신경망을 적용한 침입탐지시스템의 설계

이종혁*, 한영주**, 정태명**
*성균관대학교 컴퓨터공학과
이동형 응급의료 정보 시스템 개발센터
**성균관대학교 정보통신공학부
e-mail : {jhlee, yjhan}@imtl.skku.ac.kr, and tmchung@ece.skku.ac.kr

Design of Intrusion Detection System Using Neural Networks

Jong-Hyouk Lee*, Young-Ju Han**, Tai-Myung Chung**
* Cemi: Center for Emergency Medical Informatics,
Dept. of Computer Engineering, Sungkyunkwan University
**School of Information & Communication Engineering, Sungkyunkwan University

요 약

우리는 갈수록 지능화, 분산화, 자동화 되어 가고 있는 침입에 대해 효과적으로 대처하기 위해 신경망을 적용한 침입탐지 시스템을 설계 하였다. 본 논문은 신경망을 학습시키기 위해 학습 견본과 신경망 적용 인자를 정의 하였으며 학습 기법으론 MLP(Multi Layer Perceptron)을 이용 하였다. 새롭게 설계된 침입탐지 시스템의 탐지 모듈은 기존의 패턴 매치 방식의 모듈과 신경망 모듈이 적용되어 보다 정확한 침입 탐지가 가능하다.

1. 서론

컴퓨터와 네트워크 기술이 빠르게 발전함에 따라 대부분의 시스템들은 네트워크로 연결되어 있으며 기업 및 조직들은 그들의 중요 업무를 이들 시스템을 통해 처리하고 있다. 이와 더불어 시스템에 대한 공격도 폭발적으로 증가하여 2000 년 1943 건, 2001 년 5333 건, 2002 년 15192 건, 2003 년 26179 으로 전년 대비 200% 이상의 해킹 신고가 접수되고 있으며 공격의 유형 또한 지능화, 분산화, 자동화되어 가고 있다 [1].

침입탐지시스템(IDS: Intrusion Detection System)은 네트워크로 연결된 기업 및 조직의 정보 보호를 위해 사용되는 대표적인 보안 기술 중에 하나이다. 침입탐지시스템의 효율성은 침입에 대한 올바른 정의와 이를 기반으로 하는 침입 탐지의 정확성에 기반한다. 지금까지의 침입탐지시스템에서 사용되어 온 탐지 기술은 대표적으로 오용 탐지 기법과 비정상 행위 탐지 기법이 각각 혹은 혼합되어 적용되어 왔다. 그러나 false positive 또는 false negative 가 많이 발생될 뿐만 아니라 관리자를 통한 룰 적용에 따른 새로운 침입에 대한 신속한 대처가 어려운 상황이다 [2].

이에 따라 다양해 지는 침입에 대하여 신속하게 대처할 수 있는 새로운 침입탐지 기술의 필요성이 대두되고 있으며 이러한 기술 중에 하나가 바로 신경망이다.

본 논문에서는 자체 개발중인 침입탐지시스템인 SecureFortress 에 신경망을 적용하기 위한 제반 기술에 대한 고찰과 설계 및 탐지 메커니즘을 논한다. 2 장에서는 관련 연구로 침입탐지시스템의 개요에 대해 살펴본 후 현재 개발중인 침입탐지시스템의 특징 및 개선의 필요성을 분석하고 3 장에서는 신경망의 개요와 특성에 대해 알아 보며 4 장에서는 신경망을 적용한 침입 탐지시스템의 설계 및 탐지 메커니즘을 논하고 5 장에서는 결론 및 향후 연구과제에 대해서 기술 한다.

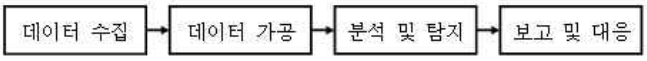
2. 관련연구

2.1 침입탐지시스템의 개요

침입탐지시스템은 시스템의 불법적인 사용이나 오

본 논문은 보건복지부 보건의료기술진흥사업회 지원에 의하여 이루어진 것임(과제번호: 02-PJ3-PG6-EV08-0001)

용, 남용 등에 의해 발생되는 침입을 탐지해 내는 시스템이다 [3]. 일반적으로 침입 탐지시스템은 [그림 1]과 같이 4 개의 모듈로 이루어져 있다 [2].

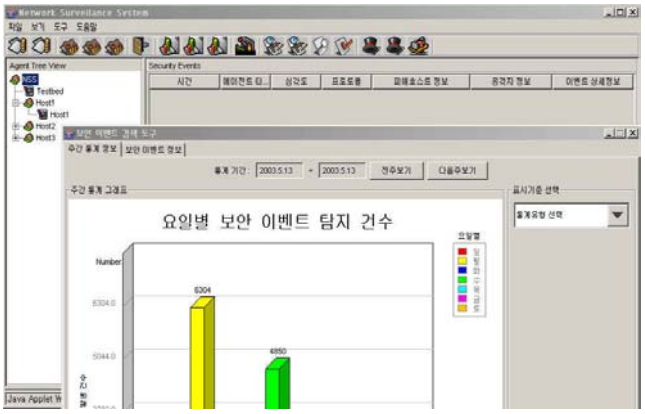


[그림 1] 침임탐지시스템의 절차

- 데이터 수집(Data Auditing): 호스트와 네트워크 기반에 따라 수집되는 데이터의 종류가 다른데, 호스트 기반의 경우 로그 및 프로세스 정보 등을 수집하며 네트워크 기반의 경우 관리 대상 네트워크 내 패킷 정보를 수집한다.
- 데이터 가공(Data Processing): 데이터 수집 모듈로부터 수집된 데이터를 가공 혹은 축약하여 분석 및 탐지 모듈로 보낸다.
- 분석 및 탐지(Analysis & Detection): 가공된 데이터를 미리 정해진 정책(policy) 및 규칙(rule)에 따라 침입여부를 결정하게 된다.
- 보고 및 대응(Reporting & Response): 침입으로 판단되었다면 관리자에게 보고하며 정해진 대응 한다.

2.2 SecureFortress

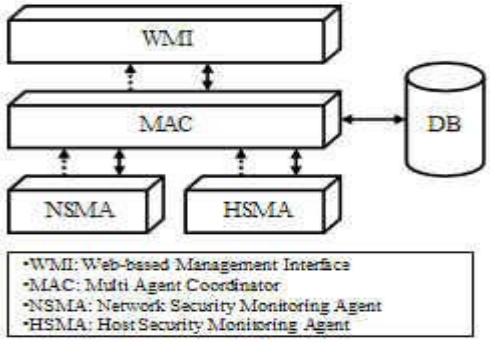
현재 개발중인 침입탐지시스템인 SecureFortress 의 관리 인터페이스는 [그림 2]와 같다.



[그림 2] SecureFortress 의 관리 인터페이스

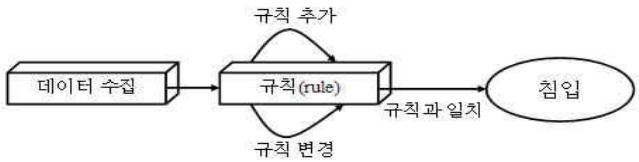
SecureFortress 는 시스템의 확장성을 고려하여 탐지 기능을 수행하는 에이전트와 관리 기능을 수행하는 매니저와 관리 인터페이스의 3-tire 구조로 이루어져 있으며 각각의 기능들이 모듈화 되어 있어 개발과 확장이 용이하다. 이러한 SecureFortress 는 호스트와 네트워크 기반의 침입에 대해서 오용탐지 방식을 적용하여 침입을 탐지 한다.

[그림 3]은 SecureFortress 의 구성 모듈을 나타낸다. 구성 모듈은 크게 4 개로 이루어져 있으며 에이전트를 관리하는 MAC 은 [그림 1]의 보고 및 대응 기능에 해당하며 보안 이벤트 수집 및 하위 수준의 보안 분석 기능을 수행하는 NSMA, HSMA 를 통합 관리하며 에이전트로부터 수신되는 보안 이벤트를 WMI 에게 통보하는 기능을 수행한다. WMI 는 [그림 3]의 MAC 과 함께 보고 및 대응 기능에 해당하며 보안 관리자가 보안 관리 기능을 수행할 수 있도록 지원하기 위한 웹 기반의 인터페이스이다.



[그림 3] SecureFortress 구성 모듈

NSMA 와 HSMA 는 SecureFortress 에서 실질적으로 침입을 탐지하는 핵심 기능로 [그림 1]의 데이터 수집, 데이터 가공, 분석 및 탐지 기능에 해당한다. 각각의 탐지 기법을 살펴보면, NSMA 는 보호대상 네트워크로부터 패킷 정보를 수집하고, 미리 설정되어 있는 오용 패턴을 수집된 패킷 정보와 비교하여 침입을 탐지하는 패턴 매칭 방식을 이용하여 침입을 탐지 한다. HSMA 또한 대상 호스트가 생성하는 로그 정보로부터 보안 관련 감사 정보를 수집하고, 이를 미리 설정된 패턴과 비교 분석하여 사용자의 오용을 탐지한다. 아래 [그림 4]는 일반적인 패턴 매칭 방식을 설명하고 있다 [3].



[그림 4] 패턴 매칭 방식

이와 같이 SecureFortress 는 확장성과 모듈화를 바탕으로 다양한 환경에서 적용 가능하도록 만들어져 있지만 탐지 기술이 안고 있는 한계로 인하여 계속적으로 발전하고 있는 침입에 대해서 능동적으로 대처할 수가 없다. 이에 침입탐지시스템의 핵심인 탐지 기법의 추가나 변경이 필요하다.

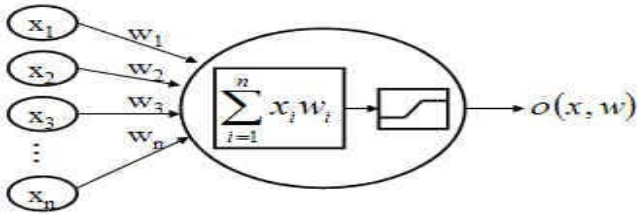
본 논문에서는 네트워크로부터 패킷 정보를 수집하고 분석하여 침입을 탐지해 내는 NSMA 의 패턴 매칭 방식을 신경망으로 재설계한다. 패턴 매칭 방식은 계속적인 규칙(rule)의 갱신이 필요하며 새로운 침입에 대해서 탐지가 불가능 하지만 신경망은 계속적인 규칙의 갱신이 필요 없으며 새로운 침입에 대해서도 탐지가 가능하다는 장점이 있다.

3. 신경망(Neural Networks)

신경망은 생물의 신경전달 과정을 단순화하고 수학적으로 모델링한 것을 말하며 상호 작용을 하는 각각의 뉴런(neuron)을 통과시키면서 뉴런끼리의 연결 강도를 조정하는 학습을 통해 문제를 분석하게 된다 [4].

$$sum = \sum_{i=0}^{n} x_i w_i$$

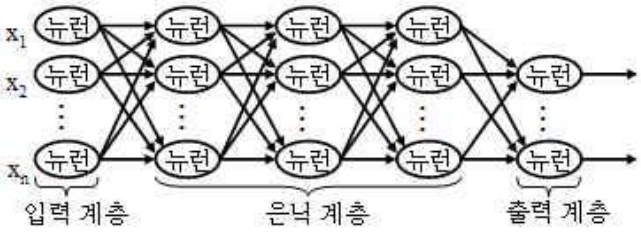
$$o(sum) = \frac{1}{1+e^{-sum}}$$



[그림 5] 뉴런의 구조

각각의 뉴런은 입력과 출력을 할 수 있으며 자체적으로 데이터를 분석하고 다른 뉴런들과의 상호 작용을 통해 데이터를 분석 할 수 있다. [그림 5]는 뉴런의 구조를 나타내고 있다. 입력 벡터와 각각의 가중치를 기반으로 한 입력으로 출력을 하게 되며 출력의 범위는 0.0 에서 1.0 사이가 된다 [5].

본 논문에서 사용될 신경망은 뉴런의 계층적 구조로 이루어져 있는 MLP(Multi Layer Perceptual)를 사용한다. [그림 6]의 MLP 계층은 입력 계층(input layer), 은닉 계층(hidden layer), 출력 계층(output layer)로 이루어져 있으며 다수의 은닉 계층은 입력 계층으로부터 전달 받은 데이터를 상호작용을 통해 연산을 하게 되고 출력 계층으로 전달한다. 실제 연산을 하게 되는 은닉 계층에 존재하는 뉴런들은 서로 연결되어 있으며 이 상호 연결 강도를 조절하고 연결에 대한 구성을 정의하는 것이 신경망의 핵심이라고 할 수 있다.



[그림 6] MLP

본 논문에서 신경망을 형성하고 있는 뉴런들을 학습 시키고 학습된 데이터를 바탕으로 또 다시 학습시키는 과정을 반복하여 최적의 신경망을 구성하기 위해 사용되는 알고리즘으로 역전파(Backpropagation) 알고리즘을 이용 한다 [6]. 역전파 알고리즘은 신경망에 학습 표본을 적용하고 출력과 표본을 비교하여 각 출력 뉴런에서의 에러를 계산 한다. 각 뉴런에 대해서 에러, 실제 출력, 스칼라 팩터를 계산하고 요구 수치보다 얼마나 높거나 낮은지를 계산하여 각 뉴런의 연결 가중치를 조절하는 방법이다. 실제 출력과 요구 수치과의 계속적인 비교를 위해서 뉴런의 연결 가중치를 조절하는 단계를 반복적으로 수행 한다.

이러한 싱경망을 침입탐지시스템에 적용하기 위해 필요한 요구사항은 학습 표본과 신경망 적용 인자이다.

4. 설계

본 장에서는 개발 중인 SecureFortress 의 NSMA 에 신경망을 적용한 시스템의 설계에 대하여 논하고 자 한다. 우리는 본 시스템에 신경망을 적용하기 위한 학습 표본으로 DARPA(Defence Advanced Research Project Agency)의 침입탐지 결과 데이터를 이용하도록 한다. 이 결과 데이터에는 다음과 같은 항목으로 이루어져 있다 [7].

[표 1] DARPA 침입탐지 결과 데이터 항목

항목	설명
Session ID	세션 구별 이름
Start Date	세션이 시작된 날짜
Start Time	세션이 시작된 시간
Duration	세션이 지속된 시간
Attack Score	공격의 강도
Attack Name	공격의 이름
Source Port	출발지 포트
Destination Port	목적지 포트
Source IP	출발지 주소
Destination IP	목적지 주소
Service	서비스명

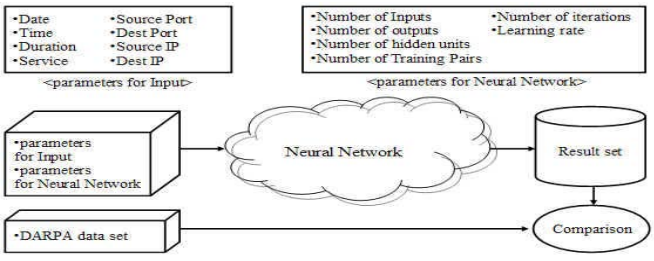
[표 1]의 결과 데이터에는 세션별로 나누어져 있으며 12 개의 침입이 분류되어 있다. 학습 데이터로 사용하기 위해 세션별로 침입으로 탐지된 세션 50 개와 평상시 세션 50 개를 입력으로 사용하도록 한다. 신경망 학습 인자로는 [표 2]와 같다.

[표 2] 신경망 학습 인자

인자	설명
Number of inputs	입력 비트의 수
Number of outputs	출력 비트의 수
Number of hidden units	은닉 계층의 뉴런의 수
Number of training pairs	학습 세션의 수
Number of iterations	학습 반복 횟수

[표 1]의 입력 데이터로 침입 세션과 평상시 세션을 입력한 후 주어진 입력값에 의해 나온 결과와 학습 표본 데이터를 비교하여 뉴런들과의 연결 강도를 조절하는데 시간 Δt 이후의 가중치 값은 아래와 같이 구한다 [8].

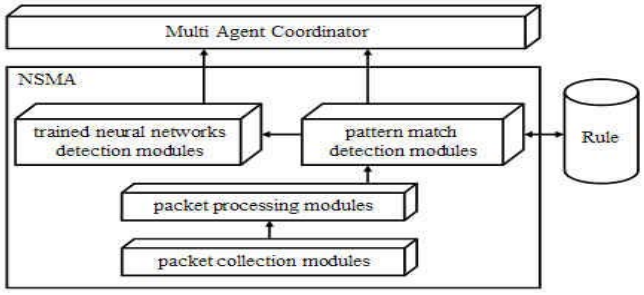
$$w_i(t+\Delta t) = w_i(t) + \eta \delta x_i$$



[그림 7] 학습 과정

[그림 7]은 학습 과정을 나타내고 있다. 입력 인자들과 신경망에 적용할 인자들을 이용해 결과를 만들어 내고 만들어진 결과와 학습 견본과 비교하게 된다.

이때 학습 과정은 역전파 알고리즘을 이용해 만족할 만 결과가 나올때까지 계속해서 학습과정을 거치게 된다.



[그림 8] 신경망이 적용된 NSMA 모듈

위의 [그림 8]은 신경망을 적용한 NSMA 모듈이다. 패킷을 수집하고 가공하는 모듈을 거쳐 패턴 매치 방식의 탐지 모듈을 이용해 침입을 탐지하며 패턴 매치 방식의 탐지 모듈이 정상 데이터라고 판단한 데이터들은 다시 신경망이 적용된 모듈에서 침입을 판단하게 되어 최종적으로 MAC 에게 침입 유무를 보고 하게 된다. 아래의 [그림 9]는 신경망이 적용된 NSMA 모듈의 의사 코드이다.

```
1: /* 초기 입력 패킷은 가공 */
2: packet = preprocess(raw_packet);
3: /* 규칙(R)에 매치 되는 패킷 */
4: if((the pattern of packet) ∈ R) {
5:    packet is dropped;
6:    report(packet->info);
7: }
8: /* 신경망에 의해 탐지된 패킷 */
9: if((neural_network(packet) == attack) {
10:     packet is dropped;
11:     report(packet->info);
12: }
13: /* 침입을 탐지하는 신경망 함수 */
14: int neural_network(packet) {
15:     /* 입력층에서 은닉층으로 */
16:     for i=0 to input_unit {
17:        for j=0 to hidden_unit {
18:           act[i] = init_act(packet); /* 입력값을 저장 */
19:           /* 연결 가중치와 활성값 곱의 합을 은닉층
                 으로 신호 전달 */
20:           value[j] += connectivity(hidden_unit) * act[i];
21:        }
22:     }
23:     /* 은닉층에서 출력층으로 */
24:     for i=0 to hidden_unit {
25:        /* 은닉층의 활성화 결정 */
26:     act[input_unit + i] = (value[i] > neu[i]);
27:     for j=0 to out_unit { /* 출력층으로 신호 전달 */
28:           value[hidden_unit + j] =
29:           connectivity(hidden_unit, out_unit) *
              act[input_unit + i];
30:     }
31: }
32: /* 출력층의 활성값 결정 */
33: for i=0 to out_unit {
34:     act[input_unit + hidden_unit + 1] =
          (value[hidden_unit + i] > neu[hidden_unit + i]);
35:   }
36:
37: return act;
38:}
```

[그림 9] 신경망이 적용된 NSMA 모듈 의사 코드

18 라인의 init_act 함수는 입력 패킷으로 활성화 값으로 만드는 함수이며 20, 29 라인의 connectivity 함수는 연결 가중치를 구하는 함수이다.

5. 결론 및 향후 과제

본 논문에서는 신경망을 적용한 탐지 모듈과 기존의 패턴 매치 탐지 모듈의 장점을 수용한 탐지 모듈을 설계 하였다.

신경망이 적용된 NSMA 는 관리자에 의한 추가적인 규칙의 갱신이 불필요하며 새로운 침입에 대해서도 탐지가 가능하다는 장점이 있다.

하지만 학습 표본으로 사용하게 되는 DARPA 의 침입탐지 결과 데이터은 침입탐지시스템을 적용하게 될 네트워크에 최적화된 학습 표본이 아니므로 실제 사용하게 될 네트워크에서 학습 표본 데이터를 추출하여 신경망에 적용해야 할 것이다.

이에 향후 연구 과제로 학습 데이터를 실제 사용하게 될 네트워크에서 효율적으로 추출하는 방법에 대해 실험 할 것이다.

참고문헌

[1] 한국정보보호진흥원, *2003 해킹 바이러스 동향*, http://www.kisa.or.kr, 2003.

[2] Theuns Verwoerd, Ray Hunt, “Intrusion detection techniques and approaches”, Computer Communications Volume 25, Issue 15 , Pages 1356-1365, September, 2002.

[3] Aurobindo Sundaram, *An Introduction to Intrusion Detection*, http://www.acm.org/crossroads/xrds2-4/intrus.html, 2001.

[4] Alexander Novokhodko, “A Survey of the Applications of Neural Networks to Intrusion Detection”, April, 2002.

[5] J.Ryan, M. J. Lin and R. Mikkulainen, “Intrusion Detection with Neural Networks”, Advances in Neural Information Processing System, Vol. 10, Cambridge, MA: MIT Press, 1998.

[6] Charles Anderson, "On the Use of Neural Networks to Guide Software Testing Activities", International Test Conference, October, 1995.

[7] Lincoln Laboratory, Massachusetts Institute of Technology, *DARPA Intrusion Detection Evaluation*, http://www.ll.mit.edu/IST/ideval/index.html

[8] 노영주, 조성배, “기계학습 기법에 의한 비정상행위 탐지기반 IDS 의 성능 평가”, 정보처리학회, November, 2002.