷을 위조된 출발지 IP 주소로 보내게 되고, 만약 이러한 UDP 패킷의 양이 많아진다면, 대상 호스트의 대역폭이 소모되어 정상적인 서비스를 제공할 수 없게 된다.

2) ICMP flood 공격- 사용자가 원거리 호스트가 살아있는 지 여부를 알아보기 위해 사용하는 ICMP echo request 패킷을 이용한 공격으로, echo request를 받은 호스트는 echo reply 패킷을 보내게 된다. 이 때 대상 호스트에 많은 양의 echo request 패킷이 도착한다면 echo reply 패킷을 보내는 데 많은 대역폭을 소모해야 될 것이다.

② 프로토콜 사용 공격: 인터넷 프로토콜의 취약점을 이용한 공격.

TCP SYN 공격- TCP 연결을 위한 셋업 단계에서 사용되는 3 way handshake 고유의 약점을 이용한 것으로 두 호스트 간에 통신을 시작하려면, 접속을 희망하는 호스트가 접속을 희망한다는 SYN 패킷을 보내고, 서버가 허락한다는 SYN과 ACK를 보낸 후, 다시 상대 호스트가 ACK를 보내면서 통신이 시작된다. TCP SYN 공격은 가짜 출발지 IP 주소를 이용해 많은 양의 SYN 패킷을 보내어, 대상 호스트가 SYN/ACK를 보내고, 또 ACK를 기다리도록 하여 다른 접속을 처리하지 못하도록 한다 [3].

3. 군집분석

군집분석은 관측 대상들 간에 어떤 공통 특징을 찾아 비슷한 특징을 갖는 관측들 끼리 군집을 형성하는 방법으로, 군집의 개수, 내용, 구조 등이 사전에 정해지지 않은 상황에서 상사성에 근거하여 군집으로 나눔으로써 식별된 군집 간의 관계 등을 연구, 분석하는 것이 군집분석의 목적이다 [4].

군집분석에서 군집의 형성 과정은 크게 계보적 기법, 분배 기법(partitioning)으로 나뉜다.

계보적 기법에는 각 객체를 하나의 군집으로 보고, 가장 가까운 객체를 묶어 결국에는 한 개의 군집을 만들어내는 병합적 방법과 그것의 반대 개념으로 모든 객체를 한 개의 군집으로 취급하여, 각 개체를 모두 다른 군집으로 분류하는 분할적 방법이 있다. 분배 기법은 군집의 수가 사전에 미리 결정되는 경우에 사용되며, 각 군집에 대한 판정기준을 결정하고 이에 따라 객체를 군집에 할당하는 방법을 취하는 것으로 대표적인 방법은 K-means가 있다.

군집분석에서 어려운 점은 각 관측 값 사이의 거리와 군집들 사이의 거리를 어떻게 정의하는가하는 것과 어떤 기준으로 군집을 병합 혹은 분할 할 것인가 하는 것이다.

관측점 간의 거리를 구하는 데에는 유클리드 거리와 민코브스키 거리, 마할라노비스 거리가 사용된다.

이러한 거리 기준을 바탕으로 군집을 병합 혹은 분할할 때 사용하는 기준은 최단 연결법, 최장 연결법, 평균 연결법, Ward의 연결법 등이 있으며, 최단 연결법은 군집 간의 거리를 정의할 때 군집 내의 가장 가까운 객체를 기준으로 삼는 것이고, 최장 연결법은 가장 멀리 있는 객체를 기준으로, 평균 연결법은 거리의 평균을 기준으로 삼는 것이다. Ward의 방법은 보편적으로 많이 사용되는 방법으로 각 관측 값들과 관측 값이 속해있는 군집의 평균과의 거리 제곱의 총합인 잔차 제곱합의 증가를 최소화하여 군집을 묶음으로써 생기는 정보의 손실을 최소화 하도록 하는 방법이다.

4. Detection Model

위에서 알아본 DDoS 공격의 특성에 미뤄 보면 DDoS 공격은 주로 많은 패킷을 특정 호스트에 집중시킴으로써 그 호스트의 기능을 마비시키며, 이 때 주로 사용되는 패킷은 UDP ICMP, TCP SYN, TCP ACK 등의 패킷임을 알 수 있다. 하지만 단순히 이들 패킷의 빈도가 높다는 것만을 기준으로 이상여부를 판단하는 것은 적절치 못하다. 이런 패킷들이 다양한 출발지 IP 주소를 가지고, 특정 호스트를 목적지 IP 주소를 가진다는 정보까지 추가된다면 DDoS 공격이라는 판단을 내리는 데 무리함이 없을 것이다.

4.1 군집분석에 사용될 변수

DDoS 공격은 여러 개의 가짜 출발지 IP를 가진 패킷들이 특정 호스트에 집중됨으로써 일어나는 공격이다. 그러므로 패킷이 다양한 출발지 IP주소를 가지고 있음에도 불구하고, 그 패킷들이 하나 또는 적은 수의 목적지 IP 주소를 가진다면 DDoS 공격으로 볼 수 있다.

이러한 정보를 포함하기 위해 여기에서는 출발지 IP 주소와 목적지 IP 주소의 수렴과 발산정도를 쉽게 알 수 있도록 엔트로피의 개념을 이용하였다.

엔트로피 H는 다음과 같이 정의되며,

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

네트워크를 지나는 패킷이 다양한 IP 주소를 가진다면, 엔트로피는 증가할 것이고, 특정 IP 주소에 집중된다면, 엔트로피는 감소할 것이다.

DDoS 공격은 다양한 가짜의 출발지 IP 주소를 가진 패킷이 특정한 하나의 목적지 IP 주소로 전달되기 때문에 공격이 일어날 때 출발지 IP 주소의 엔트로피는 증가하고, 목적지 IP 주소는 감소할 것이다 [5].

또한 DDoS 공격은 주로 임의의 port를 대상으로 공격하거나 특정 port를 대상으로 하게 되는데, 이는 공격의 특성에 따라 달라지므로 port의 엔트로피는 공격의 종류를 구분 짓는 데 도움을 줄 것이라고 예상된다.

앞에서 언급한 DDoS 트래픽의 특성에 따라 특정 패킷, 예를 들어 UDP, ICMP, TCP SYN 패킷이 나타나는 빈도, 즉 단위시간(1초) 마다 위의 패킷의 확률을 탐지 모형의 변수로 사용하고, 단위시간 당 발생하는 패킷의 수가 매우 작을 때(1~10개 :평균 패킷 수는 40)는 위의 패킷이 한 두 차례 나타나는 것만으로도 이상으로 판단될 위험이 있고, 또한 DDoS 공격 시는 네트워크를 지나는 패킷의 양이 비정상적으로 많아지므로(약 100배 혹은 그 이상) 단위시간당 패킷의 수도 변수로 도입하였다.

DDoS 공격은 특정 패킷을 이용한 경우가 많으므로 위에서 고려한 UDP, ICMP, TCP SYN 외의 패킷을 고려하기 위해, 패킷의 종류에 따른 엔트로피를 변수로 도입하였다.

5. Data Set

분석에 사용된 Data Set은 2000 DARPA Intrusion Detection Scenario Specific Data Set으로 이 때 사용된 DDoS 공격의 종류는 Mstream이며, 공격 기간은 5개의 공격 단계로 나뉘어 지며, 그 5 단계의 공격 단계는 다음과 같다.

1. 원거리 사이트로부터의 IP 탐색.
2. 공격에 이용할 sadmind daemon이 사용되는 IP 탐색.
3. sadmind 취약점을 통해 침입.
4. trojan mstream DDoS software를 설치.
5. DDoS 시작.

위의 공격 단계로 보아, 단계 1에서는 IP 탐색을 위한 ICMP request ("ping")의 발생이 비정상적으로 많을 것으로 예상되고, 단계 2에서는 취약점으로 사용할 sadmind이 작동하는 host를 찾기 위해 sadmind 서비스를 연결하기 위한 port가 무엇인지를 물어보는 rcp request을 이용하고, 이로 인해, port unreachable의 ICMP 패킷이 많이 발생할 것이다. 단계 3과 4는 호스트 상에서 일어나는 현상이므로 네트워크 트래픽 상의 변화는 알아내기 힘들 것이고, 단계 5는 실제 DDoS 공격이 일어나는 시점이므로 앞에서 제시한, 각각의 엔트로피 관련 변수와 패킷 수가 정상상태와 뚜렷이 구분될 것으로 예상된다.

앞에서 제시된 변수들이 이런 특징을 잘 설명하여 공격 단계별 군집을 형성할 수 있다면, DDoS 공격이 시작되기 전에 미리 공격을 탐지할 수 있을 것이라고 생각된다.

이 Data Set에서 제공된 자료는 TCP dump 자료로 이 자료로부터 얻을 수 있는 정보는 출발지 및 목적지의 IP 주소와 port, 패킷의 종류, 윈도우 사이즈로 제한되어 있다. 또한 실험에 사용된 네트워크가 DMZ와 내부 네트워크로 분리가 되어있는 형태이다. 여기서 DMZ란 사설 네트워크와 외부 공중 네트워크 사이에 중립지역의 역할을 하기 위해 삽입된 컴퓨터 호스트나 소형 네트워크를 말하는 것으로, 외부 사용자들이 내부의 중요한 데이터를 보유하고 있는 서버에 직접 접속하는 것을 방지하기 위한 것이다. 즉 DMZ를 사용할 경우 외부의 공중 네트워크 사용자들은 오직 DMZ 호스트에만 접근할 수 있다.

즉 이 Data Set의 외부 공격자는 핸들러를 찾기 위해서 DMZ 내의 호스트로만 접속이 가능하고, 실제 내부 네트워크로의 접근은 DMZ 내의 핸들러를 통해서만 가능하게 됨을 뜻한다.

그렇기 때문에 각 공격의 단계를 파악할 수 있기 위해서 내부 네트워크의 트래픽을 살피는 대신 DMZ의 트래픽을 살펴보았다. 대신 DMZ 내에는 공격 대상 호스트가 존재하지 않고, 핸들러 호스트만이 존재하기 때문에 위에서 언급한 출발지 IP 주소와 목적지 IP 주소에 대한 수렴, 발산의 성질이 반대로 나타나게 된다.

6. Detection

군집분석을 시행하기 위해 먼저 간단히 변수의 기초 통계량을 알아본 결과는 <표 5.1>과 같다.

변수	평균	표준편차	최소값	최대값
출발지IP	1.5818	0.6230	0	3.466
출발지Port	1.6073	0.7224	0	12.486
목적지IP	1.5844	0.6874	0	12.668
목적지Port	1.4924	0.7118	0	12.668
패킷엔트로피	1.1149	0.4141	0	4.428
패킷 수	40.172	151.533	1	6524
SYN 확률	0.0249	0.0604	0	1
UDP 확률	0.0047	0.0595	0	1
ICMP 확률	0.0073	0.0500	0	1

<표 5.1>

각각의 변수들이 단위가 다르기 때문에 변수값을 수정 없이 그대로 군집분석에 사용한다면 그 스케일이 큰 것의 영향이 크게 작용할 수밖에 없으므로 여기서는 각 변수에서 평균을 뺀 후 표준편차로 나눈 z값으로의 변환을 사용 하였다.

군집분석은 통계 프로그램인 SAS 9.1 버전의 Enterprise Miner를 이용하였고, 사용한 군집 방법은 계보적 기법의 병합적 방법을 이용하였으며, 병합의 기준은 가장 보편적으로 사용되는 Ward의 방법을 이용하였다.

이런 병합적 방법에서 가장 적절한 군집의 수를 결정하는 데 이용한 것은 CCC(Cubic Clustering Criterion)로 군집의 수에 대해 CCC 값을 플로팅하여 형성하는 최고점을 관찰하여 판단한다. CCC>3 일 때, 국부적 최고점이 있을 면, 이 점에서의 군집의 수가 가장 적절하다는 것이 알려져 있다 [5].

이런 방법을 이용해 군집분석을 시행했을 때, 결과는 <표 5.2>와 같다.

군집	군집 내 객체 수	가장 가까운 군집
1	5	4
2	9589	6
3	32	2
4	1	1
5	21	2
6	56	2

<표 5.2>

각 군집의 변수별 평균값을 관찰해보면 <표 5.3>과 같다

군집 / 변수	1	2	3	4	5	6
출발지IP	0.02	1.59	0.08	0.13	0.71	1.06
출발지port	12.4	1.61	0.12	11.4	0.56	1.07
목적지IP	12.6	1.58	0.07	11.5	4.91	1.06
목적지port	12.6	1.50	0.12	11.5	0.55	1.07
패킷엔트로피	0.02	1.12	0.04	0.12	0.53	1.36
패킷 수	6225	37.0	1.19	2876	41.4	4.70
SYN 확률	0	0.02	0	0	0	0.44
UDP 확률	0	0.00	0.99	0	0	0
ICMP 확률	0	0.00	0	0	0.87	0

<표 5.3>

군집 1의 경우의 특성을 보면, 출발지 IP의 엔트로피 값이 매우 낮고, 즉 출발지 IP 엔트로피 값이 특정 호스트로 수렴하고, 반대로 목적지 IP 엔트로피 값이 발산함을 알 수 있다. 이는 DMZ 네트워크의 트래픽을 관찰한 것이므로, 내부 네트워크에서 외부 네트워크로 나가는 트래픽을 뜻한다. 즉 내부에 있는 공격 대상 호스트가 많은 양의 TCP ACK 패킷을 받고, TCP RST를 가짜의 소스 주소로 회신하는 패킷을 의미하므로 DMZ 네트워크의 트래픽은 앞에서 설명한 DDoS 공격의 특성과 반대되는 엔트로피 특성을 가지게 된다. 또한 공격 시 네트워크를 지나가는 패킷의 양이 일반적인 상황보다 월등히 많음을 알 수 있고, 이 동안의 패킷 종류의 엔트로피가 매우 작다는 것은 동일한 종류의 패킷, 즉 TCP RST 패킷이 대부분을 차지함을 의미한다.

군집 2의 특성은 출발지, 목적지의 IP주소와 port가 평균에서 크게 벗어나지 않고, 패킷의 수도 평균에서 크게 벗어나지 않는 정상의 상태라고 보여 진다.

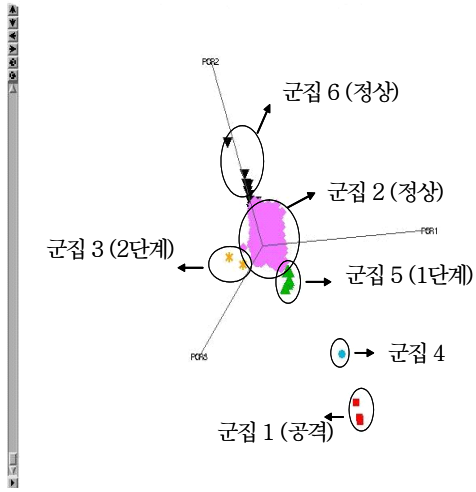
군집 3은 출발지, 목적지 IP 주소 및 port 번호 패킷의 종류에 대한 엔트로피가 상대적으로 매우 낮고, 이상적으로 UDP 패킷의 확률이 높은 것으로 보아 위의 5단계 중에 2단계에 속함을 알 수 있다.

군집 4는 다른 군집 중 군집 1과 가장 가까운 군집으로 대체적으로 비슷한 특성을 가지고 있다. 이는 실제 공격 상황은 아니지만, DMZ 네트워크에서 관찰한 트래픽이므로 공격당한 후, 그에 대한 RST 패킷을 회신으로 보내고 있기 때문에 공격 후에도 그 특성이 남아있음을 뜻한다.

군집 5는 비교적 낮은 출발지 IP 주소 엔트로피와 높은 목적지 IP 엔트로피를 가지고 있으며, ICMP 패킷의 비율이 상대적으로 높다. 이 경우 IPsweep이 일어나는 1단계를 나타냄을 알 수 있다.

군집 6은 군집 2와 가장 가깝고, 그 특성이 매우 비슷한데, 단지 단위시간당 패킷 수가 적은 편이다. DDoS 공격의 경우 많은 패킷을 이용하여 시도되는 공격이므로, 패킷의 수만이 비정상적일 경우는 정상으로 구분할 수 있다.

<그림 5.1>은 변수 9개로 설명되는 위의 분석을 주성분 분석을 통하여 반응 변수들을 선형 변환하여 3개의 주성분으로 단순화하여 3차원 공간에 표현해 본 것이다.



<그림 5.1>

위의 그림으로 보아 공격의 단계별 군집이 뚜렷이 구분됨을 알 수 있다.

7. Conclusion

DDoS 공격은 핸들러 혹은 에이전트를 분산된 형태로 배치함으로써 공격의 효과를 극대화하는 네트워크 공격 중 가장 위협적인 공격 방식이다. 본 논문에서 제시한 모형은 단순하게 공격과 비공격 상태의 군집으로 분류하여 공격시점을 탐지하는 것 이상으로 각 공격의 단계에 따라 달라지는 네트워크 트래픽의 특성에 따라 군집을 분류함으로써 공격이 일어나기 전 단계에서 공격의 기미를 탐지할 수 있다는 장점이 있으며, 또한 각 관측치의 값이 어느 군집과 가장 가까운 지만을 계산하면, 객체가 해당되는 군집을 파악할 수 있으므로 수행하기 쉽고, 간단하며 빠른 시간 내에 탐지 할 수 있다는 장점이 있다.

8. References

[1] Stephen Northcutt, Judy Novak, (2003). Network Intrusion Detection, 3rd edition.

[2] Shun-Chieh Lin, Shian-Shyong Tweng, (2004). Constructing detection knowledge for DDoS intrusion tolerance, Expert Systems with Applications 27 pp379-390.

[3] Christos Douligeris, Aikaterini Mitrokotsa, (2004). DDoS attacks and defense mechanisms: classification and state-of the-art, Computer Networks 44 pp643-666.

[4] 조인호, (2005). SAS 강좌와 통계컨설팅 2nd edition.

[5] 김민택, A statistical approach to network attack detection on backbone links.

[6] 김기영, 전명식, (1991). SAS 군집분석