주성분 분석과 서포트 벡터 머신을 이용한 침입탐지 시스템

정성윤*, 강병두*, 김상균**
인제대학교 전산학과*
인제대학교 컴퓨터공학부**

An Intrusion Detection System Using Principle Component Analysis and Support Vector Machines

Sung-Yoon Jung*, Byung-Doo Kang*, Sang-Kyoon Kim**
Dept. of Computer Science, InJe University*
Dept. of Computer Engineering, InJe University**

요 약

기존의 침입탐지 시스템에서는 오용탐지모델이 널리 사용되고 있다. 이 모델은 낮은 오판율(False Alarm rates)을 가지고 있으나, 새로운 공격에 대해 전문가시스템(Expert Systems)에 의한 규칙추가를 필요로 한다. 그리고 그 규칙과 완전히 일치되는 시그너처만 공격으로 탐지하므로 변형된 공격을 탐지하지 못한다는 문제점을 가지고 있다. 본 논문에서는 이러한 문제점을 보완하기 위해 주성분분석(Principle Component Analysis; 이하 PCA)과 서포트 벡터 머신(Support Vector Machines; 이하 SVM)을 이용한 침입탐지 시스템을 제안한다. 네트워크 상의 패킷은 PCA를 이용하여 결정된 주성분 공간에서 해석되고, 정상적인 흐름과 비정상적인 흐름에 대한 패킷이미지패턴으로 정규화 된다. 이러한 두 가지 클래스에 대한 SVM 분류기를 구현한다. 개발하는 침입탐지 시스템은 알려진 다양한 침입유형뿐만 아니라, 새로운 변종에 대해서도 분류기의 유연한 반응을 통하여 효과적으로 탐지할 수 있다.

1. 서론

정보통신서비스가 급속히 발전함에 따라 해킹이나 웜 바이러스, 대규모 서비스 거부공격과 같은 네트워크상의 범죄가 급증하고있다. 이에 대한 대책으로 방화벽이나 침입탐지 시스템(Intrusion Detection System; 이하 IDS)과 같은 기술이 보급되고 있다[1].

IDS는 알려져 있거나 잠재적 위협으로부터 네트워크를 보호하는데 목적이 있다. 네트워크 기반 IDS 모델에는 비정상 침입탐지 기법(Anomaly Detection)과 오용 침입탐지 기법(Misuse Detection)이 있다[2]. 비정상 침입탐지 기법은 사용자의 패턴을 분석하여 입력패턴과 비교하여 정해진 모델을 벗어나는 경우를 침입으로 탐지한다. 그리고 알려진 침입행위를 이용하여 침입을 탐지하는 오용탐지 기법은 낮은 오판율(False alarm rates) 때문에 가장 널리 사용되고 있다[3]. 이 기법은 대부분 알려진 공격탐지를 위해 규칙기반 시스템을 사용한다. 이 시스템은 새로운 공격 발생 시 전문가 시스템(Expert Systems)에 의해 규칙이 추가되어야 한다. 그리고 그 규칙과 완전히 매칭되는 시그너처만을 공격으로 탐지하기 때문에 변형 또는 우회 공격에 대한 유연성이 없다. 기존의 IDS는 스니핑(Sniffing)속도의 한계로 인해 네트워크 상의 모든 패킷을 수집하고 분석하는 것이 불가능하다[4]. 따라서 새로운 변종의 공격뿐만 아니라 패킷 일부분만으로 공격을 탐지할 수 있는 연구가 필요하다. 본 논문에서는 이러한 문제점을 해결하기 위해 PCA와 SVM을 이용한 침입탐지 시스템을 제안한다.

본 논문에서 제안하는 침입탐지 시스템은 입력 값으로 패킷 헤더의 특정 값이나 감사자료를 통한 접근방식과는 달리 PCA를 통한 패킷이미지패턴을 사용함으로써 더욱 광범위한 공격 유형을 탐지할 수 있다. 이러한 이유로 다변량 통계 분석(Multivariate Statistical Analysis)방법 중 하나인 PCA를 이용한다. 패킷정보를 최대한 설명할 수 있는 독립적인 인공변수(Artificial Variable)들을 유도하여 주성분을 구한다. 이러한 주성분은 일련의 패킷정보들의 선형결합으로 표시되며, 단순한 구조로 요약되는 패킷이미지패턴을 만든다. 정상적인 흐름과 비정상적인 흐름에 대한 패킷이미지 패턴을 학습하는 SVM 분류기를 구현한다. SVM은 두가지 클래스를 효율적으로 분류하는 방법으로 분류 및 인식 능력이 뛰어나 최근 다양한 패턴인식 분야에서 적용되는 알고리즘이다[5].

2. 침입탐지시스템의 전체구성

본 논문에서 제안하는 PCA와 SVM을 이용한 침입탐지

시스템의 전체 구성은 그림 1과 같다.

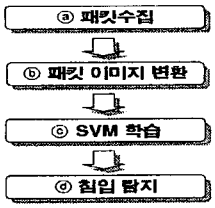

그림 1. IDS의 구성

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

ⓑ PCA를 통해 패킷이미지패턴을 만든다.

ⓒ 패킷이미지패턴을 정규화하고 SVM 학습을 위해 사용된다. 학습이 완료되면 서포트 벡터(Support Vector; 이하 SV)가 생성된다.

ⓓ 생성된 SV로 실시간 네트워크 침입탐지를 하게되고, 그 결과를 정상과 비정상으로 구별한다. 패킷을 연속한 패킷이미지패턴으로 처리함으로써 공격 이미지 일부분만을 감지하더라도 이와 가장 유사한 패턴으로 구별해 낼 수 있다.

2.1 패킷수집

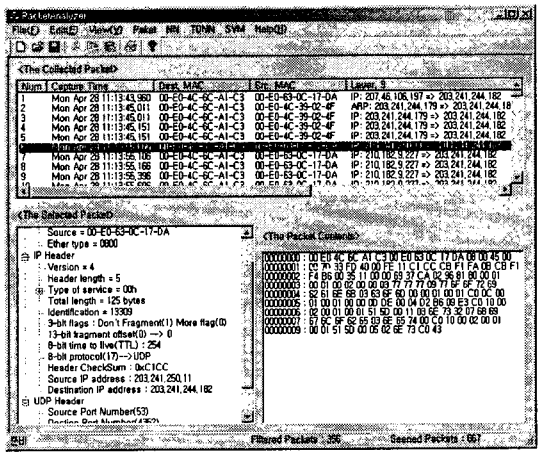

그림 2. 패킷 수집 과정

패킷 수집을 위해 libpcap 라이브러리를 사용한다. 이것은 Berkeley 대학에서 개발한 것으로 시스템 독립적으로 사용자 레벨에서 개발한 패킷 수집을 효과적으로 할 수 있도록 만든 공용 라이브러리이다. 패킷을 수집하려는 시스템에서 무작위 모드(Promiscuous mode)의 상위수준 인터페이스를 제공한다[6]. 네트워크 패킷은 libpcap 라이브러리에서 제공하는 인터페이스를 사용하여 접근할 수 있다. 이 라이브러리는 거의 모든 유닉스 시스템에서 사용가능하며 tcpdump와 같은 네트워크 모니터링 도구가 이 패킷 수집 라이브러리를 사용하고 있다[6,7,8,9].

정상적인 흐름과 비정상적인 흐름에 대한 패킷 패턴을 학습하기 위해 한 대의 공격 시스템, 다섯 대의 정상사용자 시스템 그리고 패킷 수집 시스템으로 실시간 패킷 수집을 한다. 그리고 PCA를 통하여 패킷이미지패턴을 만든다. 그림 2의 상단부분은 수집된 패킷의 정보로써 수집 시간과 패킷헤더의 정보를 나타낸다. 그리고 하단부분 중 왼쪽은 선택한 패킷의 세부 정보를 보여주고, 오른쪽은 선택한 패킷의 내용을 16진수의 바이트 코드 값으로 나타낸다.

2.2 주성분분석 (Principle Component Analysis)

PCA는 여러 개의 변수들 사이의 관계를 분석하여 이 변수들의 선형결합으로 표시되는 새로운 주성분(Principal Components)을 찾고, 이 중에서 중요한 몇 개의 주성분으로 전체변동을 설명하고자 하는 다변량 통계 분석 방법이다. PCA는 자료의 요약이나 선형관계식을 통하여 차원(Dimension)을 감소시켜 해석을 용이하게 하는데 목적이 있다. 자료가 갖고 있는 전체적인 변동의 대부분을 원래변수 p개보다 적은 m개의 주성분으로 설명할 수 있다고 하면, p개의 변수가 갖고 있는 정보의 대부분을 m개의 주성분으로 대체할 수 있다. 이렇게 함으로써 변수의 차원을 감소시킬 수 있으며, 상관관계가 있는 변수들의 경향이나 변동을 몇 개의 주성분으로 파악할 수 있게 한다[10].

2.3 패킷이미지패턴 변환

각 주성분에 의하여 설명되어지는 전체 패킷의 분산에 대한 누적설명비율이 80% 이상인 것으로 주성분 개수를 결정하게 된다. 따라서 하나의 패킷을 크게 설명하는 주성분 여덟 개를 선정하여 순차적으로 60개를 모아 패킷이미지패턴을 만든다. 그림 3은 수집한 패킷을 PCA를 통해 패킷이미지패턴을 만드는 과정을 보여주고 있다.

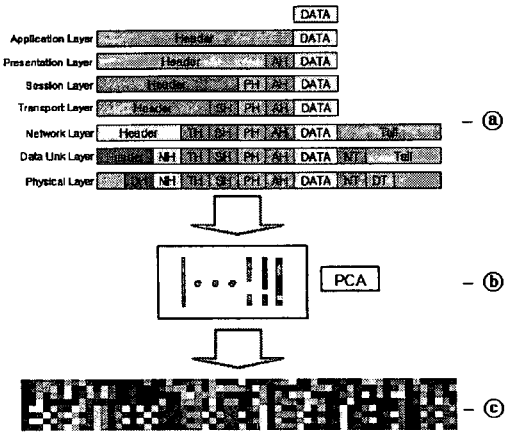

그림 3. 패킷의 학습패턴 변환

ⓐ 패킷 수집 시스템으로 패킷을 수집한다.
ⓑ PCA를 사용하여 패킷을 가장 잘 설명하는 주성분 여덟 개를 선정하고 순차적으로 60개를 모아 패킷이미지패턴을 만든다
ⓒ 연속된 패킷이미지패턴을 SVM의 입력으로 사용한다.

패킷이미지패턴을 만들면 그림 3의 ⓒ와 같은 Spectrogram을 생성한다. 세로줄은 하나의 패킷을 설명하는 주성분들을 나타내고 그것을 순차적으로 60개(time step)를 본 것이다.

3. 서포트 벡터 머신 (SVM)

SVM은 1995년에 Vapnik에 의해 제안되었고 VC(Vapnik-Chervonenkis)이론에 근간을 두고 있으며, 뛰어난 일반화 성능을 보여준다[11,12]. 구조적 에러를 최소화하는 기법으로 기존의 경험적 에러 최소화 기법인 다층신경망과 비교하여 학습에 필요한 파라미터의 일부가 자동적으로 결정된다.

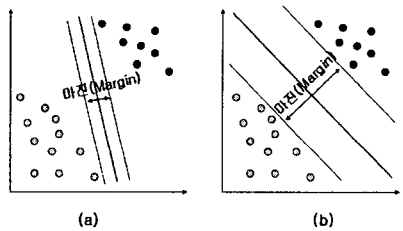

그림 4. 선형분리 가능한 경우

선형분리 가능한 이진 클래스의 경우, 입력값들을 다른 클래스간에 데이터 사이를 최대거리로 분리할 수 있는 초평면(Hyperplane)을 찾기 위해서 고차원의 특징공간(Feature space)로 변환시킨다. 두 클래스 군집을 선형 분리하는 초평면과 가장 가까운 점을 'Support Vector(SV)'라고 한다. SV와 결정 평면간의 거리를 '마진(Margin)'이라고 한다. SVM은 마진을 최대화하는 최적의 초평면을 찾는다. 그림 4는 선형분리 가능한 데이터에 대한 초평면의 예를 보여준다. (a)는 마진의 거리가 작은 경우이고, (b)는 마진의 거리가 최대화되는 경우이다.

최대의 마진을 가지는 초평면을 구하기 위해 식(1)의 최소값을 구해야 한다. 동시에 데이터를 분류하기 위한 조건식(2)를 만족하여야 한다.

$$\tau(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2 \qquad (1)$$

$$y_i \cdot ((\boldsymbol{w} \cdot x_i) + b) \geq 1, \quad i = 1, \cdots, l \qquad (2)$$

식(2)의 제약조건을 만족하면서 식(1)의 최소값을 구하는 문제는 라그랑즈 승수(Lagrange Multiplier) α를 사용한 식(3)으로 표현된다.

$$L(\boldsymbol{w}, b, \alpha) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{l} \alpha_i (y_i \cdot ((\boldsymbol{x}_i \cdot \boldsymbol{w}) + b) - 1) \qquad (3)$$

선형적으로 분리 가능하지 않은 경우, 커널함수(Kernel function)를 사용하여 특징공간(Feature space)에서 선형분리를 수행한다.

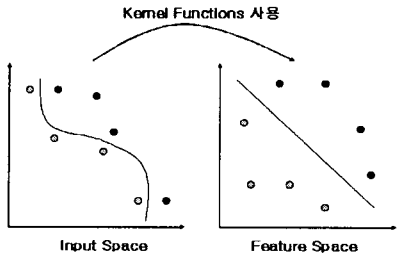

그림 5. 선형분리 가능하지 않은 경우

그림 5는 선형분리 가능하지 않은 데이터에 대한 초평면의 예를 보여준다. 이 데이터를 커널함수를 사용하여 특징공간(Feature space)으로 이동한다. 본 논문에서는 PCA를 사용하여 패킷이미지패턴으로 정규화한다. 이러한 패킷이미지패턴은 선형분리가 가능하므로 SVM의 커널함수를 사용하지 않았다.

4. 실험

제안한 PCA와 SVM을 이용한 침입탐지 시스템은 Pentium-4 PC, Windows 환경에서 Visual C++로 구현하였다. 제안한 시스템의 성능을 평가하기 위해 수집된 각 패킷 데이터 중 학습에 사용된 데이터와 학습에 사용되지 않은 데이터로 구분하여 실험하였다. SYN Flooding, Land, TearDrop, New TearDrop공격 각각에 대한 공격패킷 6000개, 정상패킷 6000개를 수집하여 랜덤하게 섞었다. 그리고 PCA를 통해 100개의 비정상 패킷이미지패턴과 100개의 정상 패킷이미지패턴을 만들어 학습하였다.

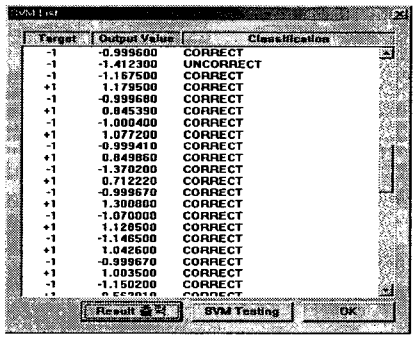

그림 6. 변종공격 테스트

TearDrop을 학습한 SVM으로 변종공격인 New TearDrop을 테스트 한 결과는 그림 6과 같다. 'Target'값은 데이터의 출력 값으로 '-1'인 것은 정상 데이터이고, '1'인 것은 비정상 데이터이다. 'Classification' 항목은 SVM이 구분해낸 결과로

'CORRECT'는 제대로 탐지한 것이고, 'UNCORRECT'는 제대로 탐지하지 못한 것이다.

그림 7은 그림 6에서 얻은 실험 결과를 그래프로 나타낸 것이다. 하단에 분포되어 있는 "●"표시는 정상패킷, 상단에 분포되어 있는 "■"표시는 비정상패킷을 말한다.

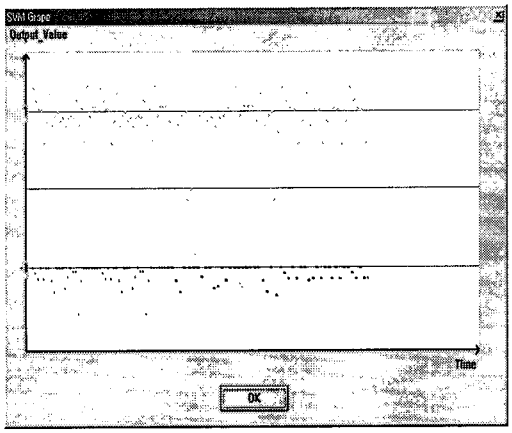

그림 7. 변종공격 결과 그래프

표 1은 제안한 시스템의 실험결과를 나타낸다. SYN Flooding, Land, TearDrop, New TearDrop공격은 정상을 비정상으로 탐지(False Positive)하거나 비정상을 정상으로 탐지(False Negative)하는 것 없이 정확하게 탐지해냈다. 뿐만 아니라, 변종공격 테스트(학습: TearDrop, 테스트: New TearDrop)에서는 비정상 패킷 이미지 4개를 제외하고는 모두 탐지해 내었다.

표 1. 제안된 시스템의 실험 결과

	False Positive	False Negative
SYN Flooding	0	0
Land	0	0
TearDrop	0	0
New TearDrop	0	0
변종공격 테스트	0	4

5. 결론

본 논문에서는 다변량 통계 분석(Multivariate Statistical Analysis)방법인 PCA와 통계적 학습 이론(Statistical Learning Theory)방법인 SVM을 이용한 침입탐지 시스템을 제안하였다.

오분류(Misclassification)를 유도하는 불순한 요소를 제거하기 위해, PCA를 사용하여 패킷정보를 여덟 개의 주성분으로 구성된 8차원으로 축약하였다. 그리고 시간적 요소를 부가하기 위해, 주성분 값들을 순차적으로 60개를 모아 패킷이미지패턴으로 정규화하였다. 마지막으로 이러한 패킷이미지패턴을 SVM의 입력값으로 사용하였다.

실험 결과에 따르면, 제안한 시스템은 학습된 패킷 이미지를 정확히 분류해 내었고, 기존의 학습된 시스템으로 학습시키지 않은 변형된 패킷 이미지를 잘 분류해 내었다. 이는 크래커들이 변형 또는 우회하는 공격을 하더라도 탐지해 낼 수 있으므로 탐지 시스템의 성능을 높일 수 있다. 그러므로 제안한 시스템은 기존의 오용탐지모델과 상호 연동 시 보다 높은 성능의 탐지 시스템이 될 것이라 기대된다.

참 고 문 헌

[1] J. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Journal of Data Mining and Knowledge Discovery, Vol. 2, No. 2, pp.121-167, 1998

[2] R. Seker, A High-Performance Network Intrusion Detection System, ACM ISBN: 1-58113-148-8 pp. 8-17 Oct. 1999

[3] Vern Paxson, Bro: A System for Detecting Network Intruders in Real-time, Computer Networks, 31(23-24), pp. 2435-2463, 14 Dec. 1999.

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

[5] Giorgio Giacinto, Fabio Roli, Luca Didaci, Fusion of multiple classifiers for intrusion detection in computer networks, Pattern Recog. Lett. 2003

[6] 이장헌, 신경회로망을 이용한 비정상적인 패킷 탐지, 정보보호학회, vol.11 no.5 pp. 105-117, october 2001.

[7] 포항공대 유닉스 보안 연구회, Security PLUS for UNIX, 영진출판사, ISBN: 89-314-1490-0, pp. 251-254, pp. 383-400, 2001.

[8] S. McCanne and V. Jacobson, The BSD Packet Filter: A New Architechure for User-level Packet Capture, USENIX conference, January pp. 25-29, 1993, San Diego, CA.the 1993 Winter

[9] Bob Quinn, Dave Shute, Windows Sockets Network Programming, Addison Wesley Publishing Company, ISBN 0-201-63372-8, 1996.

[10] Richard A. Johnson, Dean W. Wichern, Applied Multivariate Statistical Analysis, Prentice Hall, ISBN : 0-13-092553-5, pp. 356-395, 2002.

[11] V. Vapnik, "The Nature of Statistical Learning Theory", Springer-Verlag, New York, 1995

[12] V. Vapnik, An Overview of Statistical Learning Theory, IEEE Trans. on Neural Network, Vol 10, No. 5, pp. 988-999, 1999