情報保護學會論文誌
第12券 第3號, 2002. 6

Adaptive Resonance Theory 2를 이용한 네트워크 기반의 침입 탐지 모델 연구*

김 진 원**, 노 태 우**, 문 종 섭**, 고 재 영***, 최 대 식***, 한 광 택***

Network based Intrusion Detection System using Adaptive Resonance Theory 2

JinWon Kim**, TaeWoo Noh**, JongSub Moon**,
JaeYoung Koh***, DaeSik Choi***, KwangTaek Han***

요 약

인터넷의 확장에 따라서 네트워크를 통한 침입이 증가되고 있다. 이에 따라 네트워크를 통한 침입에 대하여 즉각적으로 탐지하고 대처하는 기술이 필요하게 되었다. 본 논문은 인터넷의 특성을 악용하여 침입하는 공격들을 탐지하기 위하여 Adaptive Resonance Theory2(ART2) 이론을 이용한다. 정상적인 packet과 여러 가지 공격툴을 사용하여 생산한 인위적인 공격 패킷에 대하여 ART2를 학습한 후 실험한 결과와 기존의 방식들과 비교 분석하였다.

ABSTRACT

As internet expands, the possibility of attack through the network is increasing. So we need the technology which can detect the attack to the system or the network spontaneously. The purpose of this paper proposes the system to detect intrusion automatically using the Adaptive Resonance Theory2(ART2) which is one of artificial neural network. The parameters of the system was tunned by ART2 algorithm using a lot of normal packets and various attack packets which were intentionally generated by attack tools. The results were compared and analyzed with conventional methods.

Keyword : *IDS, ART2, neural network*

Ⅰ. 서 론

최근에는 네트워크로 연결된 시스템의 가용성, 기밀성, 무결성을 해치는 침입을 탐지하기 위한 많은 연구가 진행중이다.[1] 불법적인 침입에 대응하려면, 네트워크나 시스템의 각종 사항을 파악하여야 하고, 새로운 형태의 공격 패턴에 대해서도 탐지하여야 한다. 대표적인 방식으로는 침입탐지 시스템이 있다. 일반적으로 침입을 판정하기 위한 방법으로는 오용탐지(Misuse detection)와 비정상행위 탐지(Anomaly detection)가 있다. 오용탐지 방법은 공격패턴에 대하여 정해 놓고, 현재 감시 대상 패턴을 이 패턴과 비교함으로써, 정상패턴인지 공격패턴인지 판정하는 방법이다. 따라서 시스템 설치 초기에 공격패턴들에 대한 모델을 충분히 설정하지 않은 상태에서는, 감시 대상 패턴 대부분은 정상으로 판정된다(False

* 본 연구는 국가보안기술연구소의 지원으로 수행하였습니다.
** 고려대학교 정보보호대학원(mirujohn@korea.ac.kr, semirough@cist.korea.ac.kr, jsmoon@korea.ac.kr)
*** 국가보안기술연구소(jykoh@etri.re.kr, dschoi@etri.re.kr, kthan@etri.re.kr)

Negative). 이 시스템들에서는 초기에 가능한 많은 공격 패턴들을 설정하기 위한 비용과 노력이 소모되며, 최근의 경향과 같이 공격의 패턴 종류가 급속히 늘어나는 경우는, 설정비용과 노력이 무한히 증가하기 때문에 거의 효력을 발휘하지 못한다.

비정상행위 탐지 방법은, 정상적인 사용 패턴을 정해 놓고, 현재 감시 대상 패턴을 정상적인 사용 패턴들과 비교하여 공격인지 아닌지 판정하는 방법이다. 이 방법은 시스템 설치 초기에 정상적인 사용 패턴이 구축되어 있지 않으면 감시 패턴이 정상적일지라도 공격으로 판정하는 비율이 상당히 높다. 따라서 각 시스템마다, 각 사용자마다 초기 설비 비용과 많은 노력이 소모된다.

지금까지 상용화 되어진 대부분의 침입탐지 시스템에서는 정상적인 사용 패턴을 구축하든(Misuse detection), 또는 공격 패턴을 구축하든(Anomaly detection)간에, 이 패턴을 구축하는데 많은 노력과 비용이 사용된다.[8] 구축된 패턴 대상으로 한 판정 방법은 일반화된 방법을 사용하지 않고, Rule based system,[18,19] State transition diagram,[18,19] 또는 petri net[18,19] 같은 정형화된 비교 방법을 사용함으로써, 약간의 패턴 변화에도 적응하지 못하고, 따라서 보존해야 하는 패턴의 양이 급속히 증가하고, 판정에 많은 시간이 걸린다. 이러한 정형화된 방법을 사용하는 시스템들은, 새로운 패턴이 나타날 때마다 새로운 규칙이나 패턴을 등록시켜야 하는 단점이 있다.

이에 반하여, 신경망 모델은 일반화 성능이 탁월하다. 신경망 모델은 샘플 패턴으로부터 일반적인 규칙을 생성하는 능력이 아주 우수하다.[5,7] 따라서 생성해야 할 규칙이나 패턴이 아주 많거나, 어떤 규칙을 기준으로 약간씩 변경된 패턴이 많거나, 시스템이 다양하게 변하는 패턴에 그때그때 잘 적응해야 할 경우에는 신경망 모델이 대표적으로 사용된다. 적절한 샘플 패턴으로부터 잘 학습된 신경망은 학습되어진 패턴이거나, 또는 학습되어지지 않은 패턴에 대해서도 고정된 규칙을 사용하는 시스템보다도 비교적 판정이 양호하고, 규칙을 설정하는 비용이나 노력이 거의 들지 않으며, 설치시 바로 적용이 된다. 또한 학습용 데이터 선정시, 정상적인 패턴과 비정상적인 패턴을 모두 학습에 사용함으로써, 오용탐지와 비정상행위 탐지를 따로 구분하지 않고 동시에 사용하는 역할을 한다. 본 논문에서는 신경망 모델의 한 분야인 ART2를 사용하는데, 이는 특히 적응적인 시스템으로써, 새로운 공격패턴이 나타나면, 재학습이 필요 없이•즉시 적응 가능하다. 본 논문은 네트워크 상의 패킷들에 대하여 알려지지 않은 공격과 비정상 행위들을 탐지하는 실시간 침입탐지 시스템을 제안 하고자 한다.[9] 본 논문의 구성은 다음과 같다.

2장에서는 기존의 침입탐지 시스템의 관련연구에 대하여 알아보고, 3장에서는 본 연구에 사용된 신경망 알고리즘 인 Adaptive Resonance Theory2 (ART 2)에 대하여 서술한다. 4장에서는 신경망을 이용하여 개발한 침입 탐지 시스템의 구성, 특징, 장점을 살펴보고, 5장에서는 개발한 시스템의 실험환경 및 결과에 대하여 서술한다. 6장에서는 결론 및 앞으로의 연구방향에 대해 서술한다.

II. 관련 연구

침입탐지시스템은 탐지 대상 데이터에 따라 크게 시스템기반 침입탐지 시스템과 네트워크기반 침입탐지시스템으로 나뉘어 진다. 시스템기반 침입탐지 시스템은 많은 연구가 진행되고 있으나, 네트워크기반 침입탐지시스템은 미리 침입에 대한 정의를 내리고 탐지여부를 결정하는 오용침입탐지 시스템 연구에서 크게 벗어나지 못하고 있다.[7] 비정상행위에 대한 탐지 시스템에 대해서는 아직 초기 단계에 있는 상태이다.

기존의 연구를 살펴보면, 미국의 SRI(Stanford Research Institute) International/CSL(Computer Science Laboratory)[2]의 경우, 1980년대부터 침입탐지 시스템에 대한 연구를 시작하여 지금까지 계속적으로 수행하고 있다. SRI의 침입탐지에 관한 연구의 기초는, 감사데이터 축약(reduction)과 분석(analysis)에 통계학적 기술을 사용하는 것이다. 1세대 통계학적 요소(Statistics Component)는 1980년대 초반에 IBM 메인프레임 시스템의 SMF (System Management Facility) 레코드들을 분석하는데 사용되었다. 이후에, 이 연구는 알려진 악의적인 행위를 탐지하기 위하여 규칙기반 전문가 시스템(Rule-Based Expert System)의 사용을 시도하였다. 이러한 기초 연구는 단일 호스트의 보안 위배 사항을 실시간으로 탐지하는 능력을 가진 IDES (Intrusion Detection Expert System)의 원형을 개발을 이끌어 냈다.[2,3]

IDES는 실시간 침입탐지 전문가 시스템으로 사

용자의 행동을 관찰하고 각 사용자와 사용자 그룹에 있어 정상적인 행동양식의 통계값을 사용자별 속성 파일(Profile)로 관리하고 단기간의 사용자 행동과 장기간 행동 패턴에 대한 도수 분포를 비교하여 비정상행위를 판정하고 있다. 오늘날 일반적 침입탐지 모델로 많이 인용되고 있는 모델이 1987년 Dorothy Denning에 의해 제시되었던 침입탐지 모델이며, 이것이 IDES의 기본 설계 모델이다.

IDES를 기반으로 하여 개발된 분석 방법론이 원숙기에 접어들면서, SRI 초기 IDES 프로토타입을 발전시키고 최적화 시키려는 노력을 기울여, NIDES (Next generation IDES)라 불리는 고 수준의 침입탐지 시스템 개발을 시작하였다.

SRI는 DARPA 프로젝트의 일환으로 NIDES의 후속 시스템인 EMERALD(Event Monitoring Enabling Responses to Anomalous Live Disturbance)를 개발하였다.[4] EMERALD는 대규모 분산 네트워크 환경에서 침입탐지 및 대응 시스템이며, 확장이 용이한 시스템이다. 시스템은 네트워크 환경에 분산된 모니터들로 구성되며, 이들은 각자 독립적인 감시, 대응 등의 역할을 수행하고 조절이 가능하다.

Ⅲ. Adaptive Resonance Theory2(ART2)

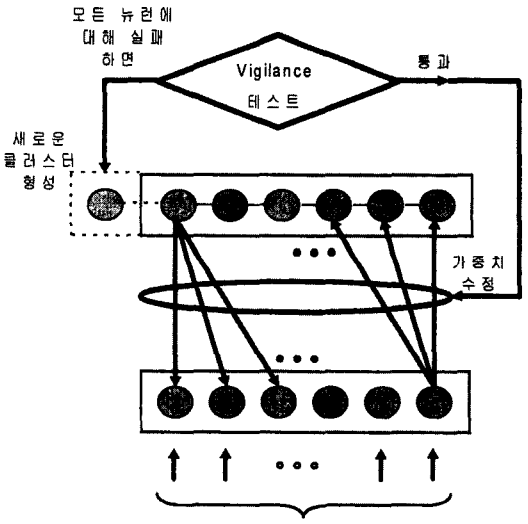
3.1 ART의 개요

Stephen Grossberg와 G.A.Capenter에 의해 제안된 Adaptive Resonance Theory(ART)[6,10] 신경망 모델은, 임의의 입력패턴에 대해 이미 학습된 패턴을 간직한 채 새로운 패턴을 학습할 수 있는 안정성과 적응성을 가지는 신경망 모델이다. 이 모델은 또한 일반적인 신경망의 장점인 일반화 특성도 그대로 가진다. 본 연구에서 ART2를 도입한 이유는 기존 신경망의 문제점 중 하나인 Stability-Plasticity 문제를 해결할 수 있는 특징 뿐 아니라, Vigilance Parameter를 통한 분류강도의 조정이 가능하기 때문이다.[6] 즉, 적응성이 우수하다.

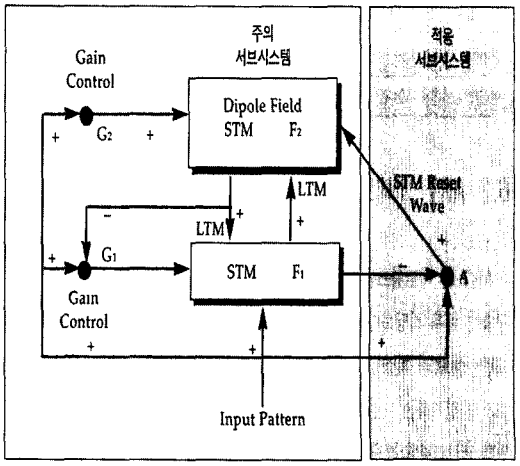
3.1.1 ART 구조

ART 모델의 구조는 입력 레이어(F0), 비교 레이어(F1), 클러스터링 레이어(F2)로 구성된 주의 서브시스템(Attentional Subsystem)과 Reset Unit으로 구성된 적응 서브시스템(orienting Subsystem)으로 되어 있다.

F1과 F2층은 각 노드에 활성화된 패턴을 저장하기 때문에 Short Term Memory(STM)이라 하고, F1과 F2사이의 상향, 하향 연결 가중치를 Long Term Memory(LTM)으로 나타낸다. F0는 입력 패턴을 받아들여 정규화 작업을 수행하며, F2는 입력 패턴들에 의해 생성된 클러스터를 저장한다. F1은 F0에서 입력받은 패턴과 F2가 기억하고 있는 클러스터와의 비교를 통해 유사한 클러스터를 선택하며 Reset Unit은 비교 결과에 따라 F2의 클러스터를 비활성 하는 기능을 수행한다.[6]



〔그림 1〕 ART 알고리즘

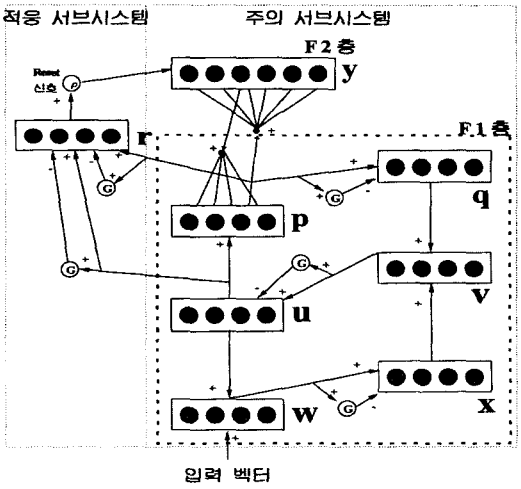


STM : Short Term Memory, LTM : Long Term Memory

〔그림 2〕 ART 시스템 구조

3.2 ART2

ART2는 ART와 유사한 구조를 가지나 입력되는 패턴의 형태가 ART와 다르다.[16] ART2의 입력패턴은 이진 입력 패턴뿐만 아니라 아날로그(또는 gray-level) 입력 패턴에 대해서도 학습이 가능한 모델이다. 아날로그 입력 패턴을 처리하기 위해서 ART2에서는 [그림 3]에서 보듯이 F1층은 정규화 작업을 수행하기 위한 w,x,u,v,p,q 노드로 구성된다.[6,7,17] 각 노드 간에는 해당 뉴런을 활성화 시키기 위한 gain control이 연결되어 있다. F2층은 ART와 동일하다.



[그림 3] ART2 시스템 구조

3.2.1 ART2 신경망에서 사용된 변수

- a, b : F1층에서의 가중치
  - a : node u와 node w 간의 가중치
  - b : node q와 node v 간의 가중치
- c : reset test에서 사용되는 가중치
- d : F2층의 winning node의 activation 값

$$\frac{c \cdot d}{1-d} < 1 \qquad (1)$$

식 (1)을 만족하도록 선택한다.

Signal Function은 식 (2)와 같은 Piecewise Linear 함수를 사용한다.

$$f(x) = \begin{cases} x & \text{if } x \geq \theta \\ 0 & otherwise \end{cases} \qquad (2)$$

이는 F1의 q, x 노드로부터 v 노드로의 출력값을 계산하는데 사용한다.

- θ(Threshold): 식 (2)에서 사용되는 임계치이다. 식 (3)을 만족해야 한다.

$$0 < \theta \leq \frac{1}{\sqrt{n}} \qquad (3)$$

(n : dimension of input pattern)

- ρ(vigilance parameter) : 얼마만큼의 클러스터를 형성할 것인가를 결정한다.[13,16] ($o < \rho < 1$)

3.2.2 ART2 신경망 알고리즘 단계

1) Step1 - 초기화 단계

ART2에 수행에 필요한 파라미터들을 a b θ c d ρ값과 함께 식 (4)에 따라 초기화한다.

$$T_{ij} = 0,\ B_{ji} \leq \frac{1}{(1-d)\sqrt{n}} \qquad (4)$$

(F1 : i=1…n, F2 : j=1…m)

T_{ji} : Top-down weight(F2 to F1)

B_{ji} : Bottom-up weight(F1 to F2)

a, b, c는 각각 F1의 입력 벡터 정규화 작업의 수행을 위해 필요한 노드간의 입출력 값을 조정하는 계수이다.

2) Step2

- F1 Layer의 계산, F2 Layer의 입력 계산

F1의 각 노드는 새로운 입력 벡터 s에 대하여 식 (5)에 의해 초기화 된다. 따라서 새로운 입력 벡터가 네트워크에 투입되면 이전 벡터에 대한 기억은 F1으로부터 삭제된다. 이러한 특성 때문에 F1은 Short Term Memory라 불린다.[6,16,17]

$$\begin{aligned} &u=0, w=s, p=0, x=\|s\|, q=0, v=f(x) \\ &u=\|v\|, w=s+au, p=u, x=\|w\| \\ &q=\|p\|, v=f(x)+bf(q) \end{aligned} \qquad (5)$$

식 (5)에 의해 초기화가 끝나면 F2의 각 j 노드의 실제 입력을 식 (6)에 의해 계산한다.

$$y_j = \sum_{i=1}^{n} B_{ij} p_i \quad \text{for } j = 1 \ldots m \tag{6}$$

3) Step3
- Winning node select & Reset check

F2의 노드 중 최대 출력값을 갖는 J를 선택한다. 그리고 Reset 여부를 확인하기 위하여 식 (7)에 의해 r_i를 계산한다.

$$u_i = \| v \|$$
$$p_i = u_i + dT_{ji},\ r_i = \frac{u_i + cp_i}{\| u \| + c \| p \|},\ i = 1..n \tag{7}$$

$$\text{if } \| r \| < \rho \text{ then } y_J = -1 (reset = true)$$
$$\text{if } \| r \| < \rho \text{ then}$$
$$w = s + au, x = \| w \|, q = \| p \|$$
$$v = f(x) + bf(q) (reset = false) \tag{8}$$

식 (8)에 의해 reset이 true가 되면 Step3을 반복한다. 그렇지 않을 경우 다음 단계로 넘어간다.

4) Step4 - Weights의 갱신
식 (9)에 의하여 노드 J의 가중치 벡터를 갱신한다.

$$T_{ji} = adu_i + \{1 + ad(d-1)\} T_{ji}$$
$$B_{ji} = adu_i + \{1 + ad(d-1)\} B_{ji} \tag{9}$$

그리고 식 (10)에 의하여 F1 노드들의 값을 계산해 낸다.

$$u = \| v \|,\ v = s + au,\ p = u + dT_J$$
$$x = \| w \|,\ q = \| p \|,\ v = f(x) + bf(q) \tag{10}$$

이러한 학습과정을 통해 노드 J는 입력 벡터를 기억하게 된다. 이러한 기억은 새로운 입력 벡터가 투입되어도 F2에 의해 지속적으로 유지된다. 이러한 특성 때문에 F2를 Long Term Memory라 부른다.

5) Step5 - 반복
새로운 입력 벡터에 대하여 Step2 - Step4의 과정을 반복한다.

3.2.3 ART2 신경망 특징

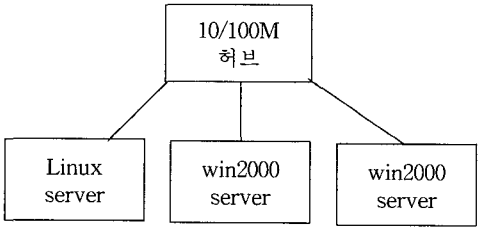
ART2 신경망 모델은 다음과 같은 장점을 갖는다.

- ART2는 비교사 학습(unsupervised learning)에 의해 입력 패턴을 클러스터링 하기 때문에 사전에 교사데이터 없이 새로운 입력 패턴을 학습할 수 있다.
- ART2는 기존 신경망 모델들의 단점인 Stability (안전성) - Plasticity(유연성) 문제를 해결하였다. 신경망에서 Stability란 이전에 학습한 패턴들에 대한 기억을 안정적으로 유지하는 능력을 말하며, Plasticity란 이전에 학습한 적이 없는 새로운 패턴을 처리할 수 있는 능력을 말한다. ART2는 입력패턴과 학습된 클러스터에 영향을 미치지 않으면서 학습을 수행할 수 있는 Reset 메커니즘을 사용하여 이 문제점을 해결하였다. 이러한 능력에 의해 Adaptive learning과 Incremental learning을 수행한다.
- ART2는 경계변수(vigilance parameter, ρ) 값에 따라 클러스터링의 분류 결과를 조정할 수 있다. 즉, 경계변수의 값을 크게 주면, 좀더 세분화되고 구체적인 클러스터들을 얻을 수 있다. 학습이 완료된 클러스터에 대한 가중치 값들은 해당 클러스터에 속해 있는 패턴들에 대한 대표 벡터로 해석될 수 있다.

Ⅳ. ART2 신경망 침입탐지 시스템

4.1 실험 환경

본 논문은 자체 구축한 네트워크 환경을 이용하여, 침입을 위한 패킷들을 수집하였으며, [그림 4]와 같이 10/100M허브에 연결된 세 가지 시스템으로 실험을 진행하였다.



[그림 4] 실험을 위한 네트워크 상황도

4.2 전처리(preprocessing) 단계

전처리 과정이란 그림 4와 같은 실험실 상황 하에 공격 패킷을 생성, 원시데이터(raw-data)를 추출

하여 신경망 입력으로 적합하게 변경하는 과정으로서 IP, TCP, UDP의 세 가지 프로토콜의 헤더정보를 사용하였다.

사용된 패킷정보는 다음과 같다.
- IP 헤더 정보(5차원)
  · Total Length
  · Identification Number
  · Flag(3 bit) - 각 bit별로 1차원씩 처리
- TCP 헤더 정보(12차원)
  · Source Port Number
  · Destination Port Number
  · Sequence Number
  · Acknowledgment Number
  · Header Length
  · Reserved(6bit) - 6차원
  · URG, ACK, PSH, RST, SYN, FIN
- 각각 1차원
- UDP 헤더 정보(3차원)
  · Source Port Number
  · Destination Port Number
  · UDP Length

위와 같은 패킷정보들을 각각 ASCII Code화 하여 신경망의 입력으로 사용하기 위해서 IP+TCP의 경우 17차원, IP+UDP의 경우 8차원인 두 가지 경우로 따로 분리하여 입력 벡터를 형성하였다. 이는 다음 단계인 신경망에의 적용에서 각기 다른 입력 벡터 값으로 쓰이게 된다.

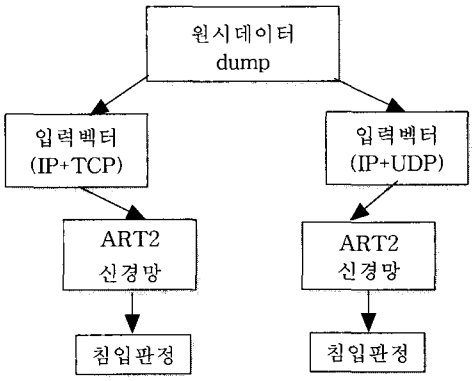
IP	TCP또는 UDP
5차원	12차원 또는 3차원

〔그림 5〕 전처리 후 신경망에서 사용 할 입력벡터 형태

4.3 실험

〔그림 4〕와 같은 환경에서 수집한 각 패킷들은 전처리과정을 통하여 신경망의 입력으로 ART2에 각각 IP+TCP형태의 벡터, IP+UDP형태의 벡터를 생성하게 되고 다음과 같은 〔그림 6〕과 같이 각각의 신경망에 입력후 침입/정상 판별을 하게 된다.

ART2의 학습과 시험을 위하여 〔표 1〕과 같은 9개의 공격 도구 중 선별하여 사용하였다. 공격 유형은 서비스 거부 공격에 중점을 두었다.



〔그림 6〕 신경망을 이용한 침임탐지 시스템

〔표 1〕 공격 유형별 사용 툴

공격 도구 번호	사용 공격 도구 명	공격유형
1	Stachelantigl	DoS,SYN,UDP Flooder
2	Syndrop	DoS, SYN Flooder
3	Synk,Synk4	DoS, SYN Flooder
4	Synhose	DoS, SYN Flooder
5	Synful	DoS, SYN Flooder
6	UDPdata	DoS, UDP Flooder
7	UDPFlood	DoS, UDP Flooder
8	Arnup100	DoS, UDP Flooder
9	rc8	DoS, UDP Flooder

4.3.1 IP+TCP

학습 데이터로서 〔표 2〕와 같이 정상 데이터 6000개와 2개의 공격 도구를 사용하여 만들어낸 비정상 패킷 200개, 총 6200개의 패킷으로 ART2신경망을 학습시켰다.

시험 데이터로서 〔표 3〕과 같이 학습 데이터로 쓰여지지 않은 3종류의 공격 툴을 사용하여 비정상 패킷 300개, 정상패킷 200개 총 500개로 시험 하였다.

〔표 2〕 IP+TCP 학습 데이터

데이터		개 수
정상		6000개
비정상	공격 도구1	100개
	공격 도구2	100개
총		6200개

※ 공격도구1 : Stachelantigl
공격도구2 : SYNdrop

〔표 3〕 IP+TCP 시험 데이터

데이터		개 수
정상		200개
비정상	공격 도구3	100개
	공격 도구4	100개
	공격 도구5	100개
총		500개

※ 공격도구3 : synk, synk4
공격도구4 : synhose
공격도구5 : synful

4.3.2 IP+UDP

학습 데이터로서 〔표 4〕와 같이 정상 데이터 6000개와 2개의 공격 도구를 사용하여 만들어낸 비정상 패킷 200개, 총 6200개의 패킷으로 ART2 신경망을 학습시켰다.

시험 데이터로서 〔표 5〕와 같이 학습 데이터로 쓰여지지 않은 2종류의 공격 툴을 사용하여 비정상 패킷 200개, 정상패킷 200개 총 400개로 시험 하였다.

〔표 4〕 IP+UDP 학습 데이터

데이터		개 수
정상		6000개
비정상	공격 도구6	100개
	공격 도구7	100개
총		6200개

※ 공격도구6 : UDPdata
공격도구7 : UDPFlood

〔표 5〕 IP+UDP 시험 데이터

데이터		개 수
정상		200개
비정상	공격 도구8	100개
	공격 도구9	100개
총		400개

※ 공격도구8 : Arnup100
공격도구9 : rc8

V. 실험 결과 및 분석

침입탐지 시스템의 성능을 객관적으로 비교하는 것은 쉽지 않다. 객관적이고 공통적으로 실험에 사용될 수 있는 데이터도 없거니와, 하루가 다르게 변하는 공격 패턴을 모두 다 수용하여, 각 시스템을 비교 평가 할 수도 없기 때문이다.[14] 따라서 지금까지 연구되어온 시스템들은 각자의 실험 환경에서 인위적으로 만들어서 각자의 성능평가를 위하여 사용하였다. 그 중에서 신경망을 이용한 첫 번째 논문[20]에서는 실험환경에 필요한 parameter를 destination IP, source IP, source port, source Flag를 확률적 방식으로 사용하였다. DoS에 대한 공격에 대한 실험결과는 20,000개의 packet중 5,000개를 학습모드로 15,000개를 테스트 모드로 사용하였다.

그 결과 정상적인 패킷은 14,165개이고, Syn-Flood 공격에 대한 패킷은 835개이고 오류 없이 탐지하였다.

두 번째 논문[9,15]의 실험에서는 〔표 6〕에서처럼 4가지의 공격유형별로 나누었다. 분류 실험 결과는 〔표 7〕과 같다. 클래스별 Best 탐지율 중 DoS에 대한 탐지율이 93.18%임을 알 수 있다.

〔표 6〕 Wenke Lee 클래스 분류

클래스번호	클래스
0	Normal
1	DoS
2	R2L
3	U2R
4	Probing

〔표 7〕 Wenke Lee의 실험 결과

클래스번호	탐지율
0	96.59
1	93.18
2	73.86
3	87.50
4	100

실험 데이터 수, parameter, 실험 방법 면에서 차이가 있지만, 두 논문 모두 DoS에 대한 실험결과 제안된 연구도 기존의 방법과 흡사하게 실험결과가 나타난 점을 알 수 있고, 본 연구에서는 ART2의 특성중 하나인 적응적으로 새로운 공격에 대해 반응하는 것을 다음 실험결과로부터 알 수 있다.

본 논문에서의 Error율과 탐지율은 다음과 같이 정의하였다.

$$\text{Error율}(\%) = \frac{B}{A} \times 100$$

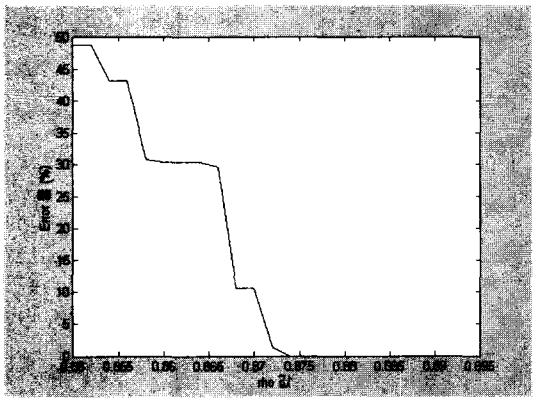
$$\text{탐지율}(\%) = \frac{C}{A} \times 100$$

A : 전체 test data 수
B : 정상(혹은 비정상)으로 정확히 판별하지 못 한 data 수
C : 정상(혹은 비정상)으로 정확히 분류된 data수

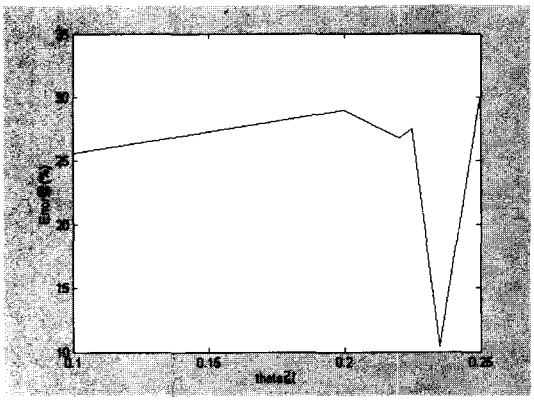
5.1 IP+TCP

학습데이터를 구성하여 경계변수(ρ)나 임계 변수(θ) 값을 바꿔가며 클러스터의 개수를 조정하였다. 즉, 경계변수와 임계변수에 의해 클러스터의 분류가 영향을 받게 되므로, 가장 최적의 상태를 찾기 위해 다음과 같이 실험을 하였다.

식 (3)을 만족하게 θ=0.235로 고정시키고, ρ값을 0.850부터 0.02씩 증가시켜 실험한 결과는 그림 7과 같다. 또한 ρ=0.875로 고정하고 θ값을 달리 하면서 실험한 결과는 [그림 8]과 같다. [그림 7, 8]에서 볼 수 있듯이 ρ=0.875, θ=0.235로 설정하였을 때, 최상의 결과가 형성되었다. 이에 따라 경계변수(ρ)와 임계 변수(θ) 값을 각각 0.875와 0.235로 고정한 실험결과에 대해서만 분석한다.



[그림 7] ρ(rho)값에 따른 Error율(%)(θ=0.235)



[그림 8] θ(theta)값에 따른 Error율(%)(ρ=0.875)

그 결과 [표 7]과 같이 총 6000개의 정상데이터는 42개의 클러스터를 형성하였고, 200개의 공격 데이터는 3개의 클러스터가 만들어졌다.

이에 대한 시험 데이터의 적용 결과, 학습에서 쓰여지지 않은 공격에 대해서는 각각 42~46과 같은 새로운 클러스터 형성을 보여줬으며, 학습에서 사용되지 않은 정상 데이터에 대해서는 기존의 0~41, 새로운 47~49의 클러스터 형성을 보였다. 위와 같은 시험의 결과는 [표 8]과 같다.

정상패킷을 공격으로 오인하는 경우와 비정상을 정상으로 판단하는 경우가 전혀 없었지만, [표 8]을 [표 7]과 비교해보면 학습을 통하여 형성되지 않은, 새로운 것이라고 판단되는 클러스터 번호(정상의 경우 47-49, 공격의 경우 45-46)가 만들어 진 것을 볼 수 있다.

이러한 새로운 클러스터 번호를 가지는 패킷들은 관리자에 의해 직접적인 정상/비정상 클러스터 판단이 필요하다. 이것은 새로운 입력값에 대하여 새로운 패턴을 적응적으로 만들어 내는 ART2의 특성을 잘 나타내고 있다.

[표 7] IP+TCP 학습데이터 클러스터 결과

데이터	클러스터 번호
정상	0~41(총 42개)
비정상	42~44(총 3개)
총	45개

[표 8] IP+TCP 시험 데이터 결과

데이터	클러스터 번호	탐지율
정상	0~41	92%
	47~49	8%
비정상	42~44	98%
	45~46	2%

5.2 IP+UDP

IP+TCP와 같이 다양하게 ρ와 θ를 변경하면서 실험하였을 때, ρ=0.89, θ=0.28로 설정하였을 경우가 최적의 결과를 보여주었으며, 학습을 통하여 형성된 클러스터 번호는 [표 9]와 같다. 6000개의 정상 학습데이터들은 0~3까지 총 4개의 클러스터, 200개의 공격 데이터는 4~5까지 총 2개의 클러스터를 형성하였다.

〔표 9〕 IP+UDP 학습데이터 클러스터 결과

데이터	클러스터 번호
정상	0~3(총 4개)
비정상	4~5(총 2개)
총	6개

〔표 10〕 IP+UDP 시험 데이터 결과

데이터	클러스터 번호	탐지율
정상	0~3	92.5%
	6	7.5%
비정상	4~5	100%

이에 대한 시험 데이터의 적용 결과, 〔표 10〕과 같이 정상에 대해서 새로운 클러스터 6의 생성되는 걸 볼 수 있었다. 이 또한 새로운 입력에 대한 새로운 패턴값을 나타낸다.

Ⅵ. 결론 및 향후 연구 과제

기존의 시스템들은 Rule을 생성하기 위해서 원시 데이터로부터 축약과정과 Rule 생성시 비용과 노력이 많이 든다. 또한 약간이라도 변경된 패턴에 대하여는 탐지가 불가능하다. 다양하게 공격 패턴이 커져 가는 현실에는 효능이 급속히 떨어지고, 시스템 구축 초기시에, Rule을 구축하는 비용과 노력이 무척 많이 든다. 본 연구에서는 이러한 단점을 극복하기 위해, 비정상행위에 대한 탐지를 자율 신경망중 하나인 ART2를 사용하였다. 또한 비정상행위에 대한 탐지에 있어서 서비스 거부 공격(Denial of Service)에 초점을 맞추었다.

제안한 침입탐지 시스템의 실험 결과를 정리하면 다음과 같다.

첫째, 클러스터링의 최적의 상태가 경계변수(ρ)와 임계변수(θ) 값의 변화에 따라 〔그림 7, 8〕과 같은 결과가 도출되었으며, 실험환경에서 100%에 가까운 공격패킷과 정상패킷을 구분해 내었다.

둘째, 제안된 시스템은 클러스터링 기법의 자율학습 방식이기 때문에, 초기 학습을 제외한 추가적인 학습을 시키기 위한 학습 시간이 필요 없는 장점을 가진다.

셋째, 제안된 시스템은 ART2의 클러스터링 분류를 이용하여, 〔표 8, 10〕을 보면 알 수 있듯이, 새로운 입력 패턴에 대해 스스로 새로운 패턴을 만들어 내는 특징인 적응적인 학습과 증가적인 특성을 잘 보여주고 있다.

이러한 결과로 미루어 볼 때 ART2를 이용한 본 연구는 새로운 공격에 대해서 높은 탐지와 적응적, 능동적으로 탐지하는 것을 볼 수 있다.

향후 연구 과제로는 Signature 선정시 각 패킷의 헤더부분의 관계를 이용하였으나, 패킷 사이의 시간 관계에 대한 분석과 ART2의 신경망과 다른 신경망을 결합한 Hybrid 형태에 대한 연구와 SVM(Support Vector Machine)을 이용한 Concept ART의 연구를 수행할 계획이다.

참 고 문 헌

〔1〕 신대철, 이보경, 유동영, 김홍근 "네트워크 비정상행위 탐지를 위한 클러스터링 모델", *The 13th Workshop on Information Security and Cryptography*(WISC2001), pp. 187~201, Sept 2001.

〔2〕 D.Anderson, T.Frivoid and A.Valdes, "Next generation intrusion detection expert system(NIDES)", *Technical Report SRI-CLS-95-07*, May, 1995.

〔3〕 Harold S.Javitz and Alfonso Valdes, "The NIDES Statistical Component Description and Justification", Annual Report, *SRI Interational*, 333 Ravenwood Avenue, Menlo Park, CA 94025, March 1994.

〔4〕 Porras, A. and Neumann, P. G. EMERALD : Event Monitoring Enabling Responce to Anomalous Live Disturbances. *In Proceedings of the National Information Systems Security Conference*, October 1997.

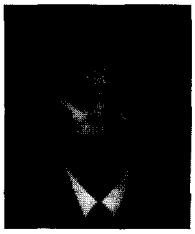
〔5〕 James A. Freeman and David M. Skapura, Neural Networks, Addison Wesley Publishing Company, pp. 291~339, 1991.

〔6〕 G. A. Capenter, and S. Grossberg "ART 2 : Self-Organizing of stable category recognition codes for analog input patterns", *Applied Optics, Vol.26*, pp. 4919~4930, December 1987.

〔7〕 Jake Ryan, Meng-Jang Lin, Risto Miikkulainen, "Intrusion Detection with Neural Networks", in Advanced in Neural Information Processing System 10, Cambridge MA : MIT process, 1998.

[8] Wenke Lee, "Data Mining Framework for Const-urctiong System", 1999.
[9] Wenke Lee, Salvatore J. Stolfo, Philip K Chan, Eleazar Eskin, Matthw Miller, Shlomo Hershkop, Junxin Zhang, "Real Time Data Mining-based Intrusion Detection", *IEEE*, 2001.
[10] Gail A. Carpenter and Stephen Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Networks", *IEEE Computers*, pp. 77~88, March 1998.
[11] P. G. Neumann and P. A. Porras, "Experience with emerald to date", *1st USENIX Workshop on IDS*, Santa Clara, pp. 11~12, April 1999.
[12] James Cannady, "Artificial Neural Networks for Misues Detection", 1988.
[13] Laurence Fausett, "Fundamentals of Neural Network Architecture, Algorithm, and Applications", *Prentice Hall*, pp. 218~288, 1994.
[14] 유신근, 이남훈, 심영철, "침입탐지시스템 평가 방법론", *한국정보처리학회 논문집*, 2000.
[15] Wenke, Lee, Salvatore J. Stolfo, Kui W Mok, "A Data Minining Framework for Building Intrusion Detection Models", *IEEE* Symposium on Security and Privacy, 1999.
[16] Filippidis, A., L. C. Jain, and P. Lozo. "Degree of Familiarity ART2 in Knowleged-Based Landmine Detection," *IEEE Transactions on Neural Networks*, January 1999, Vol. 10, No. 1, pp. 186~193.
[17] http://www.syssim.ecs.soton.ac.uk/vhdlams/examples/art2/art2.htm
[18] 멀티 호스트 기반 침입탐지 시스템 개발, *한국정보보호진흥원*, 1998
[19] 실시간 네트워크 침입탐지 시스템, *한국정보보호진흥원*, 1998
[20] 이장헌, 김성옥 "신경회로망을 이용한 비정상적인 패킷탐지", *정보보호학회논문집*, 2001

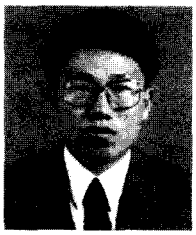
〈著者紹介〉



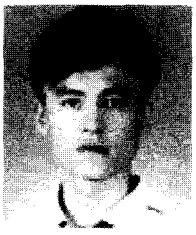
김 진 원 (Jin-Won Kim)
2001년 2월 : 고려대학교 전자계산학과 졸업
2001년 3월~현재 : 고려대학교 정보보호대학원 석사과정
〈관심분야〉 네트워크 보안, IDS, VPN

노 태 우 (Tae-Woo Noh)
2001년 2월 : 고려대학교 정보공학과 졸업
2001년 3월~현재 : 고려대학교 정보보호대학원 석사과정
〈관심분야〉 네트워크 보안, 침입탐지 시스템, Water Marking

문 종 섭 (Jong-Sub Moon)
1981년 2월 : 서울대학교 계산통계학과 졸업
1983년 2월 : 서울대학원 계산통계학과 석사
1992년 2월 : Ilinois Institute of Technology 박사
1993~현재 고려대학교 전자 및 정보공학부 교수 , 고려대학교 정보보호대학원 교수
〈관심분야〉 IDS, 신경망, 생체인식, 운영체제



고 재 영 (Jae-Young Koh)
1984년 2월 : 전북대학교 전자공학과 졸업
1992년 2월 : 전북대학교 전자공학과 석사
1998년 2월 : 전북대학교 전자공학과 박사
1984년~2000년 : 국방과학연구소 선임연구원 팀장
2000년~현재 : 국가보안기술연구소 책임연구원 팀장
〈관심분야〉 네트워크 보안, VPN, 방화벽 시스템



최 대 식 (DaeSik Choi)
1997년 2월 : 강원대학교 전자계산학과 졸업
1999년 2월 : 강원대학교 전자계산학과 석사
2000년~현재 : 국가보안기술연구소 연구원
〈관심분야〉 네트워크 생존성, 분산시스템, 알고리즘

한 광 택 (KwangTaek Han)
1998년 2월 : 고려대학교 전자계산학과 졸업
2001년 2월 : 고려대학교 전자계산학과 석사
2000년~현재 : 국가보안기술연구소 연구원
〈관심분야〉 네트워크 보안, IDS, PKI, 데이터마이닝