HTTP 트래픽 기반의 비정상행위 탐지 시스템

김효남[0] 장성민 원유헌
청강문화산업대학[0], 홍익대학교
hnkim@ck.ac.kr[0], {smjang, yhwon}@cs.hongik.ac.kr

HTTP Traffic Based Anomaly Detection System

Hyonam Kim[0], Sungmin Jang, Yuhun Won
ChungKang College[0], Hongik University

요 약

최근 인터넷 공격은 웹 서비스 환경에서 다양한 공격 유형들이 인터넷상에서 나타나고 있는 실정이다. 특히 인터넷 웜이나 기타 알려지지 않은 공격이 대중을 이루고 있어 기존의 정보 보호 기술로는 한계에 다다르고 있으며 이미 알려진 공격을 탐지하는 오용탐지 기술로는 적절하게 대응하기 어려워진 상태이다. 또한, 웹 서비스 이용이 확대되고 사용자 요구에 맞게 변화하면서 인터넷상의 노출된 웹 서비스는 공격자들에게 있어 주공격 대상이 되고 있다. 본 논문에서는 웹 기반의 트래픽 유형을 분석하고 각 유형에 따른 이상 징후를 파악할 수 있는 비정상 탐지 모델을 정의하여 정상 트래픽 모델과 비교함으로써 현재 트래픽의 이상 정도를 평가하고 탐지 및 규칙생성, 추가하는 HTTP 트래픽 기반의 비정상행위 탐지 시스템을 설계하고 구현하였다.

1. 서 론

인터넷이 급속하게 보급되면서 웹 서비스 이용이 확대되고 사용자 요구에 맞게 스트리밍이나 P2P와 같은 새로운 애플리케이션과 같은 다양한 서비스 유형들이 제공되고 변화하고 있다. 이러한 서비스들은 대량의 네트워크 트래픽을 유발시키는 특징을 갖고 있다. 이와 더불어 최근에는 인터넷 웜이나 기타 알려지지 않은 공격들이 대량의 네트워크 트래픽을 발생시켜 네트워크 가용성을 심각한 수준으로 떨어뜨리는 위협이 빈번하게 발생하고 있다. 대량 트래픽에 의한 네트워크 가용성 저하 문제를 최소화하기 위해서 미리 위협 요인을 탐지하여 적절한 조치를 취해야 하는 네트워크 탐지 시스템이 요구된다.. 기존의 공격 탐지 기술로는 이미 알려진 공격기법들을 시그너처(Signature)로 가지고 있어 공격을 탐지하는 오용탐지(Misuse Detection)기반이 주류를 이루고 있다[1]. 그러나 오용탐지 기술만으로 최근 유형의 공격들을 차단하기에는 한계가 있다. 최근에는 오용탐지 방법을 보완하기 위하여 트래픽의 추이를 분석하여 정상적이지 않은 트랙픽 즉, 비정상행위를 탐지하는 방법들로 활발하게 연구되고 있다[2].

본 논문에서는 HTTP 트래픽 정보를 수집하고 분석하여 미리 준비된 정상 트래픽 패턴과 비교함으로써 현재 트래픽의 위험 정도를 평가하여 경고하는 웹 침입 탐지 시스템을 설계 및 구현하였다. 2장에서는 제안된 웹 비정상행위 탐지시스템과 탐지 모델을 설명하고, 실험방법과 결과를 분석한다. 3장에서는 결론과 향후 연구 과제를 기술한다.

2. 본 론

2.1 트래픽 분석시스템(Traffic Analysis System)

인터넷이 규모와 복잡도 측면에서 급속하게 성장하면서 네트워크를 통해 전달되는 데이터 량이나 데이터의 복잡도 또한 증가하고 있으며 악성코드인 웜 등은 트래픽 분석에 대한 특징이나 이해, 그리고 모델링하는데 어려움을 초래하고 있다. 따라서 인터넷 상의 트래픽의 모델링은 실험실이나 테스트베드 수준의 제한된 데이터 량으로 시뮬레이션 하는 것이 일반적인 추세이다[2].

네트워크 트래픽 감시 및 분석 정보는 ①네트워크 상태와 문제점을 파악, ②트래픽이 증가하는 경우 원인을 분석, ③ 네트워크 회선 계획 등으로 필요하다. 하지만 기존의 네트워크 분석의 문제점으로서 ①패킷 손실로 인한 트래픽 분석의 부정확성, ②대용량 트래픽 처리의 한계, ③장기간 트래픽 감시 및 분석을 지원하지 않음, ④주로 잘 알려진 서비스만 분석하는 등의 문제점이 있다. 최근 몇 년 동안 여러 연구자들이 인프라 구조를 보호하는 많은 방어 장비와 방법에 대해 연구하고 있지만 실세계 데이터 부재로 인해 장비의 효용성과 다양한 분야에 쉽게 사용할 수 있도록 평가하는 작업이 계속되고 있다[3]. 또한 다양한 분야의 네트워크 행위를 분석하기 위하여 한 순간의 네트워크 행위를 포획하기 보다는 진행되는 네트워크 트래픽 데이터를 수집하는 것을 시도하고 있다[4]. 일반적으로 수동적인 가장 인기 있는 측정 시스템으로는 SNMP, Tcpdump, NetFlow 등이 있다[5][6].

2.2 네트워크 트래픽 추이 형태

네트워크 트래픽 수집 분석도구로는 웹 기반의 모니터링 및 분석 애플리케이션인 Ntop을 비롯하여 Tcpdump, Ethereal,

FlowScan, CoralReef, PMA, IPMON, Snort등이 있다. 이 중에서 Snort는 실시간 트래픽 분석과 패킷 로깅을 수행하는 시그너처 기반, 규칙 기반의 네트워크 침입 탐지 시스템으로 오픈 소스 기반의 도구이다. Snort의 주요한 세 가지 기능은 다음과 같다. 첫째 Snort는 Tcpdump와 같은 패킷 스나이퍼로 바로 사용할 수 있다. 둘째 네트워크 트래픽 디버깅에 유용한 패킷 로거(logger) 기능이 있다. 마지막으로 완벽한 네트워크 침입 탐지 시스템(NIDS) 기능을 가지고 있다. 그림 1은 특정 포트별 네트워크 트래픽 추이를 보여주는 그래프이다.

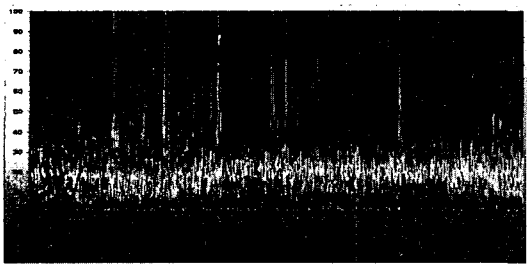

그림 1 포트별 네트워크 트래픽 추이 그래프

그림 1에서 보는 것과 같이 일반적인 네트워크 트래픽은 간헐적으로는 최고 임계치에 근접하기는 하나 전반적으로 볼 때 유순한 곡선을 나타내는 형태로 파악된다. 물론 업무량이 증가하거나 특정한 응용을 사용할 경우에 있어서는 갑작스런 네트워크의 증가를 보이기도 하나 이는 일정시간이 경과하면 다시 정상적인 행태로 복귀하는 것을 알 수 있다. 그림 2에서는 이미 알려진 포트 정보에 대해서는 Snort에서 필터를 사용하여 정상적이라고 판단되는 트래픽은 제거한 후에 네트워크 트래픽 추이를 보여주는 그래프이다.

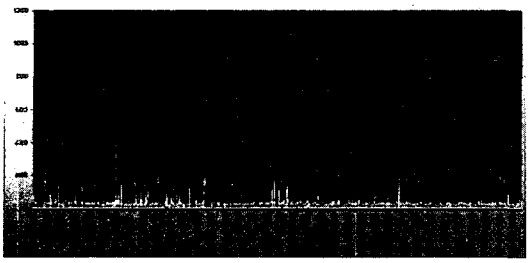

그림 2 필터를 통한 네트워크 추이 그래프

그림 2의 그래프에서는 뚜렷하게 이상 징후에 대한 내용을 확인할 수 있다.

2.3 비정상행위 탐지 시스템의 구조

본 논문에서 제안하는 비정상행위 탐지시스템은 그림 3과 같은 상세 구조로 구성된다. 그림 3에서 보는바와 같이 데이터 생성부, 데이터 수집부, 데이터 분석부 그리고 데이터 표현부로 구성된다.

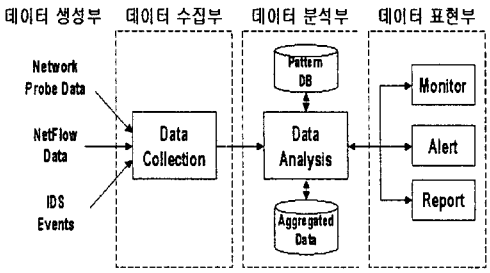

그림 3 비정상행위 탐지시스템 구조

비정상행위 탐지시스템의 데이터 생성부는 Network Probe를 통한 실시간 트래픽 데이터, 라우터 혹은 NetFlow 호환 시스템을 통한 NetFlow 데이터, 그리고 IDS 나 IPS 등을 통한 공격 이벤트 데이터를 생성한다. 데이터 수집부는 NetFlow 데이터와 공격이벤트 데이터를 수집하여 저장한다. 1차 수집된 Raw 데이터로부터 전체 트래픽(inbound /outbound), 프로토콜별, 서비스별, 패킷크기별, 출발지IP 및 목적지IP별 트래픽 량(flows, packets, bytes)을 추출하여 로그로 저장한다. 본 연구에서는 NetFlow 데이터를 수집한다. 데이터 분석부는 Snort의 패킷분석 모듈을 수정하여 구현하였으며, 수집한 데이터로부터 트래픽의 추이분석을 수행하고 프로파일링 기간에 정의된 정상적인 트래픽 패턴과 비교분석을 통해 트래픽의 비정상 여부를 평가한다. 데이터 표현부는 실시간 트래픽 모니터링, 통계 및 분석, 비정상행위 분석, 공격이벤트 분석, 라포팅 기능을 수행한다.

2.4 네트워크 트래픽 추이분석

위험 지표는 네트워크 트래픽의 이상 징후에 따른 위험의 정도를 나타내는 측정값이다. HTTP 트래픽에 대한 위험 지표는 트래픽 추이 값인 t_p으로 나타내며, 다음 식(1)에서와 같이 계산한다.

$$t_p = log(\frac{R^* r_p}{r^* R_p}) \qquad (1)$$

R = 최근 한달(4주)간 트래픽량(flows, packets, bytes)

r = 하루동안 트래픽량

R_p = 특정 포트에 대한 최근 한달(4주)간 트래픽량

r_p = 특정 포트에 대한 하루동안 트래픽량

트래픽 추이 값(t_p)은 특정 서비스(포트)의 트래픽이 급격히 증가할 경우, 양(+)의 값을 가지며, 반대의 경우는 음(-)의 값을 가진다. 트래픽 량의 변화가 없을 경우 0에 가까운 값으로 나타나고, 트래픽의 급격한 변화가 있을 때 증감하므로, 트래픽 추이 값을 통해 네트워크 트래픽의 이상 징후를 탐지할 수 있다.

위험지표는 네트워크 트래픽의 이상 징후에 따른 위험의 정도

를 나타내는 측정값이다. 위험지표를 구하기 위하여 고려한 측정대상 및 측정값은 표 1과 같다.

표 1 측정값과 위험지표

대상	측정값	위험지표
전체 트래픽	인바운드 및 아웃바운드 트래픽의 량	t_i, t_o
서비스(포트)	서비스별(목적지 포트) 트래픽의 량	t_p
프로토콜	프로토콜별(TCP, UDP, ICMP 등) 트래픽의 량	t_P
패킷 사이즈	패킷 사이즈별 분포(비율)	t_{ps}
출발지 주소	출발지 주소별 트래픽의 량, 목적지 주소의 수, Top 10 목적지 주소별 트래픽의 량	t_s
공격 이벤트	공격 유형별 이벤트 수(공격시도횟수)	t_a
목적지 주소	목적지 주소별 트래픽의 량, 출발지 주소의 수, Top 10 목적지 주소별 트래픽의 량	t_d

위험경보는 트래픽 폭주 등으로 실제 위험상황이 발생하기 전에 사전대응(주의, 경계)할 수 있도록 알리는 행위 혹은 신호이다. 위험경보는 측정한 위험지표를 종합적으로 분석하여 최종적으로 4단계(정상, 주의, 경고, 위험)를 가지도록 한다. 위험경보 단계는 각 위험지표별 가중치를 부여하고, 임계값을 초과하는 위험지표와 정도, 빈도, 지속시간 및 상황지수를 고려하여 결정한다.

2.5 실험 결과 및 분석

본 연구는 특정 웹서버의 일정기간 동안 정상적인 HTTP 트래픽의 추이분석을 통하여 트래픽 추이 값(t_p)을 구하였다. 특정 날짜, 특정 시간의 HTTP 트래픽을 조작하여 제안된 비정상행위 탐지 시스템으로 일정시간 전송하여 비정상 트래픽에 대한 탐지 여부를 실험하였다. 그림 4는 콘솔에서 탐지된 결과를 보여주고 있으며, 관리자는 탐지된 결과에 대하여 정해진 네트워크 정책을 기반으로 하여 프로파일링, 규칙추가, 바이 패스와 같은 기능을 수행할 수 있도록 하였다.

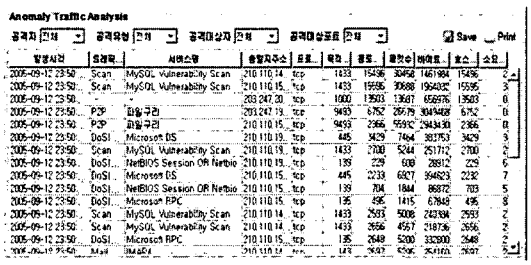

그림 4 콘솔로 출력되는 오용탐지 보고

프로파일링 기능은 해당 시간의 프로파일링 데이터베이스를 현재의 트래픽 추이 값(t_p)으로 갱신하고 적용하는 과정이며, 규칙 추가 기능은 해당 트래픽에 Snort 규칙을 추가하여 지속적으로 트래픽을 관리할 수 있도록 한다. 바이 패스는 해당 시간의 프로파일링 과정에 관여치 않고 탐지 결과를 무시하는 것을 의미한다.

웹 트래픽 분석 실험 결과 시간대별 트래픽 특성을 프로파일링 함으로써 미세한 트래픽 변화에 대하여 쉽게 탐지할 수 있음을 알 수 있었다, 이를 바탕으로 웹 기반 침입 탐지 시스템의 오탐지(false-positive) 및 미탐지(false-negative)를 감소시킬 수 있으며, 특정 웹서버의 보다 정확한 프로파일링 수행한다면 트래픽의 이상 징후를 사전에 감지하여 경보할 수 있는 기능을 제공할 수 있을 것으로 기대된다.

3. 결 론

본 논문에서는 HTTP 트래픽에 대한 정보를 수집, 분석하여 미리 준비된 정상 트래픽 패턴과 비교할 수 있는 트래픽의 파라미터 패턴 모델을 제안하였고, 이를 사용하여 비정상적인 HTTP 트래픽에 대한 리포팅기능, 규칙 생성을 제공하여 인터넷 공격 위협으로부터 네트워크 가용성을 확보할 수 있는 HTTP 트래픽 기반의 비정상행위 탐지 시스템을 설계하고 구현하였다. 향후에는 특정 웹서버가 가지는 HTTP 트래픽의 특정 URI정보만을 정규화 하는 연구가 필요하며, 웹서버에 대한 취약성 분석을 하기위한 비정상행위 트래픽 패킷의 가상 공격을 통한 탐지 규칙의 자동 생성 및 추가 방법에 대한 연구가 요구된다.

[참고문헌]

[1] U. Lindqvist and P.A. Porras. Detecting Computer and Network Misuse with the Production-Based Expert System Toolset (P-BEST). In IEEE Symposium on Security and Privacy, pages 146-161, Oakland, California, May 1999.

[2] Christopher Kruegal, Giovanni Vigna, William Robertson, "A multi-model approach to the detection of web-based attack", Computer Network: The International Journal of Computer and Telecommunications Networking, Jan 2005

[3] Jelena Mirkovic, Sven Dietrich, David Dittrich, and Peter Reiher. Internet Denial of Service: Attack and Defense Mechanisms. Prentice-Hall, Upper Saddle River, NJ, 2004.

[4] Alefiya Hussain 외 5인, Experience with a Continuous Network Tracing Infrastructure, ACM SIGCOMM'05 Workshops, 2005

[5]C. Fraleigh, C. Diot, B. Lyles, S. Moon, P. Owezarski, D. Papagiannaki, and F. Tobagi. Design and deployment of a passive monitoring infrastructure. Lecture Notes in Computer Science, 2170:556-567, 2001

[6]Anukool Lakhina 외 5인, Structural Structural Analysis of Network Traffic Flows, ACM SIGMETR ICS/CS/Performance 2004