시간지연 신경망을 이용한 침입탐지 시스템

°강병두, 문채현, 정성윤, 박수범, 김상균
인제대학교 전산학과

An Intrusion Detection System Using Time Delay Neural Network

Byoung-Doo Kang, Chae-Hyun Moon, Sung-Yoon Jung, Soo-Beom Park, Sang-Kyoon Kim
Department of Computer Science, Inje University

요 약

기존의 규칙기반 침입탐지 시스템은 사후처리식 규칙 추가로 인하여 새로운 변종의 공격을 탐지하지 못한다. 본 논문에서는 규칙기반 시스템의 한계점을 극복하기 위하여, 시간지연 신경망(Time Delay Neural Network; 이하 TDNN) 침입탐지 시스템을 제안한다. 네트워크상의 패킷은 바이트 단위를 하나의 픽셀로 하는 0에서 255사이 값으로 이루어진 그레이 이미지로 볼 수 있다. 이러한 연속된 패킷이미지를 시간지연 신경망의 학습패턴으로 사용한다. 정상적인 흐름과 비정상적인 흐름에 대한 패킷이미지를 학습하여 두 가지 클래스에 대한 신경망 분류기를 구현한다. 개발하는 침입탐지 시스템은 알려진 다양한 침입유형뿐만 아니라, 새로운 변종에 대해서도 분류기의 유연한 반응을 통하여 효과적으로 탐지할 수 있다.

1. 서론

인터넷 확산은 다양한 인터넷 서비스를 가능하게 하여 정보화 사회에 많은 이점을 가져왔다. 그러나, 이러한 이점 못지 않게 해킹을 통한 부당한 이득을 얻으려는 사이버 범죄가 증가하고 있으며, 사용법이 쉬운 해킹 프로그램의 소스 공개로 인해 다양한 해킹 방법들이 생겨나고 있어 인터넷 발전을 저해하고 있다. 야후와 같은 지명도 있는 인터넷 사이트들도 해킹 피해를 입고 있다. 그러므로, 안전한 인터넷 서비스를 위해서 침입을 탐지하고 이를 알려줄 침입탐지 시스템이 필요하다[1,2].

침입이란 비 인가된 사용자가 자원의 무결성(integrity), 기밀성(confidentiality), 가용성(availability)을 저해하는 일련의 행동들과 보안 정책을 위반하는 행위를 말한다. 이러한 침입 행위를 탐지하는 시스템을 침입탐지 시스템이라 한다.

네트웍 기반 침입탐지 모델에는 알려진 침입행위를 이용하여 침입을 탐지하며 정해진 모델과 일치하는 경우를 침입이라 하는 오용 침입탐지 기법(Misuse Detection)과 사용자의 패턴을 분석하여 입력패턴과 비교하여 정해진 모델을 벗어나는 경우를 침입으로 간주하는 비정상 침입탐지 기법(Anomaly Detection)이 있다[3].

오용 침입탐지 시스템 중 규칙기반 침입탐지 시스템은 사후처리식 규칙추가로 인하여 새로운 변종의 공격을 탐지할 수 있는 능력이 없다. 즉, 특정 공격형태가 발견되면 형태에 따른 패킷의 흐름을 규칙화하고 시스템에 추가하는 형식으로 구현하는데, 변화해나가는 다양한 공격형태를 효과적으로 탐지할 수 없다. 대표적인 예로, 가장 널리 알려진 공개용 침입탐지 시스템인 snort는 1000여개의 규칙을 가지고 있음에도 불구하고 명세화된 규칙을 조금이라도 우회하는 공격에 대해서는 감지하지 못하는 문제점을 가지고 있다. 뿐만 아니라, 기존의 네트워크 IDS는 스니핑(sniffing) 속도의 한계로 인해 네트웍크의 모든 트래픽을 분석하는 것은 불가능하다[4]. 따라서 비정상침입탐지기법

의 기본 방향을 특정 침입 유형에 종속적인 패킷의 흐름에 주안점을 둘게 아니라, 정상적인 패킷의 흐름과 비정상적인 패킷의 흐름이라는 두 가지 클래스에 대한 분류문제로 정의할 필요가 있다.

네트워크상의 패킷은 바이트 단위를 하나의 픽셀로 하는 0에서 255 사이 값으로 이루어진 그레이 이미지로 볼 수 있다. 패킷의 흐름을 시간적인 측면에서 나타내었을 때, 연속되는 여러 패킷은 일련의 연속된 이미지로 볼 수 있다. 본 연구에서는 이러한 특징을 가지는 패킷이미지들을 신경망 학습을 위한 0에서 1 사이값의 학습패턴으로 만든다. 이를 시간적 특성을 고려할 수 있는 시간지연 신경망의 입력값으로 사용하여, 정상적인 흐름과 비정상적인 흐름에 대한 패킷 이미지를 학습하는 시간지연 신경망 분류기를 구현한다.

개발하는 침입탐지 시스템은 알려진 다양한 침입 유형뿐만 아니라, 새로운 변종의 침입에 대해서도 분류기의 유연한 반응을 통하여 효과적으로 탐지함으로써 규칙기반 시스템의 한계점을 극복하고 인터넷을 통해 들어오는 다양한 패킷들을 실시간으로 탐지할 수 있다.

2. 시간지연 신경망을 이용한 침입탐지시스템

2.1 침입탐지 시스템 전체구성

본 연구에서 제시하는 시간지연 신경망을 이용한 침입탐지 시스템의 전체 흐름도는 그림 1과 같다.

ⓐ 패킷 수집기는 libpcap 라이브러리를 이용하여 패킷을 수집하고 시간, 길이, 로우(raw) 데이터 정보를 가지는 패킷 구조체를 반환한다.

ⓑ 패킷 수집기에서 반환받은 패킷 구조체의 정보를 통해 그레이 이미지 패킷으로 변환한다.

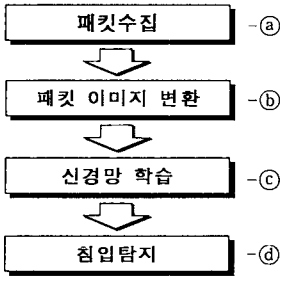

그림1. 침입탐지 시스템 흐름도

ⓒ 신경망 학습을 위한 패턴으로 전환된 후 시간지연 신경망 학습을 위해 사용된다.

ⓓ 학습이 완료된 침입탐지 엔진은 실시간 네트웍 침입탐지를 하게되고, 그 결과를 정상 패킷과 비정상 패킷으로 구별한다. 패킷을 연속한 그레이 이미지로 처리함으로써 공격 이미지 일부분만을 감지하더라도 이와 가장 유사한 패턴으로 구별해 낼 수 있다.

2.2 패킷수집

패킷 수집을 위해 사용한 libpcap 라이브러리는 Berkeley 대학에서 개발한 것으로 시스템에 독립적으로 사용자 레벨에서 개발한 패킷 수집을 효과적을 할 수 있도록 만든 공용 라이브러리이다. 패킷을 포착하려는 시스템에서 Promiscuous 모드의 상위수준 인터페이스를 제공한다[5]. 네트워크 패킷은 libpcap 라이브러리에서 제공하는 인터페이스를 사용하여 접근할 수 있는데, 이 라이브러리는 거의 모든 유닉스 시스템에서 사용가능하며 tcpdump와 같은 네트워크 모니터링 도구가 이 패킷 수집 라이브러리를 사용하고 있다[6].

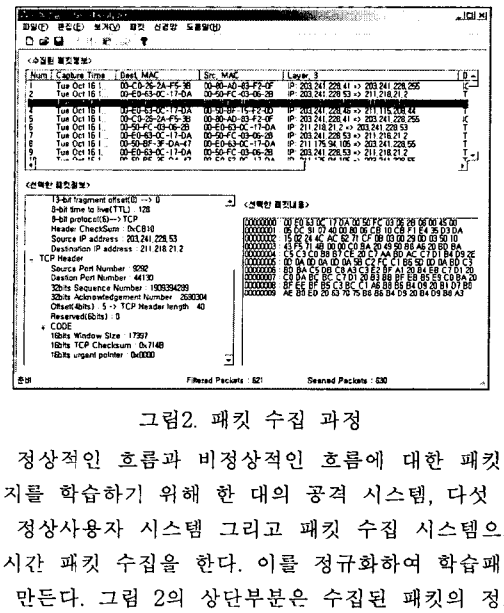

그림2. 패킷 수집 과정

정상적인 흐름과 비정상적인 흐름에 대한 패킷이미지를 학습하기 위해 한 대의 공격 시스템, 다섯 대의 정상사용자 시스템 그리고 패킷 수집 시스템으로 실시간 패킷 수집을 한다. 이를 정규화하여 학습패턴을 만든다. 그림 2의 상단부분은 수집된 패킷의 정보로써 수집 시간과 패킷헤더의 정보를 나타낸다. 하단부분중 왼쪽은 선택한 패킷의 세부 정보를 보여주고, 오른쪽은 선택한 패킷의 내용을 16진수의 바이트 코드값으로 나타낸다.

2.3 패킷 이미지 변환 및 학습패턴 생성

네트워크상의 패킷의 최대 허용길이(Maximum Transfer Unit)는 한 프래임당 1500바이트를 넘지 않는다. 그리고, 하나의 패킷은 1비트 단위로써는 0과 1로 이루어진 바이너리 이미지이지만 1바이트를 이미지의 한 픽셀로 나타내면 그레이 이미지가 된다. 그림 3은 패킷의 그레이 이미지로써 세로줄은 하나의 패킷을 나타내고 그 길이는 60바이트를 나타낸다. 그리고 오른쪽으로는 순차적으로 60개의 패킷이 나열되어 있다.

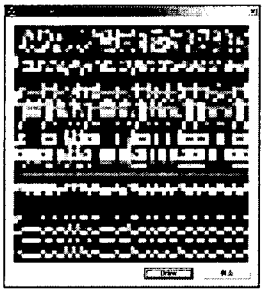

그림3. 패킷 이미지

그림 4는 패킷의 학습패턴 변환을 보여주고 있다. 신경망 학습을 위해 1바이트를 1비트의 0과 1사이값으로 바꾸기 위해 255로 나눈다. 결국 패킷은 신경망 학습을 위한 0에서 1사이의 그레이 이미지가 된다. 이러한 연속된 패킷이미지를 시간지연 신경망의 입력으로 사용한다.

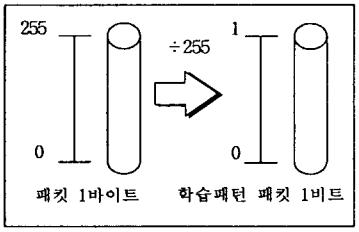

그림4. 패킷의 학습패턴 이미지 변환

2.4 침입탐지 시스템의 TDNN 구성 및 학습

신경망의 기본단위인 노드들은 입력에 연결강도(weight)가 곱해진 합을 구하여 이를 시그모이드(sigmoid)함수를 통과시킨 후 그 출력을 다음 층에 전달하는 역할을 한다. 동적 요소인 시간 지연 요소(time-delay)를 첨가한 시간 지연 신경망은 그림 5와 같이 변형된다.

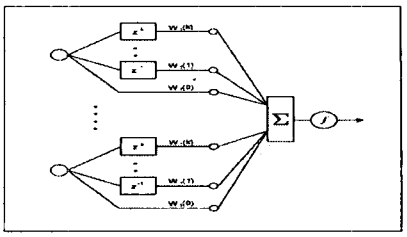

그림5. TDNN의 구성단위

이렇게 동적 요소인 시간 지연 요소를 첨가하여 시간지연 신경망은 현재의 입력과 과거의 입력을 연관시켜 비교할 수가 있게 된다[7].

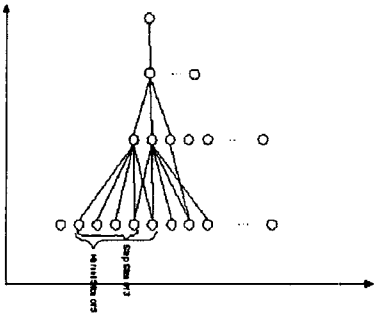

그림6. TDNN의 단면구조

본 연구에서 실험하는 시간지연 신경망을 이용한 침입탐지 시스템의 전체적인 단면 구조는 그림 6과 같다. 1개의 입력층, 2개의 은닉층, 그리고 1개의 출력층으로 일반적인 TDNN 구조와 동일하게 이루어져 있다. 그리고, 실제 사용한 각 층의 노드 구성을 보면, 입력층은 60×60개의 노드들로, 은닉층은 각각 30×30, 15×15개의 노드들로 구성하였으며, 마지막으로 출력층은 1개의 출력 노드를 가지도록 구성된다. 그리고 Kernal Size와 Step Size는 표 1과 같이 설정한다. 표 1은 시간지연 신경망을 이용한 침입탐지 시스템(TDNN)에서 각 계층에 대한 세부 구성이다.

표1. 침입탐지 시스템의 TDNN 구성

	Spatial Size	Kernal Size	Step Size	노드 개수
입력층	60	0	0	60 × 60
은닉층 1	30	4	2	29 × 30
은닉층 2	15	5	2	13 × 15
출력층	1	13	1	1 × 1

2.5 TDNN의 학습

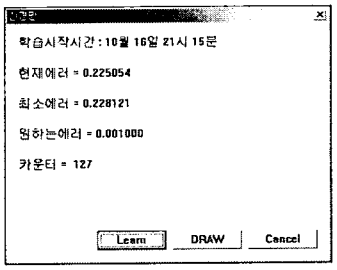

그림7. TDNN 학습

제안한 TDNN을 이용한 침입탐지 시스템은 펜티엄 PC, Windows 환경에서 Visual C++로 구현한다. 그림 7은 TDNN의 학습과정을 보여주고 있다. 구성된 시간 지연 신경망을 학습하기 위해 정상 패킷 1000개와 비정상 패킷 1000개를 수집하여 학습 패턴으로 사용한다. 원하는 에러값은 0.001로 지정하여 학습한다. 학습이 완료되면 가중치(weight) 값들이 저장된다.

3. 실험 및 결과

학습된 침입탐지 시스템의 성능을 평가하기 위해 추출된 각 패킷 데이터 중 학습에 사용된 데이터와 학습에 사용되지 않은 데이터로 구분하여 실험하였다. Teardrop 공격패킷 1000개, 정상패킷 1000개를 랜덤하게 섞어서 169개의 비정상 패킷이미지를 테스트하였다. TearDrop을 학습한 신경망으로 변종공격인 New TearDrop을 테스트 한 결과는 그림 8과 같다. Target값은 데이터의 기대값으로 0인 것은 정상 데이터이고, 1인 것은 비정상 데이터이다. Result 항목은 시간 지연신경망이 구분해낸 결과로 Normal은 정상이고, Abnormal은 공격을 나타낸다.

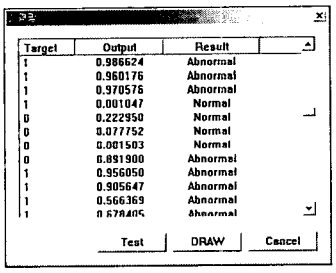

그림8. New TearDrop 결과

4. 결론 및 향후 과제

시간 지연 신경망 기반의 침입탐지 시스템은 패킷을 0에서 255사이 값으로 이루어진 그레이 이미지로 분석함으로써 공격패킷의 일부 이미지만으로도 구별해낼 수 있는 특징을 가지고 있다. 따라서, 개발하는 침입탐지 시스템은 알려진 다양한 침입유형뿐만 아니라, 새로운 변종에 대해서도 분류기의 유연한 반응을 통하여 효과적으로 탐지할 수 있다.

인터넷은 계속해서 확대될 것이며, 네트웍을 통한 침입은 그 유형만 달리할 뿐, 계속될 것이다. 따라서 신경망을 이용한 침입탐지 시스템은 기존의 규칙기반 침입탐지 시스템의 한계성을 극복하는 대안이 될 것이라 생각한다.

참고문헌

[1] http://www.certcc.or.kr

[2] R.Seker "A High Performance Network Intrusion Detection System," ACM 1-58113-148-8/99/0010, 1999

[3] Vern Paxson, "Bro: A System for Detecting Network Intruders in Real-time," Jan 14, 1998

[4] http://www.snort.org

[5] S. McCanne and V. Jacobson, "The BSD Packet Filter: A New Architechure for User-level Packet Capture," Lawrence Berkely Laboratory, Berkely, CA, 1992.

[6] Quinn, Bob외 "Windows Sockets Network Programming," Addison Wesley Publishing Company 1996.

[7] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, K. J. Lang, "Phoneme Recognition Using Time-Delay Neural Netwrks," IEEE Trans. on ASSP, vol. 37, no. 3, Mar. 1989.