Adaptive Resonance Theory2를 이용한 침입탐지 시스템

박현철,노태우,서재수,박일곤,김진원,문종섭[1],한광택,최대식,고재영[2]

(1)고려대학교 정보보호 기술연구 센터

(2)국가보안기술연구소

IDS System Using Adaptive Resonance Theory2

HyunCheal Park, TaeWoo Noh, JaeSu Seo, IlGon Park, JinWon Kim, JongSub Moon[1], KwangTaek Han, DaeSik Choi, JaeYoung Koh[2]

(1)Centet for Information Security Technologies, KOREA Univ.

(2)National Security Research Institute

요 약

본 논문은 신경망 이론중 하나인 Adaptive Resonance Theory(ART)을 사용하여 네트워크 상의 불법적인 침입을 탐지하는 기법에 대한 연구이다. ART는 비교사 학습을 하는 신경망으로써, 적응적인 학습능력이 있으며, 또 새로운 패턴에 대해서 새로운 클러스터를 생산하는 능력이 있다. ART의 이러한 특성을 이용하여, 여러 가지 침입패턴을 네트워크상에서 생산하여 학습을 시키고, 또 test 했으며, test 이후에도 on-line 상에서 새로운 공격 pattern도 찾아냄을 보였다. 따라서, 이미 알려진 침입뿐만 아니라 새롭게 발생하는 침입 기법에 대해서도 새로운 rule의 첨가 없이 적극적으로 대처할 수 있을 것으로 예측된다.

1. 서론

최근 네트워크로 연결된 시스템의 가용성, 기밀성, 무결성 등을 해치는 침입을 탐지하기 위해 많은 연구가 진행 중에 있다[1]. 이러한 연구들 중 대부분은, 미리 알려진 침입방법을 규칙 또는 특정패턴으로 정의해 놓고 온라인상의 침입에 대해서 해당 패턴과의 일치정도에 따라 침입을 탐지하고 있다. 따라서 알려진 공격에 대해서는 정확성이 있으나 알려지지 않은 공격행위나 비정상 행위에 대한 탐지여부는 미약한 상황이다[4,5,7,13]. 따라서 이러한 방법을 사용하는 시스템들은 새로운 패턴이 나타날 때마다 새로운 규칙이나 패턴을 등록시켜야 한다. 따라서 일반성이 그리 많지 않다.

이러한 문제점을 개선하기 위하여 본 연구는 정상행위의 학습을 통하여 알려지지 않는 침입과 비정상행위를 찾아내려 한다. 본 연구에서는 신경망을 이용해 정상행위를 학습시킴으로써 알려지지 않은 침입과 비정상행위를 찾아냄으로써 침입탐지 시스템의 탐지율을 향상시키고, 오판을 최소화 할 수 있는 자동화된 시스템 개발에 기본적인 목표를 두고 있다.

본 논문은 국가보안기술연구소의 지원으로 연구되었다.

1. Intrusion Detection System

침입탐지 시스템은 탐지 대상 데이터에 따라 크게 시스템기반과 네트워크 기반 침입탐지 시스템으로 나뉘어 진다. 시스템기반 침입탐지시스템은 많은 연구가 진행되고 있으나, 네트워크기반 침입탐지시스템은 미리 침입에 대한 정의를

기반으로 탐지여부를 결정하는 오용침입탐지시스템 연구에서 크게 벗어나지 못하고 있고 알려지지 않은 공격이나 비정상행위에 대한 탐지시스템에 대해서는 아직 초기 단계에 있는 상태이다. 기존의 비정상행위 판정시스템의 연구를 살펴보면 미국의 SRI (Stanford Research Institute International)의 경우 1980년대부터 현재까지 침입탐지 프로젝트를 수행하고 있다. 초기의 연구인 IDES를 기반으로 하여 NIDES (Next-generation IDES)라는 침입탐지시스템을 개발하였다[9]. 이는 시스템 기반의 침입탐지시스템으로써 통계학적인 비정상행위 탐지 및 침입행위의 특성을 규칙으로 나타낼 수 있는 P-BEST (Production-Based Expert System Toolset)[6,8] 시그니쳐 분석도구를 결합하여 높은 수준의 침입탐지시스템을 개발하였다. NIDES의 후속 시스템으로서 DARPA 프로젝트의 일환인 EMERALD(Event Monitoring Enabling Response to Anomalous Live Disturbance)가 있다[7]. 이는 네트워크 기반의 분석과 상호 연동성을 증가시키고 분산 컴퓨팅 환경으로의 통합을 편하게 하기 위하여 NIDES를 확장시킨 것이다. 이러한 노력은 프로파일을 이용한 분석, 시그니쳐를 근간으로 한 분석, 지역화된 결과의 결합에 대한 방법론 등을 통해 이루어지고 있다. 그 외 비슷한 시스템으로는, 실시간 네트워크 침입탐지 시스템, 멀티호스트 기반 침입탐지 시스템 등이 있다.

2. Adaptive Resonance Theory

침입탐지 시스템에 사용될 여러 알고리즘 중, 본 연구는 Adaptive Resonance Theory(ART)라는 신경망 모델을 사용한다.[2]
Carpenter와 Grossburg에 의해 제안된 ART는 기존에 학습되었던 것이 새로운 학습에 의해 지워지지 않도록, 새로운 지식을 전체 데이터베이스에 일관성 있는(self-consistent) 방법으로 통합한다. 실시간 클러스터링 알고리즘으로 알려진 ART는 다른 신경망에 비해 다음과 같은 장점을 갖는다. 첫째, ART는 비교사 학습(unsupervised learning)에 의해 입력 패턴을 클러스터링 하기 때문에 사전에 교사데이터 없이 새로운 입력 패턴을 학습할 수 있다. 둘째, ART는 기존 신경망들의 문제점인 Stability-Plasticity 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. ART는 입력패턴과 학습된 클러스터간의 비교를 통해, 이미 학습된 클러스터에 영향을 미치지 않으면서 학습을 수행할 수 있는 Reset 메커니즘을 사용하여 이 문제점을 해결하였다. 이러한 능력에 의해 적응적이고 증가적인 학습을 수행한다. 셋째, ART는 경계 변수(vigilance parameter)값에 따라 클러스터링의 분류 결과를 조정할 수 있다. 즉, 경계변수의 값을 크게 주면, 좀더 세분화되고 구체적인 클러스터들을 얻을 수 있다. 또한 학습이 완료된 클러스터에 대한 가중치 값들은 해당 클러스터에 속해 있는 패턴들에 대한 대표 벡터로 해석될 수 있다.
ART에는 이진 벡터를 클러스터링 하는 ART1과 아날로그 벡터를 클러스터링 하는 ART2가 있으며, Fuzzy ART는 ART1의 product 연산을 퍼지 집합 이론의 min연산으로 대치함으로써 ART1이 아날로그 벡터에 대하여 학습할 수 있도록 한 것이다. 본 연구는 여러 ART 모델 중 ART2 모델을 침입탐지시스템에 적용시키려 한다[2,3].

그림 1 과 그림 2 는 ART 알고리즘의 flow chart와 시스템의 구조를 나타낸다.

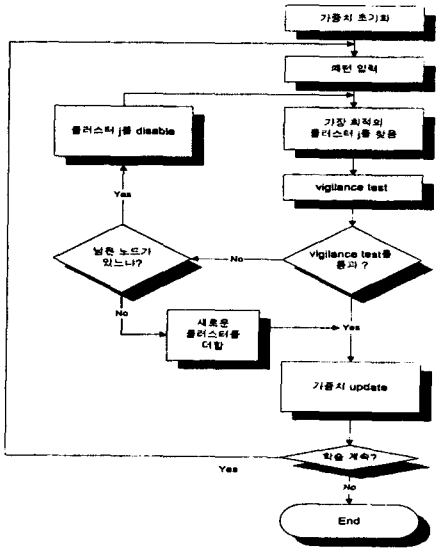


그림 1: ART 알고리즘 Flow chart

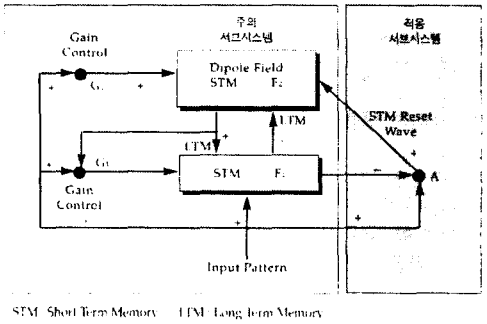


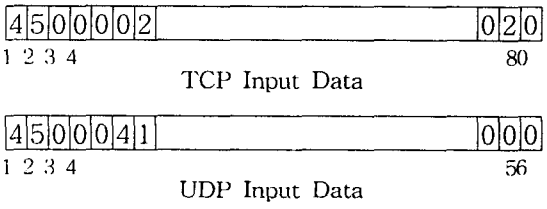
그림 2: ART 시스템의 구조[2]

II. 본문

1. 실험환경

10Mbps 이더넷 환경에서 수집한 패킷을 분석하여 ART신경망의 입력값으로 사용하였다. 네트워크 트래픽은 매 시간에 따라 다르게 변하기 때문에 샘플 데이터를 만들 때, 네트워크 트래픽 상황을 고려하여 표 1과 같이 네트워크 패킷을 TCPdump라는 tool을 사용하여 시간별로 수집하였고, 공격 패킷은, 현재 많이 사용되고 있는 DOS attack tool을 프로토콜(TCP, UDP)마다 각각 2가지씩 총 4가지 tool을 사용하여 수집하였다. 이러한 데이터는 연구실 환경에서 독립망을 구성하여 수집하여 정확도를 높였다. 여기서 얻은 패킷은 가공되지 않은 데이터이기 때문에, 이 데이터를 신경망의 입력값으로 사용하기 위해서는, 약간의 전처리 과정이 필요하다. 각각의 한 패킷에서 TCP 프로토콜은 80가지 종류를, UDP 프로토콜은 56가지 종류의 정보를 추출<그림 3>하여 각각 별도로 ART신경망의 입력값으로 사용하였다. 또한 결과의 정확도를 높이기 위해 두 종류의 프로토콜을 따로 학습시켰다.

그림 3: ART Input Pattern



2. 실험

1)TCP

• 학습데이터

시간대 별로 수집한 정상 패킷 10000개와 서로 다른 2종류의 attack tool을 사용하여 만든 공격 패킷 1000개를 생산하여, 총 12000개의 서로 다른 패킷을 섞어서 learning data set을 만들었다.

• 시험데이터

학습데이터에서 사용하지 않은 3종류의 attack tool을 사용하여, 공격 패킷을 각 tool 마다 200개씩 생성해 냈고, 정상 패킷 역시 학습데이터와는 다른 1000개의 패킷을 사용, 총 1600개의 test data set을 만들었다.

2)UDP

TCP 데이터와 같은 방법으로 learning data set 과 test data set을 만들었다.

표 1: 학습데이터, 시험데이터 생성

샘플채취 시간	Protocol	Description
05-06	TCP,UDP	Syn-Flooding,Udp-Flooding
09-10	TCP,UDP	Syn-Flooding,Udp-Flooding
14-15	TCP,UDP	Syn-Flooding,Udp-Flooding
19-20	TCP,UDP	Syn-Flooding,Udp-Flooding
24-01	TCP,UDP	Syn-Flooding,Udp-Flooding

3. 실험결과

학습데이터에 대한 결과는 다음과 같다. TCP 프로토콜의 경우 0부터 11까지 총 12개의 클러스터가 생성되었다. 그 중 정상패킷은 0-9까지 총 10 종류의 클러스터로 수렴하였고, 공격 패킷은 공격 tool에 따라 10, 11 두 종류의 클러스터로 수렴하였다. UDP 프로토콜 역시 정상 패킷은 0-5 까지 6 종류의 클러스터로, 공격 패킷은 6, 7 두 가지 공격 tool에 따라 6, 7 클러스터로 수렴하였다.

시험데이터에 대한 공격 패킷의 결과는 다음과 같다. TCP 패킷의 경우 학습데이터에서 생성된 클러스터 이외에 새로운 공격 형태에 맞추어 새로운 클러스터 12가 생성된 것을 볼 수 있다.

UDP 패킷의 경우 역시 새로운 클러스터 8, 9가 생성된 것을 볼 수 있다. 또한 세 번째 attack tool을 사용하여 얻은 패킷은 6번(14.8%)과 9번(85.2%) 클러스터로 수렴하였다. 총 600개의 시험데이터 중 정상 클러스터로 구별되어진 패킷은 단 하나도 없이 100% 공격 패킷을 구별하였다.

또한 총 2000개의 정상 패킷은 다음과 같은 결과를 얻을 수 있었다. 여기서 새로 생성된 클러스터는 무조건 공격으로 간주하였다. 정상 패킷의 경우 정상 패킷을 올바르게 구별한 경우, 공격으로 구별한 경우, 새로운 클러스터를 생성한 경우로 나누어 보았다. TCP 패킷의 경우 8.7%, UDP 패킷의 경우 11.9% 의 오차를 보였다. 그러나 결과를 모니터 한 후 새로 생성된 클러스터(TCP:13-14, UDP:10-12)를 정상패킷으로 정의한다면 오차는 TCP 패킷의 경우 1.8%, UDP 패킷의 경우 0.8%가 된다. 시험데이터의 결과는 표 4와 같다.

표 2: 학습데이터 결과

	TCP	UDP
정상 클러스터	0,1,2,3,4,5,6,7,8,9	0,1,2,3,4,5
공격 클러스터	10,11	6,7

표 3: 시험데이터(공격 패킷)

TCP		UDP	
학습데이터에	10	학습데이터에	6
사용된 tool	11	사용된 tool	7
attack 1	11	attack 1	8
attack 2	12	attack 2	6
attack 3	10	attack3	6, 9

표 4: 시험데이터(정상 패킷)

TCP		UDP	
0-9	91.3 %	0-5	88.1 %
10-12	1.8 %	6-9	0.8 %
13-14	6.9 %	10-12	11.1 %

III. 결론

1. 결론 및 향후 연구 과제

본 연구는 침입탐지시스템에 신경망을 적용할 수 있는지에 대한 가능성을 보여 줬다는 데에 의의를 둘 수 있다. 표 3을 보면 알 수 있듯이 ART학습의 특성인 adaptive learning 과 incremental learning 과정이 잘 나타나 있다. learning 되지 않은 새로운 공격(attack 1, 2, 3)에 대하여 정상 패킷으로 구별하지 않고 새로운 클러스터를 만들거나 기존 공격 클러스터에 포함시키는 것을 볼 수 있다. 표 4를 보면 정상 패킷을 새로운 클러스터를 만들어서 구별하였는데 이것은 충분히 많은 정상 패킷과 공격 패킷을 가지고 학습시킨다면 극복할 수 있을 것이다. 즉, 본 연구는 실험실 환경에서 제한된 데이터를 가지고 실험을 해서 위와 같은 연구를 얻지만 좀더 많은 양의 데이터를 가지고 실험을 한다면 위 결과보다 오차율을 줄일 수 있을 것이다.

본 연구에서 입력값으로 사용한 정보는 TCP의 경우 한 패킷마다 80가지 종류를 UDP의 경우 한 패킷마다 56가지의 종류를 추출해 내서 사용하였는데 입력값의 벡터의 수가 너무 많다. 각각의 벡터를 살펴보면 ART가 클러스터링할 때 결정적으로 작용하는 벡터도 있을 것이고, 클러스터링할 때 거의 영향을 주지 않는 벡터도 있을 것이다. 본 연구에서는 일단 패킷에서 추출해 낸 모든 벡터를 사용하였는데 클러스터링에 영향을 주지 않는 벡터를 제거하여 클러스터링 실행 시간을 단축시킬 수 있을 것으로 생각된다. 또한 본 연구는 실험실 환경에서 생성해 낸 데이터를 가지고 학습과 시험을 하였는데 보다 높은 신뢰도를 얻기 위하여 MIT Logdata와 같은 공인된 데이터 값으로 실험하여 결과를 얻는 작업도 병행하여야 한다.

2. 참고문헌

[1]신대철, 이보경, 유동영, 김홍근 "네트워크 비정상행위 탐지를 위한 클러스터링 모델", 제 13회 정보보호와 암호에 관한 학술대회, pp.187-201, Sept 2001.

[2]Gail A. Carpenter and Stephen Grossberg, A Massively Parallel Architecture for a Self-organizing Neural networks, IEEE Computers, pp. 77-88, March 1998.

[3]Gail A. carpenter and Stephen Grossberg, "ART 2: Self-organization of stable category recognition codes for analog input patterns," Applied Optics, Vol. 26, pp.4919-4930, December 1987.

[4]Jackson, K. A. NADIR: A Prototype System for Detectiong network and File System Abuse. In Proceedings of the 7th European

Congerence on Information Systems, Nov. 1992

[5]Jackson, K., Dubois, D. H., Stallings, C. A. An expert system application for network intrusion detection. In Proceedings of the 14th National Computer Security Conference, pp. 215-225, Oct.1991.

[6]P.G. Neumann and P. A. Porras, "Experience with emerald to date," 1st USENIX Workshop on IDS, Santa Clara, Cal, 11-12 April 1999.

[7]Porras, A. and Neumann, P. G. EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances. In Proceedings of the National Information Systems Security Conference, Oct 1997.

[8]J. Frank, "Machine learning and intrusion detection : Current and future directions, "Proc. 17th National Computer Security Conference, Oct 1994

[9]D. Anderson, T.frivold and A. Valdes, "Next-generation intrusion detection expert system(NIDES)," Technical Report SRI-CLS-95-07, May, 1995

[10]Wenke Lee, Salvatore J.Stolfo "Data Mining Approaches for Intrusion Detection" Computer Science Department Columbia University 500 West 120th Street, New York, NY10027

[11]Martin Ester, Hans-Peter Kriegel, Sander, Michael Wimmer, Xiaowei Xu, "Incremental Clustering for Mining in a data Warehousing Environment", Proceedings of the 24th VLDB Conference, New York, USA, 1998.