제18회 한국정보처리학회 추계학술발표대회 논문집 제9권 제2호 (2002.11)

서비스거부공격에서의 퍼지인식도를 이용한 네트워크기반 탐지 모델

이세열*, 김용수**
*대전대학교 컴퓨터공학과
**대전대학교 컴퓨터정보통신공학부
e-mail : ailab@dju.ac.kr, kysti@dju.ac.kr

A Network based Detection Model Using Fuzzy Cognitive Maps on Denial of Service Attack

Se Yul Lee*, Yong Soo Kim**
*Dept of Computer Engineering, Daejeon University
**Division of Computer & Communication Engineering, Daejeon University

요 약

최근 네트워크 취약점 검색 방법을 이용한 침입 공격이 늘어나는 추세이며 이런 공격에 대하여 적절하게 실시간 탐지 및 대응 처리하는 침입방지시스템(IPS:Intrusion Prevention System)에 대한 연구가 지속적으로 이루어지고 있다. 본 논문에서는 시스템에 허락을 얻지 않은 서비스 거부 공격(Denial of Service Attack) 기술 중 TCP의 신뢰성 및 연결 지향적 전송서비스로 종단간에 이루어지는 3-Way Handshake를 이용한 Syn Flooding Attack에 대하여 침입시도패킷 정보를 수집, 분석하고 퍼지인식도(FCM : Fuzzy Cognitive Maps)를 이용한 침입시도여부를 결정하는 네트워크 기반의 실시간 탐지 모델(Network based Real Time Scan Detection Model)을 제안한다.

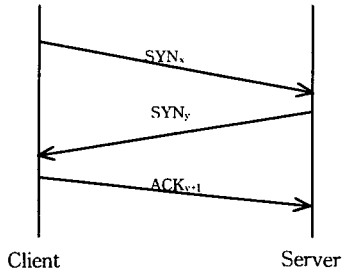
1. 서론

최근 네트워크 기술 발전으로 인하여 사회 전반에 걸쳐 인터넷 활용 의존성이 매우 높아지고 있는 추세이다. 이러한 네트워크 기술 발전의 반대급부로 악의적 목적을 둔 침입을 위한 서비스 거부 공격이 점차 늘어나는 추세이다. 여기서 서비스 거부 공격이란 일반적으로 시스템의 자원을 고갈 또는 마비시켜 서비스지원을 방해하는 일련의 침입을 위한 시도라고 볼 수 있다. 이들 중 가장 대표적인 서비스 거부 공격은 Syn Flooding이라 불리는 공격형태이다. Syn Flooding 서비스 거부 공격은 TCP 신뢰성 연결 지향적 전송서비스에 의하여 이루어지는데, 이러한 서비스 거부 공격은 인터넷환경에서 가장 많이 사용되어지는 TCP 기반의 프로토콜 서비스를 지원하는 시스템에 크게 영향을 미치고 있다. 이 공격은 TCP 프로토콜의 구조적 약점을 이용하는데 이를 해결하기 위해서는 프로토콜의 전반적 수정이외에는 현실적 정확한 해결책이 없다고 본다. 서비스 거부 공격은 크게 주요 파일을 훼손시켜 목적 시스템의 동작을 방해하는 우회적 서비스 거부 공격과 목적 시스템의 자원 및 네트워크 데이터전송을 위한 흐름제어 자원을 고갈시키는 공격으로 나눌 수 있다[1]. 현재 이를 해결하기 위한 여러 대안이 연구되어 지고 있다. 본 논문에서는 제2장에서 서비스 거부 공격에 대해서 살펴보며 이를 해결하기 위한 방안을 알아본다. 제3장에서는 이러한 방안 중 TCP의 3-Way Handshake 연결과정에서 발생하는 half open 연결상태를 실시간으로 탐지하여 퍼지 인식도(FCM:Fuzzy Cognitive Maps)를 이용하여 침입여부를 결정하는 탐지모델을 제안하며 마지막장에서 향후 연구방향과 결론을 제시한다.

2. Syn Flooding Attack

2.1 TCP Syn Flooding Attack

TCP Syn Flooding 공격은 앞에서 거론되었듯이 TCP의 약점을 이용한 공격형태이다. 일반적으로 TCP는 신뢰성 지향적 연결이므로 서버와 클라이언트간에 연결 설정에는 그림 1과 같은 '3 way Handshaking'라는 정상적 연결 흐름이 이루어진다.



[그림 1] 3 Way Handshake

만약, 여기서 클라이언트가 SYN_x를 요청하고 서버로부터 SYN_y와 ACK_{x+1}을 받은 후 ACK_{y+1}을 보내지 않으면 서버에서는 클라이언트로부터 응답이 올 것을 기대하고 반쯤 열린 'Half Open State'가 된다. 물론 얼마간 이런 상태가 유지된 후 다음 요청이 오지 않으면 해당 연결을 reset하게 된다. 이때 reset되기 전까지 메모리에는 backlog queue가 계속 쌓이게 되는데 이러한 reset이 되기 전에 지속적으로 이와 같은 요청이 아주 빠르게 이루어진다면 Syn Packet은 backlog queue에 쌓이게 되어 결국 메모리 용량을 넘어서게 되면 해당 포드에 대한 연결을 받아들일 수 없는 상태인 서비스 거부 상태가 된다.

2.2 해결 방안

2.2.1 backlog queue

실제 서비스 거부가 발생하는 원인으로 backlog queue에 더 이상 받아들일 수 있는 조건이 되지 않기 때문이다. 이를 해결하기 위해서 backlog queue 크기를 증가시켜주는 방법이다. 그러나 H/W 및 OS마다 서로 다른 메모리 용량과 backlog queue 크기가 할당되어 있어 정확한 크기증가 선정이 어려워진다. 예를 들어 Redhat Linux 7.x 시스템에서 메모리가 256MB인 경우 backlog queue 수치를 2048 이상으로 설정했을 때 TCP_SYNQ_HSIZE 와 tcp_max_syn_backlog의 수치를 조절하여야 한다. 그러나 이러한 대안은 지속적인 공격과 비용측면에서 볼 때 효율적이지 못하므로 적절한 대안이라 할 수 없다.

2.2.2 syncookies

syncookies에는 크게 Berkeley, Linux, Reset cookie가 있으며 '3 way handshake'에서 TCP 헤더의 Syn's sequence number, 소스 및 목적주소에 단방향 해쉬함수를 적용한 암호화 알고리즘을 이용한 방식으로 연결 설정이 정상적으로 이루어지지 않으면 더 이상 소스 경로를 따라 가지 않고
정상적 연결 요청에 대해서만 연결 설정을 하여 자원의 낭비를 줄이는 방법이다[2].

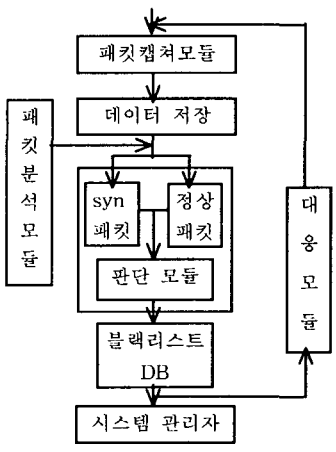
2.2.3 packet monitoring

라우터 및 게이트웨이를 통과한 후 시스템 접근에 앞서서 모니터링을 하는 방법으로써 들어오는 패킷을 잡아 분석하여 'half open state'를 요청하는 포트 및 IP Address를 탐지하여 RST 등으로 연결 해제하는 방법이다. 본 논문에서는 제안하는 모니터링을 통한 탐지 또한 이 범주에 속한다[3].
이외에도 임의의 라우팅 테이블을 변경하여 트래픽이 전달되지 못하도록 ICMP redirects를 허용하지 않는 방법과 IP 소스 라우팅을 사용하여 목적지의 경로를 지정하여 믿을 수 있는 IP로 위장하지 못하도록 하는 소스 라우팅 패킷 허용 불능법 등이 있다.

3. 탐지모델제안

3.1 탐지모델구조

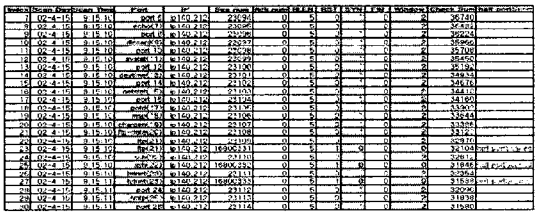
본 논문에서 제안하는 실시간 탐지모델은 256MB 메모리의 펜티엄4 리눅스 시스템과 공격시스템 3대를 연결한 시험망에서 테스트 한 것이다.



[그림 2] 탐지 모델 구조

전체적인 탐지모델구조는 그림2에서 보듯이 들어오는 패킷을 잡아 분석하는 모듈과 DB저장, 그리고 'half open state'을 판단하는 판단모듈로 되어 있으며 추가로 대응모듈이 있다.

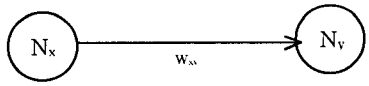
여기서 패킷캡쳐모듈은 promiscuous mode에서 데이터링크층의 패킷을 캡쳐한 것이며 이를 파싱하여 로컬포트, IP Address, Sequence Num, 윈도우 크기 및 공격시간 등으로 저장시키고 패킷분석모듈을 통하여 Syn패킷과 정상패킷으로 구분하여 1차 'half open state'를 탐지하게 된다. 여기서 탐지된 IP Address는 퍼지 인식도를 이용한 판단모듈을 통하여 블랙리스트 DB에 저장되고 시스템관리자에게 통보 하게된다. 재차 공격시에는 블랙리스트 DB와 비교하여 공격을 탐지하고 대응모듈을 가동하게 된다. 그림 3은 DB에 저장된 탐지로그 항목이며 'SYN'과 'RST'의 수치가 각각 변경된 것과 해당 Sequence Number 및 윈도우 크기가 변경된 것을 알 수 있는데, 이러한 항목들의 패턴을 감시하면 실시간으로 'half open state'를 탐지 할 수 있다.



[그림 3] 탐지로그항목

3.2 판단모듈

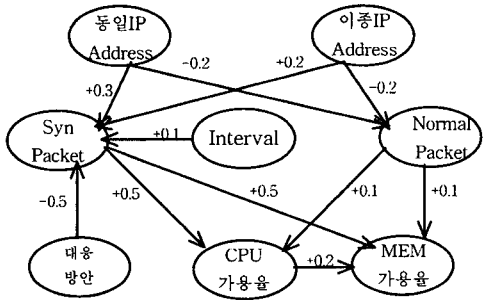
판단모듈은 퍼지 인식도의 Causal knowledge reason을 이용하여 지능적 판단모듈구조를 설계하였다. 그림 4는 퍼지 인식도를 표현한 것으로써 각 노드와 노드사이의 가중치(링크)가 W_{xy}=0인 경우에는 각 노드사이에는 아무런 관련이 없는 것을 의미하며 W_{xy}≠0 경우에는 그림4와 같은 의미를 부여한다[4].



W_{xy}>0 ; N_x 수치 증가로 인한 N_y 수치 증가인 경우
W_{xy}<0 ; N_x 수치 증가로 인한 N_y 수치 감소인 경우

[그림 4] 퍼지인식도

판단모듈에서 여러 가변 요소 중 어떤 요소에 의존성을 부여함으로써 가장 최적의 탐지를 할 수 있는 것이 가장 큰 관건이다. 그뿐만 아니라 탐지한 IP Address를 침입시도로 간주하고 블랙리스트 DB에 저장하여야 하는지도 결정하여야 한다. 퍼지인식도는 이러한 여러 가변 요소를 적용하여 최적의 판단을 내리게 한다. 그림5는 가변요소를 적용한 판단모듈의 퍼지인식도를 나타낸 것으로써, 판단모듈에 의존성을 갖는 가변요소로 IP Address의 동일성 여부와 'half open state'의 시간간격 그리고 각 프로세서의 CPU가용율과 메모리가용율 및 판단모듈 후 재차 공격시 대응모듈의 처리로 인한 공격성 IP Address에 대한 Syn 패킷 조절을 둘 수 있다.



[그림 5] 가변요소 적용한 퍼지인식도

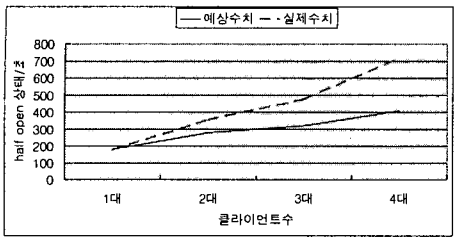
가변요소를 노드(N_x)와 다음 노드(N_y)에 두고 두 노드의 링크인 가중치(W_{xy})를적용하는 것이다. 예를 들면, Syn Packet과 CPU가용율에서는 Syn Packet의 용량이 증가할수록 CPU가용율이 증가하므로 이때 가중치는 0보다 크게 된다. 이때 임의의 노드에 가해지는 수치는 노드와 가중치를 연결한 네트워크를 통과할수록 그리고 반복횟수에 따라서 달라지게 된다. 이를 수식화 하면 다음과 같다.

$$N_k(t_{n+1}) = \sum_{i=1}^{n} W_{ik}(t_n) N_i(t_n)$$

단, 가중치(W_{xy})의 증감부호는 다음 노드에 미치는 영향에 따라서 결정을 내렸으며 수치는 의미 있는 규칙기반에 의한 수치를 연구 진행중이며 현재로는 반복적 실험에 의한 경험치를 사용하였다.

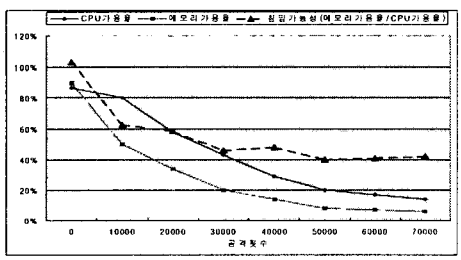
그림 6은 클라이언트에서 초(sec)당 SYN Flooding Attack인 경우 모니터링의 탐지율을 보여주고 있다. 테스트 결과 1대의 Client인 경우 180 탐지/sec 의 'half open state'를 탐지하는 것으로 나타났으며 3대 Client인 경우에는 320 탐지/sec 정도로 나타났다[5]. 그림 6와 같이 탐지율은 클라이언트 수가 증가할수록 예상탐지수치보다 조금씩 떨어지는 것을 볼 수 있는데, 이는 실시간으로 패킷을 캡쳐하고 파싱하고

분석 그리고 저장 및 탐지결정에 따른 CPU의 부하와 메모리 가용량 부족으로 인한 패킷의 폐기 및 캡쳐 도중에 발생한 손실로 보아야 한다.



[그림 6] 클라이언트 수에 대한 모니터링 탐지율

최종판단모듈은 메모리가용율과 CPU가용율로 결정하였다. 여기서 이를 판단의 중요한 근거로 선정한 이유는 'half open state'를 이용한 Syn Flooding Attack은 시스템 부하 및 자원 고갈을 이용하는 것이므로 판단의 최종 가변요소로 선정하는 것이 가장 타당하다고 본다. 아울러 시스템의 부하 및 자원고갈의 데드라인과 임계값을 설정하는 과정 역시 중요한 요건으로 보여지는데 이는 어느 정도의 'Syn Flooding Attack' 에 대한 실시간 대응할 수 있는 시간 및 침입시도가 목적이 아닌 단지 테스트 또는 예상치 못한 클라이언트의 부주의한 실수를 감안한 설정이다. 그러나, 이러한 공격에 대해서도 간과할 수 없으므로 탐지모델은 로그항목을 분석하고 DB에 저장하여 재차 이러한 공격에 대한 충분한 대비를 할 수 있게끔 되어 있다. 그림 7은 CPU가용율에 대한 메모리가용율를 나타낸 것으로써, 여기서 표현되는 수치가 50%이하인 경우에는 메모리의 가용용량을 벗어난 것으로 간주하여 위험상태로 결정되어 진다.



[그림 7] 공격횟수에 대한 하드웨어 가용율

이때, 40%-60% 정도의 가용용량이고 메모리 가용율이 CPU 가용율 보다 낮을 경우, 증가형태의 가중치를 두어 판단하였는데 이는 메모리가 backlog queue 수치와 비례관계이므로 메모리의 가용율이 낮음은 backlog queue가 포화상태로 가고 있다는 뜻을 의미하기 때문이다.

4. 결론

본 논문에서는 SYN Flooding Attack에 대해서 살펴보았으며 해결책으로 여러 대안 중에서 패킷을 분석하여 침입시도탐지기능을 수행하는 네트워크 기반 탐지모델을 제안하고 시험망에서 테스트하였다. 여기서, 탐지성능을 좌우하는 요소들간의 상호 관계로부터 퍼지인식도를 이용한 침입시도 여부를 판단하였는데, 퍼지인식도에서 가장 중요한 가중치를 결정하는 수치에 대해서는 연구가 진행중이나 현 시점에선 반복 실험치에 근거를 하였다. 아울러, 침입시도 여부를 명확히 판단 할 수 없는 하드웨어(CPU와 메모리) 가용용량 구역대(40%-60%)에서는 실시간 처리 가능한 데드라인 시간과 임계값을 설정하여 침입여부를 결정하는 실험도 병행하여 계속 진행 중이다. 시험망에서 테스트를 한 결과 하나의 시스템에서 실시간 처리로 인하여 시스템에 하드웨어 부하가 시스템의 영향을 미침을 알 수 있었다. 향후 연구과제로 패킷캡쳐와 분석 및 판단모듈을 각각의 시스템에 두어 테스트를 하고 판단모듈에 의한 학습된 지능형 대응모듈기능이 추가된 지능적 탐지 및 방지시스템 형태로 연구해 나갈 예정이다.

참고문헌

[1] Computer Emergency Response Team, "TCP Syn Flooding and IP Spoofing Attacks," CERT Advisory: CA, 96-21, 1996.

[2] Aman Garg and A.L.Narasimha Reddy, "Policy Based End Server Resource Regulation," IEEE/ACM Transactions on Networking, Vol. 8, No. 2, pp. 146-157, 2000.

[3] C. L. Schuba, I. V. Krsul, M. G. Khun, E. H. Spaford, A. Sundram, and D. Zamboni, "Analysis of a denial of service attack on tcp," 1997 IEEE Symposium on Security and Privacy, 1997.

[4] Lee, K.C and H.S. Kim, "A causal knowledge driven inference engine for expert system," In Proceedings of the annual Hawaii international conference on system science, pp. 284-293, 1998.

[5] S.Y. Lee and Y.S. Kim, "A RTSD Mechanism for Detection of DoS Attack on TCP Network," Proceedings of KFIS 2002 Spring Conference, pp. 252-255, 2002.