IP Traceback 위한 SVM기반 패킷 마킹 기법

김길한°, 이형우
한신대학교 소프트웨어학과

SVM based Packet Marking for IP Traceback

Gill-Han Kim°, Hyung-Woo Lee
Department of Software, Hanshin University

요 약

DDoS(Distributed Denial-of-Service) 공격은 인터넷을 통한 보안 위협 중 대표적인 분산 서비스 거부 공격이다. DDoS은 해킹 공격자가 공격 근원지 IP 주소를 스푸핑하여 공격목표로 하는 시스템의 가용자원을 고갈시키거나 과도한 부하를 유발시켜 서비스를 중단시킨다. 이러한 공격에 대한 대응 기술로 제시된 IP 역추적 기술은 DDoS 공격의 근원지를 판별하고 전달된 공격 패킷을 통하여 네트워크상에서 공격 패킷 전달 경로를 재구성한다. 기존의 역추적 기술인 패킷 마킹 기법에서 DDoS 공격에 대한 판별 과정 없이 임의의 패킷에 대해 역추적 정보를 생성 즉 DDoS 공격에 능동적으로 대응하고 있지 못하는 단점에 착안하여 본 연구에서는 SVM 모듈을 적용한 라우터에서 DDoS 트래픽에 대한 판별 기능을 제공하고 또한 DDoS 공격 패킷에 대해 개선된 마킹 기법을 제시하였다. 연구 실험 결과 네트워크 부하를 줄이면서도 역추적 성능을 향상시킬 수 있었다.[1)]

1. 서 론

현재 TCP SYN flooding[1] 공격과 같은 서비스 거부 공격(Dos: Denial of service)[2]을 통해 TCP/IP 체계의 취약점이 노출되어 있기 때문에 네트워크 및 인터넷에서의 해킹 공격에 대응할 수 있는 방안에 대해 연구가 진행되고 있으며, DDoS 공격과 같은 해킹 공격에 대한 대응하는 방법은 크게 백신, 침입탐지 및 침임감내 기술 등과 같은 수동적인(passive) 대응 방법과 공격 근원지 역추적(Traceback) 기법과 같은 능동적인(active) 대응 방법이 있다.

역추적 방식은 네트워크상에 패킷이 전송되는 과정에서 사전에 라우터는 역추적 경로 정보를 생성하여 패킷에 삽입하거나 패킷의 목적지 IP 주소로 전달하여 주기적으로 관리하는 방식이다. 만일 피해 시스템에서 해킹 공격이 발생하면 이미 생성, 수집된 역추적 경로 정보를 이용하여 스푸핑된 해킹 공격 근원지를 판별하는 기법이다. 패킷에 대한 확률적 마킹(PPM : probabilistic packet marking)[3,4] 기법과 ICMP 메시지를 변형한 iTrace (ICMP traceback)[5] 기법 등이 이에 해당한다. 또한 최근 제시된 SVM[6,7,8,9] 기법은 DDoS 공격이 발생하였을 경우 입력 패킷에 대해 양쪽의 비선형 분류 기능을 제공한다. 이 기법은 DDoS 공격 트래픽에 대한 제어 기능을 제공하지만 DDoS 해킹 공격 근원지를 역추적하는 기능은 제공하지 못하고 다만 다양한 패킷 전송에 있어서 분류기법을 제공함으로써 악의의 패킷을 필터링하여 DDoS 패킷들의 네트워크 전송을 제어한다.

따라서, 본 연구에서는 기존의 DDoS 공격에 대한 필터링과 제어 기능을 제공하는 SVM기술을 역추적 기능과 접목하여 스푸핑된 DDoS 패킷에 대한 IP 근원지를 역추적하는 기술을 제안하고자 한다. 라우터에서는 SVM 기법을 적용하여 트래픽에 대한 판별/제어 기능을 수행하고 만일 DDoS 공격이 발생하였을 경우 상위 라우터로 라우터의 정보를 해당 패킷의 헤더에 마킹하여 전달한다. 제시된 기법을 통해 기존의 역추적 기법보다 관리시스템 부하, 네트워크 부하 및 역추적 기능 등을 향상시킬 수 있었다.

2. Support Vector Machine

2.1 SVM 개요

Support Vector Machine(SVM)은 1995년 Vapnik에 의하여 개발되고 제안된 학습 알고리즘이다. SVM 이론에 따르면, 패턴 인식을 위한 전통적인 기법들이 경험적인 위험을 최소화하는데 기초한 반면, SVM은 구조적인 위험을 최소화하는 것에 기초하고 있다. 여

1)본 연구는 대학IT연구센터육성지원사업의 연구결과로 수행되었음.

기서 경험적 위험의 최소화는 훈련 집단의 수행도를 최적화하려는 노력을 말하고, 구조적 위험의 최소화는 고정되어 있지만 알려지지 않은 확률분포를 갖는 데이터에 대해 잘못 분류하는 확률을 최소화하는 것을 말한다. SVM의 장점은 우선 훈련 집단에 포함된 정보를 모으는 능력이 있다는 것과 상대적으로 낮은 공간의 결정 평면 집단을 사용한다는 것이다. 패턴 집단이 선형이고 분리 가능한 경우에 있어 SVM의 주요 아이디어는 간단히 설명될 수 있다.

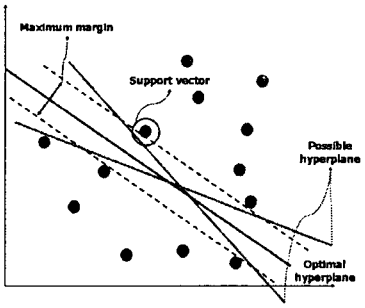

그림 1. SVM 기반 분류 기법

그림 1과 같이 기본적으로 SVM는 입력패턴들을 교사학습방법을 통하여, +1과 -1의 두 클래스로 패턴을 분류한다. 훈련 집단 S는 두 클래스로 분류되면, 각 클래스에 포함된 훈련 패턴들을 분리하는 초월면(Hyperplane)이 결정된다. 여기서 초월면이란 각 집단을 분리하는 절단 평면을 일컫는다. 이때, 초월면을 결정하는 입력 패턴들을 Support Vector라 한다. 패턴 집단이 분리 가능한 경우, 이 초월면은 면으로부터 Support Vector까지의 거리(마진)를 최대화하며, 모든 Support Vector는 초월면으로부터 같은 최소 거리에 위치해 있다. 그러나 실제로 패턴집단이 선형으로 분리되는 경우는 거의 드물고, 따라서 두 클래스는 선형적으로 분리가 불가능한 경우가 많을 것이다. 이 때의 초월면과 Support Vector는 제약식을 갖는 최적 문제의 해로부터 얻어진다. 최적해는 마진(각 클래스의 Support Vector사이의 거리)을 가장 크게 하는 것과 에러의 수를 최소화는 것 사이의 trade-off를 가지고 있고, 이는 정규화 된 파라미터에 의해 조정된다. 훈련 과정은 제약식을 갖는 이차 최적 문제를 풀기 위한 것과 기본적으로 같다[10].

2.2 SVM기반 IP Traceback

네트워크는 노드 집합 V와 에지 집합 E로 구성된 그래프 $G=(V,E)$로 정의할 수 있다. 다시 네트워크 노드 집합 V는 종단 시스템과 내부 노드에 해당하는 라우터로 나눌 수 있다. 에지는 V 집합 내에 있는 노드들에 대한 물리적인 연결에 해당한다. $S \subset V$를 공격자라고 정의하고 $t \in V/S$를 피해 시스템이라고 정의한다. 만일 $|S|=1$일 경우 단일 공격자에 의한 해킹 공격을 의미하고 공격 경로 정보 $P=(s, v_1, v_2, ..., v_d, t)$인 경우 공격 시스템 s에서 피해 시스템 t로 d개의 라우터를 통해 전달된 공격 경로를 의미한다. 이때 전달된 패킷의 수를 N이라고 하자. 만일 패킷내에 라우터에 대한 링크 정보 $(v,v') \in E$를 마킹할 수 있는 필드가 있다면 이를 확률 p로 샘플링하여 전달하게 된다. 패킷에 대해서 라우터에서는 일정한 확률로 패킷을 선택하여 에지에 대한 정보와 라우터에 대한 거리 정보를 패킷내에 포함시켜 전달할 수 있다.

기존의 기법에서는 임의의 확률 p로 패킷을 선택하여 여기에 라우터에 대한 링크 정보를 마킹하여 전달하게 된다. 만일 네트워크 상에서 노드 v_i에서 마킹하였을 경우 다른 라우터에 의해서는 재마킹되지 않고 전달될 확률 α_i을 계산하면 다음과 같다.

$$\alpha_i = Pr(x_d = (v_{i-1}, v_i)) = p(1-p)^{d-1} \ (i = 1, 2, ..., d)$$

따라서 확률 α_i는 공격자에 해당하는 패킷 정보가 다른 라우터에 의해서는 재마킹되지 않고 피해 시스템에 전달될 확률을 의미한다. 결국 피해 시스템에서 α_i 값을 높이기 위해서는 p 값을 크게 해야 하는데, 이는 라우터에서 빈번하게 마킹 과정을 수행해야 한다는 것을 의미하므로 기존의 기법에서는 결과적으로 네트워크 성능을 저하시키게 된다.

본 연구에서 제시하는 기법은 라우터에서 임의의 확률 p로 패킷을 샘플링하여 마킹하지 않고 SVM 모듈에 의해서 이상 트래픽이 발견되었을 경우 패킷에 대한 마킹 과정을 수행하게 된다. 본 연구에서 제안한 구조는 그림 2과 같다.

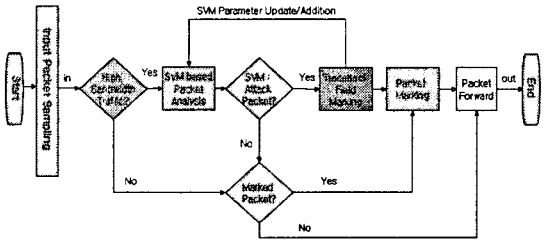

그림 2. 제안한 SVM 기반 DDoS 근원지 역추적 구조

제안한 구조에서는 라우터에 들어온 패킷에 대해 트래픽의 대역폭을 검사하고 일정 이상으로 도착하게 되면 공격 형태에 해당하는 혼잡 시그너처인지를 판단하게 된다. 만일 공격 형태 트래픽에 해당한다면 패킷에 마킹과정을 수행하고 동시에 해당 패킷에 대한 SVM 메시지를 생성하여 이를 라우터의 출력 큐로 하여금 앞단위 라우터에게 전송토록 한다. 만일 대역폭 조건을 만족하지 않을 경우에는 이전에

SVM 메시지를 통해 주변 라우터로부터 전달된 정보가 있는지를 확인하고 만일 해당된다면 마찬가지로 패킷에 대한 마킹 과정을 수행한다. 위 조건을 만족하지 않을 경우 일반적인 트래픽으로 간주하여 다음 라우터로 전달한다.

2.3 라우터에서 경로 마킹

라우터 R_x의 IP 주소를 A_x라고 하자. 그리고 R_x에 도착한 IP 패킷을 P_x라고 할 때, P_x에서의 헤더에서 마킹 정보를 저장할 수 있는 24 비트를 M_x라고 했을 때, 패킷 P_x에서 M_x는 그림 3과 같이 TOS(type of service) 필드 8비트와 ID 필드 16비트로 구성된다. TOS 필드인 경우 현재 필드에 대한 정의만 되어 있을 뿐 실제적으로 사용하고 있지 않다. 따라서 TOS 필드 값을 사용한다고 하더라도 전체 네트워크에 영향을 미치지 않는다.

현재의 TOS 필드는 상위 3비트가 우선순위 비트로 설정되어 있고, 다음 3비트는 최소지연, 최대 성능 및 신뢰성 필드로 정의되어 있으나 현재는 사용하고 있지 않다. 다만 최근에 RFC2474에 의하면 Differenciated Service 필드(DS field)로 재정의하였으며 TOS 8비트 중에서 상위 6비트만을 사용하고 하위 2 비트는 사용하지 않고 있다. 따라서 본 연구에서는 TOS 필드 중에서 현재 사용하고 있지 않은 2비트에 대해서 TM(traceback marking flag)와 CF(congestion flag)로 정의한다. 특히 CF인 경우 RFC2474에서도 네트워크상에서 혼잡 현상이 발생하였을 경우 1로 설정하도록 정의되어 있다.

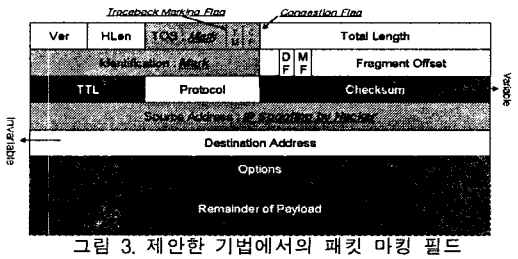

그림 3. 제안한 기법에서의 패킷 마킹 필드

앞에서 제시한 SVM 기반 역추적 모듈을 통해 이상 트래픽이 발생하였다는 것을 통보받게 되면 이제 라우터 R_x에서는 SVM 메시지 내에 포함된 혼잡 시그너쳐에 해당하는 패킷 P_x에 대해서 마킹 과정을 수행한다.

3. 역추적 경로 재구성

네트워크를 통해 전달된 패킷에 대해 피해시스템 V에서는 DDoS 공격 경로를 재구성하게 된다. 공격 패킷에 대해 라우터 R_x, R_y 및 R_z는 패킷 헤더 24비트 정보내에 라우터 자신의 IP 정보와 패킷에서의 TTL 필드 6비트 정보를 마킹하였다. 피해시스템에서는 DDoS 공격이 발생하였을 경우 도착한 패킷에 대해 아래와 같이 경로 역추적 과정을 수행한다.

우선 피해시스템 V에 도착한 패킷을 P_v 집합이라고 정의하자. P_v 값은 DDoS 공격에 해당하는 패킷들로 구성된 집합이고, 집합내에서 라우터에 의해 마킹되어 전달된 패킷의 집합을 M_v라고 하자.

피해시스템에 도착한 패킷 집합 P_v에서 M_v값을 구별하는 방식은 아래와 같이 패킷에서의 TOS 필드 값중에서 임의의 패킷 P_x에서의 패킷 PF 필드에 해당하는 P_x^{PF}와 CF 필드 P_x^{CF} 부분이 설정되어 있는 패킷을 선택하는 과정을 수행하게 된다.

$$M_v = \{P_x | P_x^{PF} == 1 \wedge P_x^{CF} == 1, x \in v\}$$

즉, 피해시스템에서 마킹되어 있는 패킷 M_v의 원소에 해당하는 임의의 패킷 M_i에 대해서 8비트 TTL 값을 $TTLof M_i$라고 정의할 수 있고, TOS 필드에 패킷된 정보 T_{Mi}값과 비교여 패킷 M_i가 라우터로부터 마킹된 후에 전송된 네트워크 홉 거리 $D(M_i)$를 다음과 같이 계산 할 수 있다.

$$D(M_i) = M_i^{TF} - (TTLof M_i \wedge 00111111)$$

만일 $D(M_i) == 1$이라면 피해시스템 바로 앞에 있는 라우터에 의해서 마킹되었다는 것을 알 수 있다. 그러나, 본 연구에서 제시하는 기법은 pushback 기법과 연계하였기 때문에, $D(M_i) == 2$인 패킷을 대상으로 바로 역추적 경로 재구성 과정을 수행할 수 있다.

$D(M_i) == 2$을 만족하는 패킷 M_i는 피해시스템 바로 앞단에 연결되어 있는 두 홉 거리 내에 있는 라우터 R_y 및 R_x에 의해서 마킹된 패킷이라는 것을 의미한다. 즉, 패킷 M_i는 피해시스템과 바로 연결되어 있는 라우터 R_y와 2 홉 거리에 있는 임의의 라우터 R_x에 의해 마킹되었기 때문에 $D(M_i)$ 값은 2가 된다. 따라서 패킷 M_i에서 우선 2 홉 거리를 갖는 라우터 R_x를 다음과 같이 판별할 수 있다.

$$M_i^{MF1} == H(M_i^{TF} | R_x), (R_x \in D(M_i) == 2) \text{ and}$$

$$M_i^{MF1} == H((TTLof M_i \wedge 00111111) + 2 | R_x),$$

$$(R_x \in D(M_i) == 2)$$

물론 패킷 M_i는 피해시스템과 홉 거리 1에 해당하는 라우터 R_y에 의해 마킹되었다는 것 역시 아래과 같은 방식으로 검증이 가능하다.

$$M_i^{MF2} == H(M_i^{TF} - 1 | R_y), (R_y \in D(M_i) == 1)$$

and

$M_i^{MF2} == H((TTLofM_i \wedge 00111111) + 1 | R_y)$,
$(R_y \in D(M_i) == 1)$

이제는 $D(M_j) == n, (n \geq 3)$를 만족하는 M_j에 대해서 위와 같은 과정을 반복하게 되면 DDoS 공격 패킷 집합 P_v에서 패킷이 전달된 실제 공격 경로를 재구성할 수 있다.

아래와 같은 네트워크 구조에 대해 본 연구에서 제시한 기법을 적용하게 되면 피해시스템에 대한 DDoS 공격 경로 AP를 다음과 같이 구할 수 있다.

$AP_1 = R_y \rightarrow R_x \rightarrow R_z \rightarrow S1$, $AP_2 = R_y \rightarrow R_3 \rightarrow R_7 \rightarrow S2$,
$AP_2 = R_y \rightarrow R_3 \rightarrow R_7 \rightarrow S2$

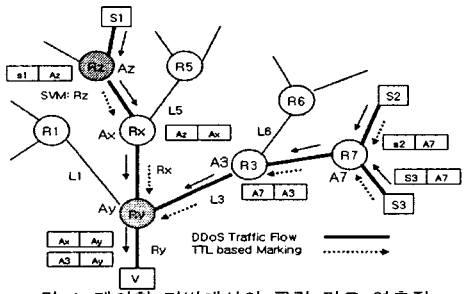

그림 4. 제안한 기법에서의 공격 경로 역추적

4. 성능 분석

본 연구에서 제시한 기법에 대한 성능을 평가하기 위해서 Linux 환경에서 ns-2 시뮬레이터를 이용하여 성능을 분석하였다.

실험 결과 기존의 패킷 마킹 기법은 DDoS 공격에 대해 각 라우터에서 확률 p로 샘플링하여 마킹하는 방식이므로 전체 마킹된 패킷의 수가 DDoS 트래픽에 비례하여 생성되는 것을 볼 수 있다. 본 연구에서 제시하는 기법인 경우 SVM 기법을 적용하여 DDoS 트래픽에 대한 마킹 과정을 수행하기 때문에 마킹된 패킷의 수가 12.8% 정도 감소하는 것을 확인할 수 있었다.

[표 1] IP 역추적 기법 성능 비교 평가

기법 \ 특성	관리 시스템 부하	네트워크 부하	피해 시스템 부하	메모리 요구	대역폭 부하	역추적 기능	적용 가능성	보안 기능	DDoS 대응	확장성	경로 재구성 패킷 수
Ingress filtering	×	×	×	×	×	×	▽	×	×	△	×
SYN flooding	×	×	↓	×	↓	×	▽	×	×	△	×
Logging	↑	×	×	↑	×	▽	▽	◇	▽	◇	1
PPM	↓	↓	↑	↑	×	△	△	◇	▽	△	↑
iTrace	↓	↓	↑	↑	↓	△	△	◇	▽	△	↑
제안한 기법	↓	↓	↓	↑	↓	△	△	△	△	△	n

×:N/AT ↑:high, ↔:middle ↓:low △:good ◇:moderate ▽:bad

4. 결론

본 연구에서는 인터넷을 통해 급격히 확산되고 있는 해킹·바이러스에 대한 대응 기술로서 DDoS 공격 등이 발생하였을 경우 스푸핑된 트래픽에 대한 실제적인 공격 근원지 IP를 피해 시스템에서 역추적하는 기술을 제시하였다. 기존 역추적 기술의 구조와 현황, 문제점 등을 고찰하여 네트워크상에서 DDoS 해킹 공격에 대한 판단/제어 기능도 제공하면서도 피해 시스템에서는 스푸핑된 해킹 공격 근원지를 효율적으로 역추적할 수 있는 새로운 패킷 마킹 기법을 제시하였다. 제시한 기법은 기존의 기법보다 부하, 성능, 안전성 및 역추적 기능에서 개선된 특징을 보인다.

[참 고 문 헌]

[1] Computer Emergency Response Team, "TCP SYN flooding and IP Spoofing attackts," CERT Advisory CA-1996-21, Sept, 1996.

[2] L. Garber. "Denial-of-service attacks trip the Internet". Computer, pages 12, Apr. 2000.

[3] K. Park and H. Lee. On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack. In Proc. IEEE INFOCOM '01, pages 338 {347, 2001.

[4] D. X. Song, A. Perrig, "Advanced and Authenticated Marking Scheme for IP Traceback," Proc, Infocom, vol. 2, pp. 878-886, 2001.

[5] Steve Bellovin, Tom Taylor, "ICMP Traceback Messages", RFC 2026, Internet Engineering Task Force, February 2003.

[6] C.J.C Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, Vol. 2, pp.121-167, 1998

[7] Jongmei Deng, Qing-An Zeng, Dharma P. Agrawal, "SVM-based Intrusion Detection System for Wireless Ad Hoc Networks", IEEE, 2003.

[8] Cortes C., "Vapnik V. Support Vector Network", Machine Learning, Vol. 20. pp.273-279, 1995.

[9] Cristianimi, N., Shawe-Taylor J., "An Introduction to Support Vector Machines", Cambridge University Press, 2000.

[10] Vapnik V., The Nature of Statistical Learning Theory, Springer-Verlag, NewYork, 1995.